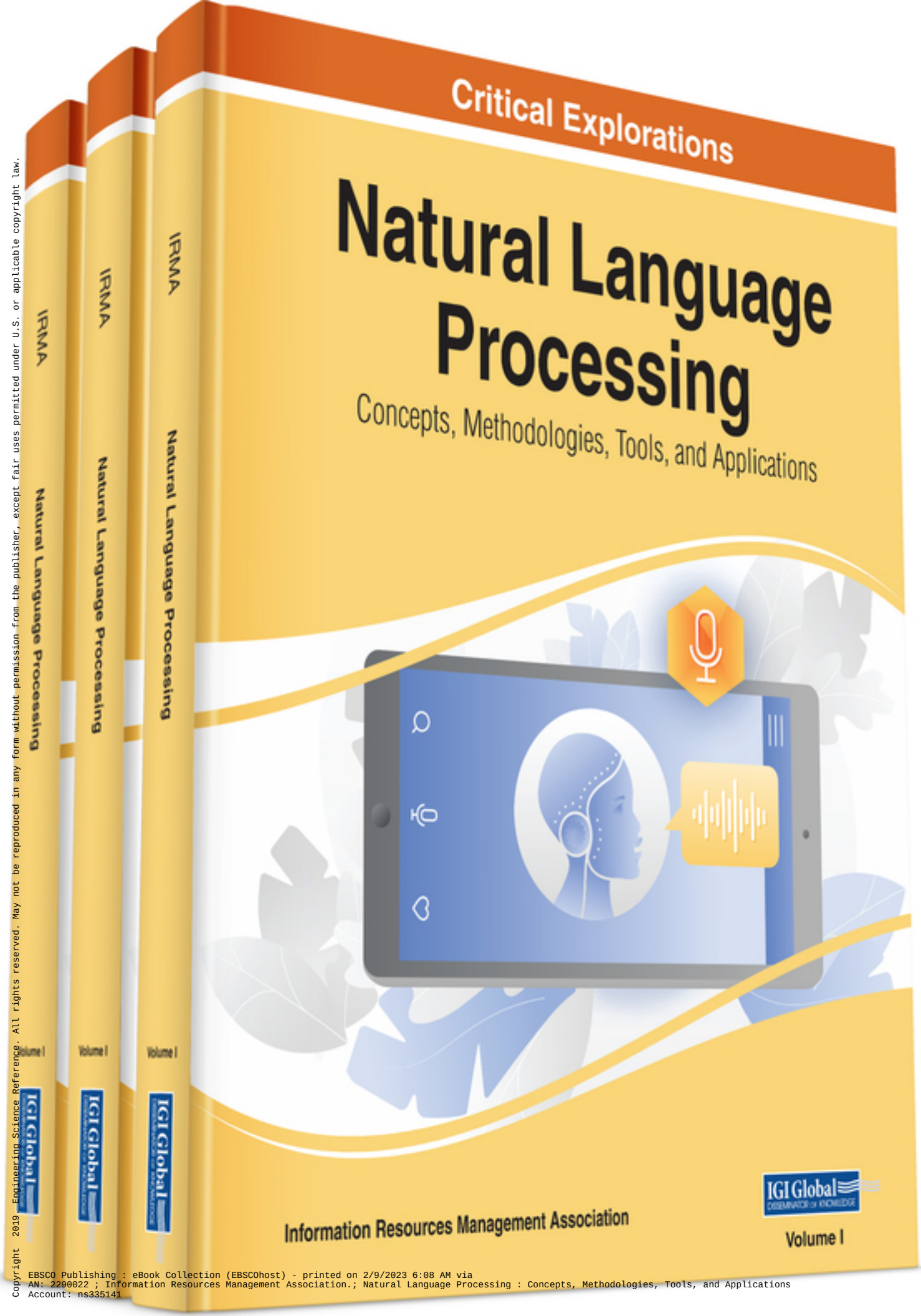


Copyright 2019 Engineering Science Reference. All rights reserved. May not be reproduced in any form without permission from the publisher, except fair uses permitted under U.S. or applicable copyright law.

EBSCO Publishing : eBook Collection (EBSCOhost) - printed on 2/9/2023 6:08 AM via
AN: 2290022 ; Information Resources Management Association.; Natural Language Processing : Concepts, Methodologies, Tools, and Applications
Account: ns335141

Natural Language Processing:

Concepts, Methodologies, Tools, and Applications

Information Resources Management Association
USA

IGI Global
DISSEMINATOR OF KNOWLEDGE

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Published in the United States of America by
IGI Global
Engineering Science Reference (an imprint of IGI Global)
701 E. Chocolate Avenue
Hershey PA, USA 17033
Tel: 717-533-8845
Fax: 717-533-8661
E-mail: cust@igi-global.com
Web site: http://www.igi-global.com

Copyright © 2020 by IGI Global. All rights reserved. No part of this publication may be reproduced, stored or distributed in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher. Product or company names used in this set are for identification purposes only. Inclusion of the names of the products or companies does not indicate a claim of ownership by IGI Global of the trademark or registered trademark.

Library of Congress Cataloging-in-Publication Data

Names: Information Resources Management Association, editor.
Title: Natural language processing : concepts, methodologies, tools, and applications / Information Resources Management Association.
Other titles: Natural language processing (Engineering Science Reference (Publisher))
Description: Hershey : Engineering Science Reference, 2019. | Includes bibliographical references and index. | Summary: “This book examines the latest concepts, processes, and techniques for communication between computers and humans”-- Provided by publisher.
Identifiers: LCCN 2019023675 (print) | LCCN 2019023676 (ebook) | ISBN 9781799809517 (hardcover) | ISBN 9781799809524 (ebook)
Subjects: LCSH: Natural language processing (Computer science) | Computational linguistics. | Human-computer interaction--Industrial applications.
Classification: LCC QA76.9.N38 N3835 2019 (print) | LCC QA76.9.N38 (ebook) | DDC 006.3/5--dc23
LC record available at https://lccn.loc.gov/2019023675
LC ebook record available at https://lccn.loc.gov/2019023676

British Cataloguing in Publication Data
A Cataloguing in Publication record for this book is available from the British Library.

The views expressed in this book are those of the authors, but not necessarily of the publisher.

For electronic access to this publication, please contact: eresources@igi-global.com.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Editor-in-Chief

Mehdi Khosrow-Pour, DBA
Information Resources Management Association, USA

Associate Editors

Steve Clarke, *University of Hull, UK*
Murray E. Jennex, *San Diego State University, USA*
Ari-Veikko Anttiroiko, *University of Tampere, Finland*

Editorial Advisory Board

Sherif Kamel, *American University in Cairo, Egypt*
In Lee, *Western Illinois University, USA*
Jerzy Kisielnicki, *Warsaw University, Poland*
Amar Gupta, *Arizona University, USA*
Craig van Slyke, *University of Central Florida, USA*
John Wang, *Montclair State University, USA*
Vishanth Weerakkody, *Brunel University, UK*

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

List of Contributors

Aguilar, Juan Antonio Rodríguez / *Artificial Intelligence Research Institute (IIIA-CSIC), Spain* 860
Ahuja, Neelu Jyothi / *University of Petroleum and Energy Studies, India* 722
Akimoto, Taisuke / *The University of Electro-Communications, Japan* 743
Alamanou, Evangelia / *Tzaneio Hospital, Greece* 266
Almajano, Pablo / *University of Barcelona, Spain* 1408
Anderson, Dennis / *St. Francis College, USA* 1298
Armstrong, Stuart / *Future of Humanity Institute, UK* 1232
Bagula, A. B. / *University of the Western Cape, South Africa* 372
Barbora, Madhumita / *Tezpur University, India* 1139
Basaca-Preciado, Luis C. / *Center of Excellence in Innovation & Design – CETYS University, Mexico* 184
Basque, Josianne / *TELUQ University, Canada* 1178
Beaudry, Eric / *University of Québec à Montréal, Canada* 1178
Ben Kharrat, Firas / *University of Carthage, Tunisia* 638
Berman, S. / *University of Cape Town, South Africa* 372
Bhattacharyya, Pushpak / *Indian Institute of Technology Patna, India* 991
Biba, Marenglen / *University of New York in Tirana, Albania* 688
Boblan, Ivo Wilhelm / *Technische Universität Berlin, Germany* 1507
Bottrighi, Alessio / *University of Piemonte Orientale, Italy* 658
Boukchina, Eya / *University of Carthage, Tunisia* 1162
Cabri, Giacomo / *University of Modena and Reggio Emilia, Italy* 302
Callies, Sophie / *University of Québec à Montréal, Canada* 1178
Cao, Yang / *Virginia Tech, USA* 939
Caro, Manuel F. / *Universidad de Córdoba, Colombia* 983
Carrascosa, Carlos / *Universidad Politécnica de Valencia, Spain* 464
Chan, Christine / *University of Regina, Canada* 708
Chan, Jonathan T. / *University of Regina, Canada* 708
Chan, Marvin T. / *University of Regina, Canada* 708
Chen, Lin-Chih / *National Dong Hwa University, Tawian* 838
Chen, Xin / *Virginia Tech, USA* 939
Chishti, Mohammad Ahsan / *National Institute of Technology, India* 1214
Choi, Youngseok / *Brunel University London, UK* 1085
Chrelias, Charalambos / *National and Kapodistrian University of Athens, Greece* 266
Chrelias, George / *National and Kapodistrian University of Athens, Greece* 266

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Cialdella, Vincent T. / *Illinois State University, USA* 791
Claudi, Andrea / *Università Politecnica delle Marche, Italy* 1652
Coners, André / *South Westphalia University of Applied Sciences, Germany* 878
Contreras, Eduardo C. / *Autonomous University of Coahuila, Mexico* 1533
Corbett, Natalie / *University of Portsmouth, UK* 1459
Corradini, Andrea / *Design School Kolding, Denmark* 561
Cui, Pingyuan / *Beijing Institute of Technology, China* 248
Dafflon, Baudouin / *Université de Lyon, France* 232
Daimary, Surjya Kanta / *Punjabi University, India* 1139
Damaskou, Vasileia / *National and Kapodistrian University of Athens, Greece* 266
Davey, Bill / *RMIT University, Australia* 1286
Deshpande, Shrinivas P / *DCPE, HVPM, India* 422
Dey, Nilanjan / *TechnoIndia College of Technology, India* 1199
Dragoni, Aldo Franco / *Università Politecnica delle Marche, Italy* 1652
D'Rosario, Michael / *Deakin University, Australia* 49, 65, 539
Duhan, Neelam / *YMCA University of Science and Technology, India* 818
Dwivedi, Sanjay K. / *Babasaheb Bhimrao Ambedkar University (BBAU), India* 1621
Ekbal, Asif / *Indian Institute of Technology Patna, India* 991
Elkhleifi, Aymen / *Paris-Sorbonne University, France* 638
Faiz, Rim / *University of Carthage, Tunisia* 638
Fang, Youjia / *Virginia Tech, USA* 939
Fernández-Martínez, José Luis / *Institute of Advanced Social Studies (IESA-CSIC) & University of Granada, Spain* 860
Flores-Fuentes, Wendy / *Autonomous University of Baja California, Mexico* 184
Fukuda, Kiyohito / *Osaka Prefecture University, Japan* 1058
Gao, Ai / *Beijing Institute of Technology, China* 248
Gechter, Franck / *LE2I, UTBM, University of Bourgogne Franche Comté, France* 232
Gelowitz, Craig / *University of Regina, Canada* 708
Gómez, Adán A. / *Universidad de Córdoba, Colombia* 983
Gopalani, Dinesh / *MNIT Jaipur, India* 619
Goyal, Vishal / *Punjabi University, India* 1139
Gravel, Mathieu / *University of Québec à Montréal, Canada* 1178
Guériau, Maxime / *Univ Lyon, ENTPE, IFSTTA, LICIT UMR_T 9401, F-69675, France* 232
Hall Jr., Owen P. / *Pepperdine University, USA* 280
Heras, Stella / *Universidad Politécnica de Valencia, Spain* 464
Hernandez-Balbuena, Daniel / *Autonomous University of Baja California, Mexico* 184
Herpich, Fabricio / *Federal University of Rio Grande do Sul, Brazil* 912
Hkiri, Emna / *University of Monastir, Tunisia* 1686
Horne, Jeremy / *Inventors Assistance League, USA* 1564
Hsieh, Calvin / *Deakin University, Australia* 65
Isafiade, O. E. / *University of Cape Town, South Africa* 372
Issertial, Laurent / *Osaka Prefecture University, Japan* 394
Jain, Goonjan / *Jawaharlal Nehru University, India* 770
Jain, Vinay Kumar / *Jaypee University of Engineering and Technology, India* 808
Jakóbczak, Dariusz Jacek / *Technical University of Koszalin, Poland* 509, 584, 1026, 1379
Jan, Rafiya / *Central university Of Kashmir, India* 1115

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Jiang, Yuncheng / *South China Normal University, China* 1436
Jordan, J. Scott / *Illinois State University, USA* 791
Julian, Vicente / *Universidad Politécnica de Valencia, Spain* 464
Juneja, Dimple / *Dronacharya Institute of Management and Technology, India* 83
Kalles, Dimitris / *Hellenic Open University, Greece* 1314, 1348
Karakitsos, Petros / *National and Kapodistrian University of Athens, Greece* 266
Karakitsou, Effrosyni / *University of Barcelona, Spain* 266
Kathuria, Mamta / *YMCA University of Science and Technology, India* 818
Khan, A. M. / *Mangalore University, India* 1199
Khan, Afaq Alam / *Central University of Kashmir, India* 1115
Kiourt, Chairi / *Hellenic Open University, Greece* 1314, 1348
Koureas, Nikolaos / *St. Savas Hospital, Greece* 266
Koutsoudis, Anestis / *University Campus at Kimmeria, Greece* 1314, 1348
Kucukasci, Emel Seyma / *Istanbul Commerce University, Turkey* 322
Kulkarni, U. V. / *SGGS IE&T, India* 1149
Kumar K., Vimal / *Jaypee Institute of Information Technology, India* 410
Kumar, Amit / *University of Petroleum and Energy Studies, India* 722
Kumar, Santosh / *IMS UNISON University, India* 23
Kumar, Shishir / *Jaypee University of Engineering and Technology, India* 808
Leonardi, Giorgio / *University of Piemonte Orientale, Italy* 658
Lester, James C. / *North Carolina State University, USA* 214
Li, Zhaoyu / *Beijing Institute of Technology, China* 248
Lindner, Lars / *Autonomous University of Baja California, Mexico* 184
Llorente, Maria Salamó / *University of Barcelona, Spain* 1408
Lobato, Emilio J. C. / *Illinois State University, USA* 791
Lobene, Eleni V. / *Aon Hewitt, USA* 214
Lobiyal, D.K. / *Jawaharlal Nehru University, India* 770
Lopez-Sanchez, Maite / *University of Barcelona, Spain* 1408
López-Sánchez, Maite / *Universitat de Barcelona, Spain* 860
Lu, Gehao / *University of Huddersfield, UK & Yunnan University, China* 1
Lu, Joan / *University of Huddersfield, UK* 1
Mallat, Souheyl / *University of Monastir, Tunisia* 1686
Mane, D. T. / *SGGS IE&T, India* 1149
Margari, Niki / *National and Kapodistrian University of Athens, Greece* 266
Martoglia, Riccardo / *University of Modena and Reggio Emilia, Italy* 302
Matthies, Benjamin / *South Westphalia University of Applied Sciences, Germany* 878
Medina, Roseclea Duarte / *Federal University of Santa Maria, Brazil* 912
Meena, Yogesh Kumar / *MNIT Jaipur, India* 619
Mehta, Manish / *Accenture Technology Lab, USA* 561
Mellouli, Sehl / *Université Laval, Canada* 1162
Menif, Emna / *University of Carthage, Tunisia* 1162
Mir, Roohie Naaz / *National Institute of Technology, India* 1214
Mishra, Vinod Kumar / *Bipin Tripathi Kumaon Institute of Technology, India* 31
Mori, Naoki / *Osaka Prefecture University, Japan* 1058
Mott, Bradford W. / *North Carolina State University, USA* 214
Mukherjee, Saurabh / *Banasthali University, India* 83

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Nagpal, Chander Kumar / *YMCA University of Science and Technology, India* 818
Nava, Jose / *Pinnacle Aerospace, Inc., Mexico* 112
Nemegyei, Berenice Zambrano / *MediaLab Prado, Spain* 860
Nguyen, Nhu Gia / *Duy Tan University, Vietnam* 1199
Niewiadomski, Robert / *NYCDOE, USA* 1298
Nishani, Lediona / *University of New York Tirana, Albania* 688
Nunes, Felipe Becker / *Federal University of Rio Grande do Sul, Brazil* 912
Ogata, Takashi / *Iwate Prefectural University, Japan* 743
Oh, Jungsuk / *Seoul National University, South Korea* 1085
Osorio, Alejandro / *Pinnacle Aerospace, Inc., Mexico* 112
Ozil, Eralp / *Zeta Information Technologies, Turkey* 139
Ozpinar, Alper / *Istanbul Commerce University, Turkey* 139, 322
Panayiotides, Ioannis G / *National and Kapodistrian University of Athens, Greece* 266
Pappas, Asimakis / *MHTERA Maternity Hospital, Greece* 266
Park, Jinsoo / *Seoul National University, South Korea* 1085
Pathak, Parth S / *ezDI, LLC, India* 991
Pavlidis, George / *University Campus at Kimmeria, Greece* 1314
Piovesan, Luca / *Università degli Studi di Torino, Italy* 658
Pouliakis, Abraham / *National and Kapodistrian University of Athens, Greece* 266
Powell, Vaughan / *University of Portsmouth, UK* 1459
Powell, Wendy A. / *University of Portsmouth, UK* 1459
Puente, Gustavo J. / *Autonomous University of Coahuila, Mexico* 1533
Puig, Anna / *University of Barcelona, Spain* 1408
Rawat, Bhupesh / *Babasaheb Bhimrao Ambedkar University (BBAU), India* 1621
Ribera, Mireia / *University of Barcelona, Spain* 1408
Rivas-Lopez, Moises / *Autonomous University of Baja California, Mexico* 184
Rivera-Castillo, Javier / *Autonomous University of Baja California, Mexico* 184
Rodriguez, Inmaculada / *University of Barcelona, Spain* 1408
Rodríguez-Quiñonez, Julio C. / *Autonomous University of Baja California, Mexico* 184
Rowe, Jonathan P. / *North Carolina State University, USA* 214
Rubio, Dionisio Sánchez / *Universitat Politècnica de València (UPV) & Escola d'Art i Superior de Disseny de València EASD, Spain* 860
Saha, Sriparna / *Indian Institute of Technology Patna, India* 991
Sanchez-Anguix, Victor / *Coventry University, UK* 464
Sawicka, Hanna / *Poznan University of Technology, Poland* 957
Schafer, Stephen Brock / *Digipen Institute of Technology, USA* 343
Selot, Smita / *SSTC, India* 1011
Sergiyenko, Oleg / *Autonomous University of Baja California, Mexico* 184
Sernani, Paolo / *Università Politecnica delle Marche, Italy* 1652
Sharma, Roopali / *Birla Institute of Technology, India* 23
Shelke, Nilesh M / *Priyadarshini Indira Gandhi College of Engineering, India* 422
Sholla, Sahil / *National Institute of Technology, India* 1214
Singh, Aarti / *Maharishi Markandeshwar University, India* 83
Singh, Ninni / *University of Petroleum and Energy Studies, India* 722
Singh, Rashmi / *Banasthali University, India* 83
Singh, Umrinderpal / *Punjabi University, India* 1139

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sioulas, Vasileios / *National and Kapodistrian University of Athens, Greece* 266
Solano, Angela M. / *Universidad de Córdoba, Colombia* 983
Song, Zheng / *Virginia Tech, USA* 939
Srivastava, Arvind K. / *Systron Donner Inertial, USA* 1199
Tatnall, Arthur / *Victoria University, Australia* 1286
Terenziani, Paolo / *University of Piemonte Orientale, Italy* 658
Terzakis, Emmanouil G / *St. Savas Hospital, Greece* 266
Tiruwa, Himanshu / *Bipin Tripathi Kumaon Institute of Technology, India* 31
Tripathi, Neeta / *SSTC, India* 1011
Tsuji, Hiroshi / *Osaka Prefecture University, Japan* 394
Tuğran, Aytaç Hakan / *Istanbul University, Turkey* 1272
Tuğran, Filiz Erdoğan / *Ondokuz Mayis University, Turkey* 1272
Tunç, Ayşegül Özbebek / *Istanbul University, Turkey* 1667
Turner, Jeremy Owen / *Simon Fraser University, Canada* 485
Ueno, Miki / *Toyohashi University of Technology, Japan* 1058
Vega, Yina M. / *Universidad de Córdoba, Colombia* 983
Vieira, Armando / *Redoctopus, UK* 440
Voss, Gleizer Bierhalz / *Federal Institute of Education Science and Technology Farroupilha, Brazil* 912
Wahlstedt, Robert / *Independent Researcher, USA* 1489
Walczak, Steven / *University of South Florida, USA* 160
Wang, Tianzi / *Virginia Tech, USA* 939
Wehle, Marko / *Technische Universität Berlin, Germany* 1507
Weidemann, Alexandra / *Technische Universität Berlin, Germany* 1507
Xu, Rui / *Beijing Institute of Technology, China* 248
Yadav, Divakar / *Jaypee Institute of Information Technology, India* 410
Yadav, Shweta / *Indian Institute of Technology Patna, India* 991
Yampolskiy, Roman V. / *JB Speed School of Engineering, USA* 1232, 1367
Yang, Mingxuan / *South China Normal University, China* 1436
Zadgaonkar, A. S. / *CV Raman University, India* 1011
Zeleznikow, John / *Victoria University, Australia* 49
Zhu, Shengying / *Beijing Institute of Technology, China* 248
Ziesche, Soenke / *Maldives National University, Maldives* 1367
Zincir, Oya / *Istanbul University, Turkey* 1667
Zohora, Syeda Erfana / *Taif University, Saudi Arabia* 1199
Zrigui, Mounir / *University of Monastir, Tunisia* 1686

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table of Contents

Preface xxi

Volume I

Section 1
Fundamental Concepts and Theories

Chapter 1
Background Review for Neural Trust and Multi-Agent System 1
Gehao Lu, University of Huddersfield, UK & Yunnan University, China
Joan Lu, University of Huddersfield, UK

Chapter 2
Applications of AI in Financial System 23
Santosh Kumar, IMS UNISON University, India
Roopali Sharma, Birla Institute of Technology, India

Chapter 3
Aspect-Based Sentiment Analysis of Online Product Reviews 31
Vinod Kumar Mishra, Bipin Tripathi Kumaon Institute of Technology, India
Himanshu Tiruwa, Bipin Tripathi Kumaon Institute of Technology, India

Section 2
Development and Design Methodologies

Chapter 4
Compliance with International Soft Law: Is the Adoption of Soft Law Predictable? 49
Michael D'Rosario, Deakin University, Australia
John Zeleznikow, Victoria University, Australia

Chapter 5
Predicting Credit Rating Migration Employing Neural Network Models 65
Michael D'Rosario, Deakin University, Australia
Calvin Hsieh, Deakin University, Australia

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 6
A Thorough Insight into Theoretical and Practical Developments in MultiAgent Systems 83
Dimple Juneja, Dronacharya Institute of Management and Technology, India
Aarti Singh, Maharishi Markandeshwar University, India
Rashmi Singh, Banasthali University, India
Saurabh Mukherjee, Banasthali University, India

Chapter 7
A Hybrid Intelligent Risk Identification Model for Configuration Management in Aerospace Systems 112
Jose Nava, Pinnacle Aerospace, Inc., Mexico
Alejandro Osorio, Pinnacle Aerospace, Inc., Mexico

Chapter 8
Smart Grid and Demand Side Management: Application of Metaheuristic and Artificial Intelligence Algorithms 139
Alper Ozpinar, Istanbul Commerce University, Turkey
Eralp Ozil, Zeta Information Technologies, Turkey

Chapter 9
Society of Agents: A Framework for Multi-Agent Collaborative Problem Solving 160
Steven Walczak, University of South Florida, USA

Chapter 10
Applying Optoelectronic Devices Fusion in Machine Vision: Spatial Coordinate Measurement 184
Wendy Flores-Fuentes, Autonomous University of Baja California, Mexico
Moises Rivas-Lopez, Autonomous University of Baja California, Mexico
Daniel Hernandez-Balbuena, Autonomous University of Baja California, Mexico
Oleg Sergiyenko, Autonomous University of Baja California, Mexico
Julio C. Rodríguez-Quiñonez, Autonomous University of Baja California, Mexico
Javier Rivera-Castillo, Autonomous University of Baja California, Mexico
Lars Lindner, Autonomous University of Baja California, Mexico
Luis C. Basaca-Preciado, Center of Excellence in Innovation & Design – CETYS University, Mexico

Chapter 11
Play in the Museum: Design and Development of a Game-Based Learning Exhibit for Informal Science Education 214
Jonathan P. Rowe, North Carolina State University, USA
Eleni V. Lobene, Aon Hewitt, USA
Bradford W. Mott, North Carolina State University, USA
James C. Lester, North Carolina State University, USA

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 12
Using Physics Inspired Wave Agents in a Virtual Environment: Longitudinal Distance Control in Robots Platoon 232
Baudouin Dafflon, Université de Lyon, France
Maxime Guériau, Univ Lyon, ENTPE, IFSTTA, LICIT UMR_T 9401, F-69675, France
Franck Gechter, LE2I, UTBM, University of Bourgogne Franche Comté, France

Chapter 13
A Geometric Dynamic Temporal Reasoning Method with Tags for Cognitive Systems 248
Rui Xu, Beijing Institute of Technology, China
Zhaoyu Li, Beijing Institute of Technology, China
Pingyuan Cui, Beijing Institute of Technology, China
Shengying Zhu, Beijing Institute of Technology, China
Ai Gao, Beijing Institute of Technology, China

Chapter 14
Artificial Intelligence for the Identification of Endometrial Malignancies: Application of the Learning Vector Quantizer 266
Abraham Pouliakis, National and Kapodistrian University of Athens, Greece
Niki Margari, National and Kapodistrian University of Athens, Greece
Effrosyni Karakitsou, University of Barcelona, Spain
Evangelia Alamanou, Tzaneio Hospital, Greece
Nikolaos Koureas, St. Savas Hospital, Greece
George Chrelias, National and Kapodistrian University of Athens, Greece
Vasileios Sioulas, National and Kapodistrian University of Athens, Greece
Asimakis Pappas, MHTERA Maternity Hospital, Greece
Charalambos Chrelias, National and Kapodistrian University of Athens, Greece
Emmanouil G Terzakis, St. Savas Hospital, Greece
Vasileia Damaskou, National and Kapodistrian University of Athens, Greece
Ioannis G Panayiotides, National and Kapodistrian University of Athens, Greece
Petros Karakitsos, National and Kapodistrian University of Athens, Greece

Chapter 15
Expanding Bloom's Two-Sigma Tutoring Theory Using Intelligent Agents: Application to Management Education 280
Owen P. Hall Jr., Pepperdine University, USA

Chapter 16
A User-Aware and Semantic Approach for Enterprise Search 302
Giacomo Cabri, University of Modena and Reggio Emilia, Italy
Riccardo Martoglia, University of Modena and Reggio Emilia, Italy

Chapter 17
Use of Chaotic Randomness Numbers: Metaheuristic and Artificial Intelligence Algorithms 322
Alper Ozpinar, Istanbul Commerce University, Turkey
Emel Seyma Kucukasci, Istanbul Commerce University, Turkey

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 18
The Media-Sphere as Dream Researching the Contextual Unconscious of Collectives 343
Stephen Brock Schafer, Digipen Institute of Technology, USA

Chapter 19
On the Use of Bayesian Network in Crime Suspect Modelling and Legal Decision Support 372
O. E. Isafiade, University of Cape Town, South Africa
A. B. Bagula, University of the Western Cape, South Africa
S. Berman, University of Cape Town, South Africa

Chapter 20
Information Extraction for Call for Paper 394
Laurent Issertial, Osaka Prefecture University, Japan
Hiroshi Tsuji, Osaka Prefecture University, Japan

Chapter 21
Word Sense Based Hindi-Tamil Statistical Machine Translation 410
Vimal Kumar K., Jaypee Institute of Information Technology, India
Divakar Yadav, Jaypee Institute of Information Technology, India

Chapter 22
Exploiting Chi Square Method for Sentiment Analysis of Product Reviews 422
Nilesh M Shelke, Priyadarshini Indira Gandhi College of Engineering, India
Shrinivas P Deshpande, DCPE, HVPM, India

Chapter 23
Business Applications of Deep Learning 440
Armando Vieira, Redoctopus, UK

Section 3
Tools and Technologies

Chapter 24
Agreement Technologies for Conflict Resolution 464
Vicente Julian, Universidad Politécnica de Valencia, Spain
Victor Sanchez-Anguix, Coventry University, UK
Stella Heras, Universidad Politécnica de Valencia, Spain
Carlos Carrascosa, Universidad Politécnica de Valencia, Spain

Chapter 25
Virtual Soar-Agent Implementations: Examples, Issues, and Speculations 485
Jeremy Owen Turner, Simon Fraser University, Canada

Chapter 26
Applications of PNC in Artificial Intelligence 509
Dariusz Jacek Jakóbczak, Technical University of Koszalin, Poland

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 27
Fair Use Defences During Copyright Litigation: Is the Success of a Fair Use Defence Strategy Predictable? 539
Michael D'Rosario, Deakin University, Australia

Volume II

Chapter 28
A Graphical Tool for the Creation of Behaviors in Virtual Worlds 561
Andrea Corradini, Design School Kolding, Denmark
Manish Mehta, Accenture Technology Lab, USA

Chapter 29
Contour Reconstruction: 2D Object Modeling 584
Dariusz Jacek Jakóbczak, Technical University of Koszalin, Poland

Chapter 30
Statistical Features for Extractive Automatic Text Summarization 619
Yogesh Kumar Meena, MNIT Jaipur, India
Dinesh Gopalani, MNIT Jaipur, India

Chapter 31
Improving Collaborative Filtering Algorithms: Sentiment-based Approach in Social Network 638
Firas Ben Kharrat, University of Carthage, Tunisia
Aymen Elkhleifi, Paris-Sorbonne University, France
Rim Faiz, University of Carthage, Tunisia

Chapter 32
Knowledge-Based Support to the Treatment of Exceptions in Computer Interpretable Clinical Guidelines 658
Alessio Bottrighi, University of Piemonte Orientale, Italy
Giorgio Leonardi, University of Piemonte Orientale, Italy
Luca Piovesan, Università degli Studi di Torino, Italy
Paolo Terenziani, University of Piemonte Orientale, Italy

Chapter 33
Statistical Relational Learning for Collaborative Filtering a State-of-the-Art Review 688
Lediona Nishani, University of New York Tirana, Albania
Marenglen Biba, University of New York in Tirana, Albania

Chapter 34
An Artificial Intelligence-Based Vehicular System Simulator 708
Marvin T. Chan, University of Regina, Canada
Jonathan T. Chan, University of Regina, Canada
Christine Chan, University of Regina, Canada
Craig Gelowitz, University of Regina, Canada

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 35
A Novel Architecture for Learner-Centric Curriculum Sequencing in Adaptive Intelligent Tutoring System 722
Ninni Singh, University of Petroleum and Energy Studies, India
Neelu Jyothi Ahuja, University of Petroleum and Energy Studies, India
Amit Kumar, University of Petroleum and Energy Studies, India

Chapter 36
Designing a Socially Open Narrative Generation System 743
Taisuke Akimoto, The University of Electro-Communications, Japan
Takashi Ogata, Iwate Prefectural University, Japan

Chapter 37
Conceptual Graphs Based Approach for Subjective Answers Evaluation 770
Goonjan Jain, Jawaharlal Nehru University, India
D.K. Lobiyal, Jawaharlal Nehru University, India

Chapter 38
Wild Architecture: Explaining Cognition via Self-Sustaining Systems 791
Vincent T. Cialdella, Illinois State University, USA
Emilio J. C. Lobato, Illinois State University, USA
J. Scott Jordan, Illinois State University, USA

Chapter 39
Predictive Analysis of Emotions for Improving Customer Services 808
Vinay Kumar Jain, Jaypee University of Engineering and Technology, India
Shishir Kumar, Jaypee University of Engineering and Technology, India

Chapter 40
A Fuzzy Logic Based Synonym Resolution Approach for Automated Information Retrieval 818
Mamta Kathuria, YMCA University of Science and Technology, India
Chander Kumar Nagpal, YMCA University of Science and Technology, India
Neelam Duhan, YMCA University of Science and Technology, India

Section 4
Utilization and Applications

Chapter 41
Finding the Semantic Relationship Between Wikipedia Articles Based on a Useful Entry Relationship 838
Lin-Chih Chen, National Dong Hwa University, Tawian

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 42
Co-Designing Participatory Tools for a New Age: A Proposal for Combining Collective and Artificial Intelligences 860
José Luis Fernández-Martínez, Institute of Advanced Social Studies (IESA-CSIC) & University of Granada, Spain
Maite López-Sánchez, Universitat de Barcelona, Spain
Juan Antonio Rodríguez Aguilar, Artificial Intelligence Research Institute (IIIA-CSIC), Spain
Dionisio Sánchez Rubio, Universitat Politècnica de València (UPV) & Escola d'Art i Superior de Disseny de València EASD, Spain
Berenice Zambrano Nemegyei, MediaLab Prado, Spain

Chapter 43
Exploring the Conceptual Nature of e-Business Projects 878
Benjamin Matthies, South Westphalia University of Applied Sciences, Germany
André Coners, South Westphalia University of Applied Sciences, Germany

Chapter 44
Three-Dimensional Virtual Environment and NPC: A Perspective about Intelligent Agents Ubiquitous 912
Fabricio Herpich, Federal University of Rio Grande do Sul, Brazil
Felipe Becker Nunes, Federal University of Rio Grande do Sul, Brazil
Gleizer Bierhalz Voss, Federal Institute of Education Science and Technology Farroupilha, Brazil
Roseclea Duarte Medina, Federal University of Santa Maria, Brazil

Chapter 45
Modelling Propagation of Public Opinions on Microblogging Big Data Using Sentiment Analysis and Compartmental Models 939
Youjia Fang, Virginia Tech, USA
Xin Chen, Virginia Tech, USA
Zheng Song, Virginia Tech, USA
Tianzi Wang, Virginia Tech, USA
Yang Cao, Virginia Tech, USA

Chapter 46
Artificial Intelligence in Stochastic Multiple-Criteria Decision Making 957
Hanna Sawicka, Poznan University of Technology, Poland

Chapter 47
Trends of Educational Informatics in Latin America 983
Adán A. Gómez, Universidad de Córdoba, Colombia
Manuel F. Caro, Universidad de Córdoba, Colombia
Angela M. Solano, Universidad de Córdoba, Colombia
Yina M. Vega, Universidad de Córdoba, Colombia

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 48
Patient Data De-Identification: A Conditional Random-Field-Based Supervised Approach 991
Shweta Yadav, Indian Institute of Technology Patna, India
Asif Ekbal, Indian Institute of Technology Patna, India
Sriparna Saha, Indian Institute of Technology Patna, India
Parth S Pathak, ezDI, LLC, India
Pushpak Bhattacharyya, Indian Institute of Technology Patna, India

Chapter 49
Neural Network Model for Semantic Analysis of Sanskrit Text 1011
Smita Selot, SSTC, India
Neeta Tripathi, SSTC, India
A. S. Zadgaonkar, CV Raman University, India

Chapter 50
Probabilistic Nodes Combination (PNC): Formulas and Examples 1026
Dariusz Jacek Jakóbczak, Technical University of Koszalin, Poland

Chapter 51
Can Computers Create Comics and Animations? 1058
Miki Ueno, Toyohashi University of Technology, Japan
Kiyohito Fukuda, Osaka Prefecture University, Japan
Naoki Mori, Osaka Prefecture University, Japan

Chapter 52
A Novel Approach to Managing the Dynamic Nature of Semantic Relatedness 1085
Youngseok Choi, Brunel University London, UK
Jungsuk Oh, Seoul National University, South Korea
Jinsoo Park, Seoul National University, South Korea

Volume III

Chapter 53
Emotion Mining Using Semantic Similarity 1115
Rafiya Jan, Central university Of Kashmir, India
Afaq Alam Khan, Central University of Kashmir, India

Chapter 54
Development of Part of Speech Tagger for Assamese Using HMM 1139
Surjya Kanta Daimary, Punjabi University, India
Vishal Goyal, Punjabi University, India
Madhumita Barbora, Tezpur University, India
Umrinderpal Singh, Punjabi University, India

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 55
A Survey on Supervised Convolutional Neural Network and Its Major Applications 1149
D. T. Mane, SGGS IE&T, India
U. V. Kulkarni, SGGS IE&T, India

Chapter 56
From Citizens to Decision-Makers: A Natural Language Processing Approach in Citizens' Participation 1162
Eya Boukchina, University of Carthage, Tunisia
Sehl Mellouli, Université Laval, Canada
Emna Menif, University of Carthage, Tunisia

Chapter 57
Logs Analysis of Adapted Pedagogical Scenarios Generated by a Simulation Serious Game Architecture 1178
Sophie Callies, University of Québec à Montréal, Canada
Mathieu Gravel, University of Québec à Montréal, Canada
Eric Beaudry, University of Québec à Montréal, Canada
Josianne Basque, TELUQ University, Canada

Chapter 58
A Study of the State of the Art in Synthetic Emotional Intelligence in Affective Computing 1199
Syeda Erfana Zohora, Taif University, Saudi Arabia
A. M. Khan, Mangalore University, India
Arvind K. Srivastava, Systron Donner Inertial, USA
Nhu Gia Nguyen, Duy Tan University, Vietnam
Nilanjan Dey, TechnoIndia College of Technology, India

Section 5
Organizational and Social Implications

Chapter 59
Eventuality of an Apartheid State of Things: An Ethical Perspective on the Internet of Things 1214
Sahil Sholla, National Institute of Technology, India
Roohie Naaz Mir, National Institute of Technology, India
Mohammad Ahsan Chishti, National Institute of Technology, India

Chapter 60
Security Solutions for Intelligent and Complex Systems 1232
Stuart Armstrong, Future of Humanity Institute, UK
Roman V. Yampolskiy, JB Speed School of Engineering, USA

Chapter 61
Homo-di-fict: Creations Turn Against Humanity in South Park Town 1272
Filiz Erdoğan Tuğran, Ondokuz Mayis University, Turkey
Aytaç Hakan Tuğran, Istanbul University, Turkey

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 62
The Internet of Things and Beyond: Rise of the Non-Human Actors 1286
Arthur Tatnall, Victoria University, Australia
Bill Davey, RMIT University, Australia

Chapter 63
The Rise of Artificial Intelligence: Its Impact on Labor Market and Beyond 1298
Robert Niewiadomski, NYCDOE, USA
Dennis Anderson, St. Francis College, USA

Chapter 64
Realistic Simulation of Cultural Heritage 1314
Chairi Kiourt, Hellenic Open University, Greece
George Pavlidis, University Campus at Kimmeria, Greece
Anestis Koutsoudis, University Campus at Kimmeria, Greece
Dimitris Kalles, Hellenic Open University, Greece

Chapter 65
Enhanced Virtual Reality Experience in Personalised Virtual Museums 1348
Chairi Kiourt, Athena Research Centre, Greece
Anestis Koutsoudis, Athena Research Centre, Greece
Dimitris Kalles, Hellenic Open University, Greece

Chapter 66
High Performance Computing of Possible Minds 1367
Soenke Ziesche, Maldives National University, Maldives
Roman V. Yampolskiy, University of Louisville, USA

Chapter 67
Data Extrapolation via Curve Modeling in Analyzing Risk: Value Anticipation for Decision Making 1379
Dariusz Jacek Jakóbczak, Technical University of Koszalin, Poland

Chapter 68
Training Infrastructure to Participate in Real Life Institutions: Learning through Virtual Worlds 1408
Pablo Almajano, University of Barcelona, Spain
Maite Lopez-Sanchez, University of Barcelona, Spain
Inmaculada Rodriguez, University of Barcelona, Spain
Anna Puig, University of Barcelona, Spain
Maria Salamó Llorente, University of Barcelona, Spain
Mireia Ribera, University of Barcelona, Spain

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 69
Semantic Search Exploiting Formal Concept Analysis, Rough Sets, and Wikipedia..... 1436
Yuncheng Jiang, South China Normal University, China
Mingxuan Yang, South China Normal University, China

Chapter 70
The Rise of the Virtual Human..... 1459
Wendy A. Powell, University of Portsmouth, UK
Natalie Corbett, University of Portsmouth, UK
Vaughan Powell, University of Portsmouth, UK

Chapter 71
How Games Improve Language in People With Language Dysfunctions..... 1489
Robert Wahlstedt, Independent Researcher, USA

Section 6
Critical Issues and Challenges

Chapter 72
Research on Human Cognition for Biologically Inspired Developments: Human-Robot Interaction by Biomimetic AI..... 1507
Marko Wehle, Technische Universität Berlin, Germany
Alexandra Weidemann, Technische Universität Berlin, Germany
Ivo Wilhelm Boblan, Technische Universität Berlin, Germany

Chapter 73
How to Identify Rheumatic Diseases by General Physicians..... 1533
Eduardo C. Contreras, Autonomous University of Coahuila, Mexico
Gustavo J. Puente, Autonomous University of Coahuila, Mexico

Chapter 74
Social Implications of Big Data and Fog Computing..... 1564
Jeremy Horne, Inventors Assistance League, USA

Section 7
Emerging Trends

Chapter 75
State of the Art Recommendation Approaches: Their Issues and Future Research Direction in E-Learning A Survey..... 1621
Bhupesh Rawat, Babasaheb Bhimrao Ambedkar University (BBAU), India
Sanjay K. Dwivedi, Babasaheb Bhimrao Ambedkar University (BBAU), India

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 76
Combining Artificial Intelligence and NetMedicine for Ambient Assisted Living: A Distributed BDI-based Expert System 1652
Paolo Sernani, Università Politecnica delle Marche, Italy
Andrea Claudi, Università Politecnica delle Marche, Italy
Aldo Franco Dragoni, Università Politecnica delle Marche, Italy

Chapter 77
An Imagination of Organizations in the Future: Rethinking McKinsey's 7S Model 1667
Oya Zincir, Istanbul University, Turkey
Ayşegül Özbebek Tunç, Istanbul University, Turkey

Chapter 78
Events Automatic Extraction from Arabic Texts 1686
Emna Hkiri, University of Monastir, Tunisia
Souheyl Mallat, University of Monastir, Tunisia
Mounir Zrigui, University of Monastir, Tunisia

Index xxiv

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Preface

Information in today's advancing world is rapidly expanding and becoming widely available. This eruption of data has made handling it a daunting and time-consuming task. Natural language processing (NLP) is a method that applies linguistics and algorithms to large amounts of this data to make it more valuable. Persistent research in the field of natural language processing enables an understanding of the world around us in addition to opportunities for manmade computing to mirror natural language processes that have existed for centuries.

Staying informed of the most up-to-date research trends and findings is of the utmost importance. That is why IGI Global is pleased to offer this three-volume reference collection that will shed light on critical issues on the methods and applications of linguistic information processing and its computational properties by providing both broad and detailed perspectives on cutting-edge theories and developments. This reference is designed to act as a single reference source on conceptual, methodological, technical, and managerial issues, as well as provide insight into emerging trends and future opportunities on the fundamental concepts and techniques of natural language processing.

Natural Language Processing: Concepts, Methodologies, Tools, and Applications is organized into seven distinct sections that provide comprehensive coverage of important topics. The sections are:

1. Fundamental Concepts and Theories;
2. Development and Design Methodologies;
3. Tools and Technologies;
4. Utilization and Applications;
5. Organizational and Social Implications;
6. Critical Issues and Challenges; and
7. Emerging Trends.

The following paragraphs provide a summary of what to expect from this invaluable reference tool.

Section 1, "Fundamental Concepts and Theories," serves as a foundation for this extensive reference tool by addressing crucial theories essential to the understanding of natural language processing. Introducing the book is "Background Review for Neural Trust and Multi-Agent System," by Prof. Gehao Lu of the University of Huddersfield, UK & Yunnan University, China and Prof. Joan Lu from the University of Huddersfield, UK: a great foundation laying the groundwork for the basic concepts and theories that will be discussed throughout the rest of the book. Section 1 concludes and leads into the following portion of the book with a nice segue chapter, "Aspect-Based Sentiment Analysis of Online Product Reviews," by Profs. Vinod Kumar Mishra and Himanshu Tiruwa from Bipin Tripathi Kumaon Institute of Technology, India.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Section 2, "Development and Design Methodologies," presents in-depth coverage of the conceptual design and architecture of natural language processing. Opening the section is "Compliance With International Soft Law: Is the Adoption of Soft Law Predictable?" by Prof. Michael D'Rosario from Deakin Business School, Deakin University, Sydney, Australia and Prof. John Zeleznikow from the College of Business, Victoria University, Melbourne, Australia. Through case studies, this section lays excellent groundwork for later sections that examine present and future applications for natural language processing. The section concludes with an excellent work by Prof. Armando Vieira of Redoctopus, UK, "Business Applications of Deep Learning."

Section 3, "Tools and Technologies," presents extensive coverage of the various tools and technologies used in the implementation of natural language processing. The first chapter, "Agreement Technologies for Conflict Resolution," by Prof. Vicente Julian from the Universidad Politécnica de Valencia, Spain; Prof. Victor Sanchez-Anguix from Coventry University, UK; and Profs. Stella Heras and Carlos Carrascosa from the Universidad Politécnica de Valencia, Spain, lays a framework for the types of works that can be found in this section. The section concludes with "A Fuzzy Logic Based Synonym Resolution Approach for Automated Information Retrieval," by Profs. Mamta Kathuria, Chander Kumar Nagpal, and Neelam Duhan of the YMCA University of Science and Technology, Faridabad, India. Where Section 3 described specific tools and technologies at the disposal of practitioners, Section 4 describes the use and applications of the tools and frameworks discussed in previous sections.

Section 4, "Utilization and Applications," describes how the broad range of natural language processing efforts has been utilized and offers insight on and important lessons for their applications and impact. The first chapter in the section is "Finding the Semantic Relationship Between Wikipedia Articles Based on a Useful Entry Relationship," written by Prof. Lin-Chih Chen from the Department of Information Management, National Dong Hwa University, Hualien, Taiwan. This section includes the widest range of topics because it describes case studies, research, methodologies, frameworks, architectures, theory, analysis, and guides for implementation. The breadth of topics covered in this section is also reflected in the diversity of its authors, from countries all over the globe. The section concludes with "A Study of the State of the Art in Synthetic Emotional Intelligence in Affective Computing," by Prof. Syeda Erfana Zohora from Taif University, Taif, Saudi Arabia; Prof. A. M. Khan from Mangalore University, Magalore, India; Prof. Arvind K. Srivastava of Systron Donner Inertial, Concord, CA, USA; Prof. Nhu Gia Nguyen from Duy Tan University, Da Nang, Vietnam; and Prof. Nilanjan Dey of TechnoIndia College of Technology, Kolkata, India, a great transition chapter into the next section.

Section 5, "Organizational and Social Implications," includes chapters discussing the organizational and social impact of natural language processing. The section opens with "Eventuality of an Apartheid State of Things: An Ethical Perspective on the Internet of Things," by Profs. Sahil Sholla, Roohie Naaz Mir, and Mohammad Ahsan Chishti from the Computer Science and Engineering Department, National Institute of Technology Srinagar, India. This section focuses exclusively on how these technologies affect human lives, either through the way they interact with each other or through how they affect behavioral/workplace situations. The section concludes with "How Games Improve Language in People With Language Dysfunctions," by Robert Wahlstedt, an Independent Researcher, USA.

Section 6, "Critical Issues and Challenges," presents coverage of academic and research perspectives on natural language processing tools and applications. The section begins with "Research on Human Cognition for Biologically Inspired Developments: Human-Robot Interaction by Biomimetic AI," by Profs. Marko Wehle, Alexandra Weidemann, and Ivo Wilhelm Boblan from the Technische Universität Berlin, Germany. Chapters in this section will investigate theoretical approaches and offer alternatives

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

to crucial questions on the subject of natural language processing. The section concludes with "Social Implications of Big Data and Fog Computing," by Prof. Jeremy Horne of Inventors Assistance League, Prescott, USA.

Section 7, "Emerging Trends," highlights areas for future research within the field of natural language processing, opening with "State of the Art Recommendation Approaches: Their Issues and Future Research Direction in E-Learning – A Survey," by Profs. Bhupesh Rawat and Sanjay K. Dwivedi from the Department of Computer Science, Babasaheb Bhimrao Ambedkar University (BBAU), Lucknow, India. This section contains chapters that look at what might happen in the coming years that can extend the already staggering amount of applications for natural language processing. The final chapter of the book looks at an emerging field within natural language processing, in the excellent contribution, "Events Automatic Extraction From Arabic Texts," by Profs. Emna Hkiri, Souheyl Mallat, and Mounir Zrigui from the Faculty of Sciences, University of Monastir, Monastir, Tunisia.

Although the primary organization of the contents in this multi-volume work is based on its seven sections, offering a progression of coverage of the important concepts, methodologies, technologies, applications, social issues, and emerging trends, the reader can also identify specific contents by utilizing the extensive indexing system listed at the end of each volume. As a comprehensive collection of research on the latest findings related to using technology to providing various services, *Natural Language Processing: Concepts, Methodologies, Tools, and Applications* provides researchers, administrators, and all audiences with a complete understanding of the development of applications and concepts in natural language processing. Given the vast number of issues concerning usage, failure, success, policies, strategies, and applications of natural language processing in countries around the world, *Natural Language Processing: Concepts, Methodologies, Tools, and Applications* addresses the demand for a resource that encompasses the most pertinent research in technologies being employed to globally bolster the knowledge and applications of natural language processing.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Section 1

Fundamental Concepts and Theories

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 1
Background Review for Neural Trust and Multi-Agent System

Gehao Lu
University of Huddersfield, UK & Yunnan University, China

Joan Lu
University of Huddersfield, UK

ABSTRACT

This chapter provides a systematic background study in the neural trust and multi-agent system. Theoretic models are discussed in details. The concepts are explained. The existing systems are analyzed. The limitations and strength of previous research are discussed. About 59 references are cited to support the study for the investigation. The study did address the research importance and significance and finally, proposed the future directions for the research undertaken.

1. INTRODUCTION

Multi-agent systems are initially introduced as a branch of distributed artificial intelligence. In 1986, Minsky and Murata proposed the concept of agents and he thought some problems could be solved through the negotiation among the social individuals (Minsky & Murata, 2004). These social individuals are agents. Agents should be highly interactive and intelligent. Hewitt thought that it is difficult to define agent as it is to define intelligence (Hewitt, 1985). Wooldridge and Jennings thought that agents should be autonomous, social interactive, proactive, and reactive (Wooldridge & Jennings, 2002). Eberhart & Shi thought that agents can react through sensing the outside environments (Eberhart & Shi, 2007).

With the increased size and complexity of the computer systems, it is nearly impossible to design a system from scratch and control every details of the system purely by human brain (Simon, 1996). It is difficult to control millions of transactions occurring in a large-scale E-market. It is also difficult to monitor an enterprise information system which encompasses huge amounts of heterogeneous devices which covers thousands of different geographical locations (Rothkopf, 2003; Coulouris, Dollimore & Kindberg, 2000a). Grid Computing, Autonomous Computing, Pervasive computing and Multi-agent systems, are

DOI: 10.4018/978-1-7998-0951-7.ch001

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

all committing themselves to challenging the design of large-scale distributed system (Coulouris et al., 2000b). Computational trust is to make an intelligent agent trust another agent and delegate part of their tasks to the target agent in a heterogeneous distributed multi-agent environment. Delegation of action is the result of trust and it also forms the foundation of future large-scale cooperative computer systems. Generally, trust toward specific agent is generated through recognition and experience under repeated transactions with that agent. Reputation is the socialized trust which can be propagated through a social network of agents. It helps agents trust the target agent without any direct interaction with the target agent. The benefits of introducing trust and reputation into multi-agent system include:

- As a lubricant, trust can eliminate much of unnecessary communications which are currently necessitates many interaction protocols thus greatly improve the performance of the multi-agent systems.
- An agent can make decision easier based upon the evaluation of the trustworthiness of another agent. Computational trust is also a very beneficial addition to the traditional decision theory.

Trust is a kind of soft security which complements the traditional hard security like encryption, authorization, and authentication. An agent that exists in complex heterogeneous environment must possess both securities in order to be safe and effective.

The mechanisms for coordinating interactions among agents are always pre-defined, that is, the designer specifies how one agent responses to another agent in a fixed protocol (Huynh, 2006). Such mechanisms are not flexible enough because of the intrinsic high openness and variability of the distributed systems (Ferber, 1999). For example, in open MAS (Multi-agent Systems), an agent cannot expect to always interact with the agents in the predefined application domain in a predetermined way (Subrahmanian, Bonatti, Dix, Eiter, Kraus, Ozcan & Ross, 2000). Agents will interact with different agents coming from heterogeneous applications and they may face challenges from lying, deceiving and accidental incidents (Ferber, 1999). Such complexity creates the following questions: can agents accept services from other unfamiliar agents? Can agents make use of the interaction history and transform them into experiences? Can agents avoid excessive negotiation with a familiar agent in an efficient way? Computational trust seems to be the answer and the next step of research for the multi-agent systems. Thus, a systematic review on the existing trust models is necessary.

Trust actually is a belief that someone or agents can delegate the host to finish some actions. There are two layers of meaning in the expression: first, agents should generate the belief of trustworthiness toward some other agents in some specific form; second, agents should make decision whether to delegate actions to the trusted agent. The first layer is actually to study how agents generate and update their belief which is part of research from computational intelligence. The second layer is an extension of the traditional decision theory which adds agent's belief and trustworthiness as one of the concerns during decision making.

There are a few computational trust models and reputation models that have been proposed from (Subrahmanian et al., 2000) to (Huynh, 2006). From a point of dimensional view, there are two types of models involved, i.e. local trust based and reputation based models. For local trust based model, the early model developed by Marsh, University of Stirling, 1994, only considers the local trust dimension which only derives trust from the agent's direct interaction without referencing to the recommendations from other witness agents (Marsh, 1994a). For reputation based models, like SPORA (Zacharia, 1999), they only consider the reputation (witness trust) dimension without looking at the local experience of

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

the agent itself. Recently RegreT, Referral Network and TRAVOS take both local trust and witness trust into account (Teacy et al., 2005; Sabater, 2003). They combine the value of the two dimensions with relative weights and finally get a sum. Some models, such as FIRE, even introduce additional dimension called as role-based trust and certified reputation (Huynh, 2006; Huynh et al., 2006).

From a point of algorithmic view, different ways of calculating trust and reputation are proposed. For example, Bayesian systems take binary ratings as input and are based on calculating reputation scores by statistical updating of beta probability density functions. Such models include TRAVOS (Teacy, 2006; Teacy et al., 2005). Models such as RegreT (Sabater & Sierra, 2001) are based on discrete trust model to represent trustworthiness and untrustworthiness as a discrete value. A detailed analysis about the composing elements of trust and reputation model will be given.

The objective of the study is finding out the current situation and future trends of computational trust for multi-agent systems. It also aims at looking for common necessary compositional elements that compose the models through extracting the essence of those representative models and summarizing their common weaknesses through comparison, discussion and analysis. Finally, a clear research path is proposed to help the community to promote the research of computational trust.

2. BACKGROUND STRUCTURE

2.1 Distributed Artificial Intelligence

Intelligence and distribution are the future direction of software design. The subject of Distributed Artificial Intelligence (DAI) is a combination of Artificial intelligence and distributed computing. The Distributed Artificial Intelligence attracts the interests from researchers. The other requirement is interoperability, that is, the capabilities to exchange information and cooperate in fast changing heterogeneous environment.

In DAI, the intelligence is not an independently existing concept as it only exists in a group of agents. The interaction and cooperation among agents are the major research areas in DAI. DAI can be divided into two parts: Distributed Problem Solving (DPS) and Multi-agent System (MAS). DPS mainly considers how to decompose a problem into tasks that can be allocated to cooperating and knowledge sharing modules or nodes. The multi-agent system mainly focuses on the coordination of behaviors among autonomous agents.

Both fields need to investigate the division of knowledge, resource and control. However there is big difference between DPS and MAS. In DPS, there is only one conceptual model, one global problem, and success criteria; while in MAS, there are multiple local conceptual models, problems and success criteria. The purpose of DPS research is to create a coarse grained cooperating module. The modules in DPS cooperate to solve a problem. The problem is usually decomposed into several tasks. Each task is executed by a specialized module. All the policies of interactions are integrated as a whole. DPS is a top down approach of designing systems. On the contrary, MAS is a bottom up approach of designing systems. There are no global controls or global criteria for agents. The dispersed agents are defined first and then the problem is decomposed into several tasks which can be assigned to different agents. The relationship among agents may be cooperative, competitive or even hostile.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.2 Intelligent Agent

2.2.1 Concepts

Although the word "agent" is widely used in different areas, "agent" is still an ambiguous concept. It is difficult to find a definition accepted by all researchers. The most common viewpoint is that an agent is a computer system designed for some purpose and the system is autonomous and flexible. Wooldridge and Jennings summarize the work done in this field by pointing out that it is better to understand agent from the broad sense and the narrow sense (Wooldridge, 2002).

From the broad sense, nearly all agent based software or hardware have the following characteristics:

- **Autonomy:** Agents are not controlled by humans or other programs. They have control over the behaviors and the internal states.
- **Social Ability:** Agents can exchange information through agent communication languages. This is also called communicability.
- **Reactivity:** Agents can sense from the environment and also influence the environment. Whether the agents in real world or the agent in virtual society, they need to sense their environments and change their environments. A program which cannot influence the environment cannot be called as an agent.
- **Pro-Activeness:** The traditional application is passively run by the users and strictly follows the instructions from the users. Agents behave actively or spontaneously. Through sensing the changes in the environments, agents can perform some goal directed behaviors.

Some researchers believe that agents should have some human like characteristics such as knowledge, belief, intention, and desire. Some researchers even propose emotional agents (Bates, 1994; Bates et al., 1992). According to (Shoham, 1992, 1993), an agent is an entity with states composed of belief, capability, choice, commitment and mental component.

1.2.2 Theoretical Models

Mind elements in order to adapt to the environment and cooperate to solve problems, intelligent agents must use knowledge to change its internal status, that is, the mental state. The mental state drives the actions of the agents. It is possible to build an agent mental state from three aspects: emotion, feeling, and intention. The most famous model based on formal method is the BDI (Belief, Desire, and Intention) model. Belief represents the basic opinions owned by agents toward the environment; an agent also can use it to represent the future possible states. Desire is derived from belief. Desire includes the estimation of the future for agents. Intention confines the activities of agents; it is the subset of goals.

Logic is the most important tool to depict the mental state. Moore is one of the first researchers who used formal logic to model agents (Moore, 1990). His research focuses on the relationship between knowledge owned by agents and the realized activities. He introduces modal operator of knowledge into logic and use dynamic logic to represent actions. Such formalization allows agents to own beliefs about how to achieve goals. And agents can acquire information through executing actions. The disadvantage

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

is his modeling is too simple and incomplete. Cohen and Levesque designed a formal model which is based on linear sequential logic (1990). The model introduces concepts like time, event, behavior, goal, belief, and intention. The model also depicts the relationship among these concepts. However, the model does not reveal the relationship between intention and goal, and the model is also very abstract and difficult to implement. Kinny proposes the BDI model of rational agent based on nonlinear branch temporal logic (Kinny, George & Rao, (1996)). They use three modal operators, belief, desire and intention, to build a BDI model. Singh uses agent logic to describe intention, belief, knowledge and communication. Its logic is also based on branch temporal logic (Singh, 1994).

The behavior of rational agent cannot be driven directly from belief, desire or their combination. There should be an intention between the desire and the plan. The reason is that the behavior of agents is confined by limited resources. Once an agent decides to do some task, it needs to build a commitment with limited form. In a multi-agent environment, there must be a commitment to coordinate the behaviors of agents. Without commitment, there are no actions. Intention is just one type of commitments. In an open environment, a rational agent's behavior must be restricted by intention. If an agent changes its own intention, the changes must have some reasons. An agent cannot ignore the changes in environment and keep unimportant or not realistic activities.

1.2.3 Structures

An agent can be looked as a black box with sensors and effecters. Agents use sensors to sense the environment and use effecters to change environment. Except when interacting with environment, agents need process and explain information received from outside. The information is merged and processed into understandable data. There are three types of agent structures: deliberative agent, reactive agent, and mixed structure.

Deliberative agent is also called cognitive agent. It is an explicit symbol based model which has abilities of logical deductions (Wooldridge, 2002). It is a knowledge based system that keeps the essence of the traditional artificial intelligence. There is an environment model implemented as the major component of the knowledge library. Agents using this structure will face two basic problems (Castelfranchi, 1999): (1) transformation problem: how to translate the real world into an accurate, appropriate symbol description; (2) Representation/deduction problem: how to use the symbols to represent the entities and processes in the real world, and how to let agents make decision through deducing the information. The BDI model (Kinny et al., 1996) is a typical deliberative agent structure. The biggest disadvantage for deliberative agent is its fixed structure. In a complex, fast changing environment, it is too stationery to be reactive. The reactive agent is an agent without complex symbol deduction and symbol based world model (Cruces, 2000). It uses the condition-action rule to connect perception and action.

Neither the symbol base deliberative agent nor condition-action based reactive agent can be built into a perfect agent. Some researchers propose the mixed structure in order to combine the advantages of both types. The most direct method is to build two sub systems in an agent: one for deliberative structure, the other for reactive structure. Usually the reactive sub system has higher priority than the deliberative sub system because the agent needs to react to important event in real-time. Each agent contains a series of modules including sensory, action, reaction, modeling, planning, communication and decision making. The agent has an agent kernel which allows additional modules be attached to the kernel dynamically. The kernel is made up of internal database, mail box, and blackboard and execution engine.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1.2.4 Communications

The commonly used communication methods include the black board system and the message/dialog system. The black board system provides a public working area; agents can exchange information, data, and knowledge there. The black board is the medium between agents. Agents do not communicate with each other directly. Message based communication is the foundation of realizing complex coordination policies. Through the specified protocol, agents can exchange messages to build communication and coordination. It is different from the black board system in that the messages are directly exchanged between two agents without any caches. Message is sent to one specific agent or specific group. Only the agent or group with that address can read the message. The protocol must specify the process of communication, format of messages, and languages of communication. Of course, agents must understand the semantics of the language.

Knowledge Query Manipulation Language (KQML) is a uniform agent communication language (Labrou & Finin, 1997). KQML specifies the format of messages and the message transmitting system. It also provides a group of protocols for recognizing, establishing connection and exchanging messages. But the semantics of message content is not specified clearly in KQML. KQML can be analyzed from three layers: communication, message and content. The communication layer specifies the techniques, and the parameters of communications. The message layer specifies message related types of speech acts. The content layer specifies the content.

FIPA (Foundation for Intelligent Physical Agents) defines a language and designs supporting tools. The language is called Agent Communication Language (ACL) (Labrou et al., 1999). The thoughts of ACL are originated from interaction of human society. It does not specify any low level service like network protocol or transmitting services. It assumes the physical infrastructure exists. The characteristics of ACL including: defining goal driven behavior, defining autonomous decision making processes for agent actions, building interaction protocol through negotiation and delegation, creating models of mind states, and enhancing the adaptability to environments and requirements. In this research, ACL is followed as the communication language among agents.

2.3 Multi-Agent System

2.3.1 Coordination and Cooperation

Coordination and cooperation is the core research topics for multi-agent systems. It is also the major difference between multi-agent systems and the other types of computer systems. The purpose of build system in a multi-agent way is to coordinate the knowledge, desire, intention, planning, and actions so that the agents can cooperate to complete the tasks.

Coordination is to change the intention of the agents. Cooperation is a special case of coordination. In an open dynamic environment, agents with different goals need to coordinate its goals and resources. If there is a conflict in using resources, without a good coordination, there will be a dead lock. In another scenario, if a single agent cannot finish its goal, it may need helps form other agents, this requires coordination.

Cooperation can improve the performance of the multi-agent system and enhance its capability to solve problems. It also improves the flexibility of the system. While individual agents are independent from other agents and only care about its own needs and goals, when in a multi-agent system, agents must fol-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

low the predefined social rules and its behaviors must meet the preset social norms. The interdependency among agents strictly restricts the interaction and cooperation of agents. Based on different mechanisms of interaction and cooperation, the implementation of agents in a multi-agent system is different.

Nowadays, the researches on agent cooperation mainly fall into two types. The first type is to migrate the theories into other domains (Game theory, classical dynamics) for agents (Kraus, 1997). The second type is to research the cooperation from the idea of mind attitude as goals, intension, planning like model named as FA/C (Kraus, 1997), joint intention framework (Levesque, 1990; Wooldridge & Jennings., 1994), and shared planning (Grosz & Kraus, 1996). The first type of mechanisms is not as widely used as the second because once the environment changes, the number of agents, types, and interaction relationships might be different from the situations that are suitable for the theories. The theory based interaction thus loses its advantages. The second type focuses on problem planning and solving.

2.3.2 Contract Net

Contract net (Smith, 1980) is widely used in coordination of multi-agent systems. The communications among agents are usually built on agreed message formats. Contract net is actually a contract or a protocol that is based on formats of contract net. The contract net specifies task assignments and roles of agents. It is made up by task processor, contract processor, communication processor and local database.

The local database contains the nodes related knowledge base, information about the current states, and problem solving processes. The other three components use the local database to execute their tasks. The communication processor is responsible for communicating with other nodes. All nodes connect to the network through this processor. The contract processor receives the tasks from bidding. It sends application and finish jobs. It also explains and analyzes the received information. The contract processor executes the coordination among nodes. The task processor is responsible for processing and solving the assigned tasks.

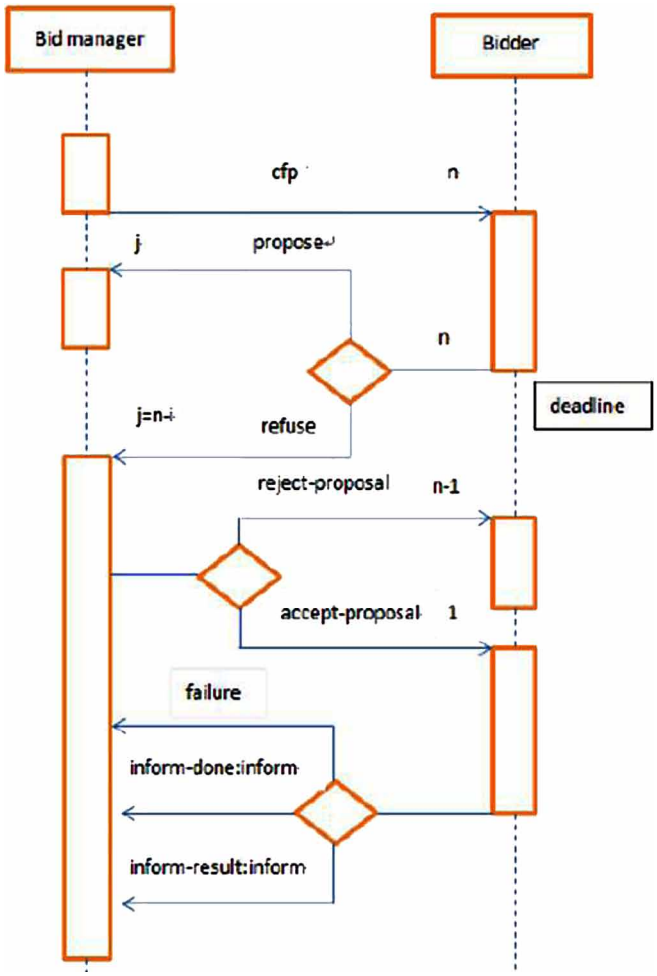
The working process of the contract net is as follows: the problem is decomposed into several sub problems. A node is chosen as the manager to manage the bidding processes. The manager encapsulates the sub problems as bids and publishes them to the nodes. The bids are open to all agents and can be processed through contract processor. The contract processor decides whether to accept the bid. If it accepts, it will inform the manager. The manager chooses the appropriate sub problem for the agent. The agent sends an acknowledge confirming it has accepted the contract. When the processing or solving is completed, the result is sent back to the manager. The essence of the contact net is task distribution. Figure 1 illustrates the process of the contract net with a sequence diagram.

2.3.3 Cooperation Planning

When an agent believes that cooperation bring benefits like improving efficiency or finishing unavailable tasks for single agent, it will have the desire of cooperating and looking for partner to do it. Or when multiple agents find that they can achieve bigger goals, they will form federations and take cooperative actions. The principle of cooperation is setting roles on demand and acquiring roles through competition. The roles are preset according to the goals and the needs of cooperation. The agents play specific roles only when they win in the competition attended by multiple agents.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. How a contract networks



The process of cooperation can be divided into six steps:

- Set goals and produce requirements;
- Cooperation planning and decide cooperation structure;
- Search for cooperating partners;
- Choose a cooperation solution;
- Realize goals;
- Evaluate results.

There are many formal methods to describe the cooperation process. Fox example, d'Inverno uses Z notation (Luck & d'Inverno, 1995), Fisher use sequential logic (Fisher & Wooldride, 1997), Kinny uses modal logic base on derived channel (Kinny et al, 1996). Although the logic based formal method can accurately describe the mind attitude and time series for agents, the action norms depicted by logics is still far away from implementation. The logic method is good at depicting the static properties and so-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

cial norms, but it is not good at depicting communication and interaction structure among agents. Some researchers use process calculus in process algebra such as calculus to describe the dynamic behavior processes, concurrent interactions and cooperating structures. Some researcher combines the formal methods and process calculus into a uniform formal structure (Werger, 1999). Using sequential logic depicts goals of cooperation while using calculus depicts interaction processes and cooperation structures.

2.3.4 Ecology Based Coordination

The distributed computing system has similar social, organic characteristics as living beings. Such open system is different from the current computing system. It can do asynchronous computing against complex tasks; it can generate nodes in different types of machines. Each node may make local decision according incomplete knowledge and delayed information. There is no central control; instead, it uses interaction among nodes to cooperate to solve problems. These characteristics form a concurrent composition. Their interactions, policies and competitions for resources are similar to pure ecology.

Ecology computing looks the computing systems as an ecosystem. It introduces many biological mechanisms such as mutation and evolution. The mutation leads to genetic variation and create diversities of species with stronger adaptation. The policies of mutation can be a method to improve the ability of agents in artificial intelligence.

Lenat and Brown (Lenat & Brown, 1984) introduce mutation mechanism into their AM and Eurisoko system through grammar mutating in Lisp programs to discover mathematical concepts. They believe the future successful systems should be a series of social systems with abilities of evolution and self-organization. Miller and Drexler (Drexler, 1988) discuss models of evolution such as biological ecosystem, commodity market and points out the difference between ecosystems and computing systems. They also think that a direct computing market is the most ideal system model. Huberman and Hog propose and analyze the process of the dynamic game (Huberman & Hogg, 1995). When different processes interact to finish the computing tasks, if there are multiple optional strategies, the dynamic gradual process may turn into nonlinear oscillation and chaos. This means that there are no stable strategies of evolution for computing ecology systems. Rosenschein and Zlotkin (1994) propose to use static game theory to solve conflicts in multiple computing nodes with different goals (Parsons, 2002).

The most famous ecosystem model includes biological ecosystem model, species evolution model, economic model, and social model of scientific group. The intelligence for the ecosystem usually supersedes the individual intelligence.

2.4 Learning

The basic learning mechanism tries to transfer a successful behavior in one circumstance to another similar circumstance. Learning is a process to gain knowledge, accumulate knowledge, improve performance, discover order, and adapt to environments. There are four basic nodes in a learning system: learning node, execution node, knowledge base and environment. The environment provides outside information. The learning node (or learning algorithm) process the information provided by the environment (Panait, 2005). The knowledge base stores the information in some knowledge representation form. The execution node uses the knowledge from the knowledge base to accomplish some tasks and provides the feedback regarding the execution to the learning node. Learning improves the performance of the system.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The quality and level of information that is provided by the environment directly influences the ways and effects of learning. The level of information refers to the degree of abstracting information. The low level information refers to detailed, concrete and raw information while high level information refers to processed abstract information. The quality of information decides the difficulty of learning. If the information is incorrect, incomplete or imprecise, it is impossible to gain the right knowledge.

The learning node is actually learning methods and algorithms; it is also called the learner. There are many methods and algorithms for machine learning such as inductive learning, learning by analogy, analytic learning, genetic learning and connectionist learning. The knowledge base is place for storing knowledge. It should select the appropriate ways of representing knowledge and it should allow modification and extension. The execution node use knowledge to guide actions and to change environments. The results of the execution are provided to the learning node so that the learning node can be improved. Only through the effect of the environments can the learning algorithm be evaluated.

In agent learning, there are two streams: one is to improve performance of the system through agent learning; the other one is to improve learning efficacy and quality of knowledge through agent learning (Alonso & Noble, 2001).

2.4.1 Active Learning

When an agent works in an environment, the method of gaining experiences of the autonomous agent is active learning. It is different from passive learning. In passive learning, the teachers or domain experts need to choose groups of training data. In active learning, agents learn the effects of actions or plans and amend their domain knowledge. There two types of learning methods for domain knowledge. The first type is theory amendment system such as EITHER (Ourston & Mooney, 1994), OCCAM (Pazzani & Kibler, 1991), CLIPS-R (Pazzani et al., 1994). These systems are representatives of the passive learning. They have preselected one group of training data which is unrelated to the status in training. The second type is reinforcement learning such as Q-learning, classifiers, back propagation. These types of systems do not explicitly focus on the correctness of the learned domain knowledge. They care about the results of executions. The reinforcement learning is widely used. Agents make a plan based on their current knowledge. The execution of the plan produces results which become a feedback to the learning algorithm.

2.4.2 Case Based Reasoning

In case based reasoning, the situations or problems that the agent face are called target case. The memorized situations or problems are called source case. CBR is a strategy that let agent search source cases according to target cases and uses the source cases to help solve the target cases.

Cases are the fundamentals of knowledge representations in CBR (de Mantaras, 2001). Acquiring cases is easier than acquiring rules. This can greatly simplify the process of knowledge acquisition and reuse the experiences of past processes. Agents do not need to start everything from scratch so that the efficiency of processing new problems is improved. Past successful or failure experiences can provide hints to solving current problems and to avoid future failures.

There are two types of CBR in multi-agent systems: process oriented multi-agent CBR and task oriented multi-agent CBR. The process oriented CBR focuses on the processes. Each process can be implemented as an agent. The process oriented CBR is made up of five agents. The case maintenance agent is responsible for maintaining the case library; The case searching agent is responsible for search-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ing and matching the new case; The case modification agent is responsible for the transformation of the case solutions; The case reuse agent is responsible for dealing with the target case according to the transformed solution; The case storing agent is responsible for putting the new case into the case library.

Task oriented CBR focuses on the tasks instead of processes. It provides a CBR agent with multiple strategies to deal with complex tasks. According to the requirements of the specific tasks, it constructs and organizes agents with different functionalities. Every agent is a small CBR system doing specified tasks. Such agent is called a CBR agent. The functionalities of the agents are relatively independent to each other and these agents have their own unique capabilities. Multiple CBR agents can cooperate to finish complex tasks through the form of contract net.

2.5 Research Areas

Design new theoretical model for intelligent agents is one of the interesting research areas. Nowadays, researchers are trying to build better theoretical agent models. The models should be based on the autonomous agents and should find the balance between social interactivity and individual intelligence. Some researchers propose to build rational agents and multi-agent system made up of the rational agents. For example, the logic rationality based BDI model (Kinny et al., 1996); the behavior rationality based Rational Agent (Wooldridge, 2000); The MAS sociality based semantics of open information systems and economic methodologies using game theory and decision making theory.

Negotiation mechanisms and organizations are also important topics that attract many researchers. Another important stream is to research the MAS frameworks and MAS related methodologies for the opening dynamic environments. Such research covers negotiations, cooperation, and task allocation mechanisms, social rules, filtering strategies, behavior normalization and federations. An agent organization is a structured agent group with clear goals. In agent organization researches, the hotspots include the models for agent organization, the formation and evolution of the organization, rules of organizations and patterns of organizations. Sociality is an important aspect to indicate the social abilities of agents in MAS. The modern organization theory based agent sociality can become the foundation of the agent theories. It also stands for the research direction of agent theories and technologies.

Implementation tools and applications are the research areas that focus on how the theoretical models can be applied into the real world. The implementing technologies for Agent and MAS are also hot research topics. The current technologies are based upon the existing social organization theory and modeling theories. Many tools have been developed to support MAS including experimental test beds, integrated systems, blackboard framework and object oriented concurrent programming OBCP (Briot & Gasser, 1998). The most representative work in this area is the Agent Oriented Program (AOP) program designing (Bresciani et al., 2004). There are lots of applications with logically distributed problems or physically distributed problems. Nowadays, the major application domains include: language processing, industry manufacturing, organizational information systems, air traffic controls, concurrent engineering design, distributed sensor systems, transportation planning, monitoring and robots. With the high speed development of the Internet, the integration technology based on mobile agents and multi-agent also become new hot research spots.

Intelligent agent should have the ability of learning from their environment. Trust and reputation are important mechanisms that support coordination and cooperation. 22 If agents can evaluate the trust or reputation of another agent, the performance and reliabilities of the coordination and cooperation can be greatly enhanced. If the agent can learn the trust and reputation through its past interaction experiences,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

the agent can then dynamically adjust its belief about other agents in the fast changing environments. Thus the framework, mechanisms and algorithms that the research proposes contributes both to the agent coordination and agent learning. The trust learning algorithms give agents learning ability while the reputation mechanisms improve the quality of coordination and cooperation.

3. EXISTING SYSTEMS

3.1 Method Employed

The investigation is carried out through viewing and analyzing conference papers, journal papers and technical reports on computational trust from varieties of sources. The criteria of choosing analyzing target are based on whether the model is representative and whether the model reflects the latest trend of the research. To achieve a comparative result, the first thing is to figure out the common basic compositional elements of different models. Taking the elements as parameters then compares the above models in the form of table; it will also try to find out the significance of the result table.

The results will be organized in the form of spider graph to show the reader a clear relationship between the evaluation criteria. The study will also differentiate research paths of different models through quantitative analysis and statistics.

3.2 Investigation Results

Results are shown from two approaches, i.e. centralized approach and distributed approach. The centralized approach saves all the rating procedure, storage of reputation, query of reputation, searching of comments to the computer server, while the distributed approach finishes all these jobs by agents themselves.

3.2.1 Centralized Approach

There are two existing models that take the centralized approach. They are eBay (Ebay, 2012) and SPORAS (Zacharia, 1999). eBay has built a feedback reputation system for its customer-to-Customer web-sites. The goal of designing such a system is to transfer the trust and reputation mechanism in the real life human market to the internet-based e-Market. SPORAS is a reputation model was developed by Zacharia in MIT, 1999. It is an evolved version of the eBay's online reputation models. In this model, only the most recent rating between two users is considered. Another important characteristic is that users with very high reputation values experience much smaller rating changes after each update than users with a low reputation. SPORAS incorporates a measure of the reliability of the users' reputation based on the standard deviation of reputation values.

3.2.2 Distributed Approaches

The research lists five existing models that take the distributed approaches and analyze their characteristics respectively. They are Marsh's Model (Marsh, 1994a), RegreT (Sabater & Sierra, 2001), Referral Reputation (Yu & Singh, 2002, 2003), FIRE (Huynh, 2006; Huynh et al., 2006), and TRAVOS (Teacy, 2006; Teacy et al., 2005).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The pioneer work on computational trust model was done by Marsh (Marsh, 1994a) in 1994. Marsh thought that knowledge, utility, importance, risk, and perceived competence are important aspects related to trust. He defined three types of trust: dispositional trust, general trust and situational trust. The trust management provided by Marsh does not treat the collection of recommendations provided by other agents; he only models direct trust between two agents. The aspect of risk is dealt with explicitly based on costs and benefits of the considered engagement. The decision making is threshold based. Among other parameters, the cooperation threshold depends on the perceived risk and competence of the possible interaction partner. If the situational trust is above the value calculated for the cooperation threshold, cooperation will take place otherwise it will not. Furthermore, the decision making can be extended by the concept of reciprocity, i.e. it is expected to be compensated at some time, if one does another one a favor.

RegreT takes trust as a multi-facet concept and a combination of pieces of information (Sabater & Sierra, 2001). In RegreT, reputation is a combinatorial product of individual dimension, social dimension, and ontological dimension. The calculation of reputation is same as the calculation in individual dimension. The only difference is that all the reputation under each sub-ontological dimension should be summarized. The model deals with three dimensions of trust or reputation. The individual dimension is based on self-made experiences of an agent. The trust values are called direct trust or outcome reputation. The social dimension is based on third party information (witness reputation), the social relationships between agents (neighbor-hood reputation), and the social role of the agents (system reputation). The ontological dimension helps to transfer trust information between related contexts. For all trust values, a measurement of reliability is introduced, which depends on the number of past experience and expected experience (intimate level of interaction), and the variability of the ratings.

The underlying computational framework for Referral Reputation (Yu & Singh, 2002, 2003) is based on Dempster-Shafer theory (Dempster & Arthur, 1968). The model has a well-defined trust network to propagate the trust value from the witnesses, the proposed approach does not concern the uncertainty which surely occurs in the interaction and there is no risk management in this model.

The FIRE model (Huynh, 2006; Huynh et al., 2006) believes that most of the trust information source can be categorized into the four main sources: direct experience, witness information, role based rules and third party references. FIRE integrates all four sources of information and is able to provide trust metrics in a wide variety of situations. The reliability value bases on the rating reliability and deviation reliability to counteract the uncertainty due to instability of agents.

TRAVOS (Trust and Reputation model for Agent-based Virtual OrganizationS) is based on probability theory (Teacy, 2006; Teacy L, W. T., Patel, J., Jennings, N. R. & Luck, M, 2005) and Bayesian system (DeGroot & Schervish, 2002). The prominent feature of TRAVOS is that it takes confidence into account and the trustor makes decisions based upon the confidence level instead of the trust value. The reputation is not simply added to the direct trust value or confidence value, the introduction of reputation depends on the value of confidence level. If the agent can achieve the minimum confidence level through checking belief of its direct interaction, then it does not have to query other agents. But if the agent cannot find enough confidence, then the agent needs to seek more evidence from the witness.

The researcher needs to build a test-bed for implementing TROVOS model. Such test-bed is also useful for future evaluation of the researchers own model and other models. The test-bed can mimic a general market where agents carry on business by automatically standing in for humans. Each agent seizes some resources and cost specific amount of resources with time elapsing (charging energy). The lifetime of the agent ends when the agent owns nothing (starving to death). The goal of each agent is to

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

maximize their assets (resources) and keep itself alive as long as possible. Transaction is the only way to accumulate resources and avoid dying because the system is designed to make agents achieve more resources in transaction than the resources elapsing with time. There are two attitudes held by the agents: benevolence and malevolence. Under the attitude of benevolence, the agent tends to give full value to the other agent in a transaction; under the attitude of malevolence, the agent tends to give partial value to the other agent in a transaction.

Before any transaction, an agent should find out a target agent and evaluate the trustworthiness of that agent. To find out a target agent, the usual way is to use Contract Net Protocol, the agent broadcast its needs and the other capable agents who are willing to sell goods will response to the request listing their preferred price and quantities. However, in our design, to simplify the scenario and focus on the effect of trust model, the target agent is selected randomly by the simulation engine. The trustworthiness is calculated by the agent through combining the reputation from other agents and the experience accumulated in past transactions with the target agents. The evaluation engine is based on TRAVOS model. Agents must consult other agents through broadcasting request for reputation. If the agent has knowledge (experience) about the target agent then the agent should respond to the request and send their experience to originated agent. The originated agent then combine these experience with its own experience according to the TRAVOS model to decide whether to make transaction with the target agent.

- Dimension is to study what the sources of the trust values are.
- The semantics focus on the meaning behind the trust in case that the outcome from the trust is a composite product.
- The mathematical model employed in calculating trust or reputation is sometime called trust computation engine (Josang, Ismail & Boyd, 2007).Trust network is the study on the topology of organizing witness agents and the host agent.
- Uncertainty refers to the management of risk which monitors the accidental incident and environmental changing, and reliability which ensures that the trusted agent is reliable enough even though it is trustworthy that is based upon the result of mathematical calculation.

Table 1. Comparison: the difference among different computational trust models

Model Name	Dimension	Semantics	Architecture	Trust Network	Uncertainty	
					Reliability	Risk
eBay	Reputation	Single	Centralized	N/A	N/A	N/A
Marsh	Local trust	Single	Distributed	N/A	N/A	N/A
SPORAS	Reputation	Single	Centralized	N/A	Exogenous	N/A
Referal System	Local trust reputation	Single	Distributed	Directed graph With depth limit	N/A	N/A
Regret	Local trust Reputation	Ontology-based	Distributed	N/A	Endogenous	N/A
FIRE	Local trust Reputation Role-based trust Certified trust	Single	Distributed	Directed graph with depth limit	Exogenous	N/A
TRAVOS	Local trust reputation	Single	Distributed	N/A	Exogenous	N/A

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.2.3 Observations

Different models are compared in Table 1. Significant observations from the table can be listed as follows.

1. There are multiple facts (cardinality of dimensions) to forge the trust or reputation. 3 models are single dimension and 4 models are multiple dimensions.
2. 6 models presume the semantics behind trust is consistent to all agents except that RegreT adds an ontological dimension to deal with the semantic difference.
3. 5 models choose distributed architecture rather than centralized architecture.
4. 6 mathematical models are based on summation or product of different dimensions with selected weights representing their influences. TRAVOS is an exception which is based on Bayesian probability theory. (The evaluation criterion is the confidence level instead of the trust or reputation value).
5. 5 models, except Referal System and FIRE, don't take trust network and trust transitivity into account.

Their hypothesis is that trust propagates from the target witness to the host agent without any distortion or loss. The findings are: semantics, risks, trust network pose weak points for most models, whereas the dimension, architecture and mathematical models are intensively studied by researchers.

4. DISCUSSIONS

Most models have two basic dimensions: local trust and reputation. But they aren't the only dimensions that can be used to deduce trust. FIRE (Huynh, 2006; Huynh et al., 2006) introduces role-based reputation, which models the trust to the specific role in the society. This is similar to the situation that people always tend to trust some group of people with occupations like professor, doctor or police in real world society. FIRE (Huynh, 2006; Huynh et al., 2006) also has a design called "certified reputation" which is reputation collected by the trustee from its previous interactions. The truster then does not need to contact its acquaintances to know about the trustee thus the design improves the efficiency of communication. Some idea can be borrowed from social science study on trust; cultural trust and mechanism trust (Sztompka, 2000) are trust that may find their places in computational trust. Their potential hypothesis is that their trust is restricted to a predefined topic. For example, in eBay, trust (reputation) to the seller implicitly means the belief held by the buyer that the seller will send the correct product to the correct location within the right time frame (Ebay, 2012). However, this is an ambiguous semantic. Some sellers may believe that the correct product to the correct location with a little delay deserves the buyer's trust. Such gaps of semantics lead to disruption between buyers and sellers. In multi-agent systems, if two agents with different trust semantics meet, do they simply refuse to trust the other because of incompatible semantics or they need to build up a consensus through negotiation? If the semantics of trust can be adjusted, how often should such adjustment happen and to what extent? It is necessary to create an ontology built upon XML and RDF that allows systems to provide machine-readable semantic annotations for the trust of specific domain. Most models do not explain how the trust network work and what mechanism the trust transitivity is based. How the trust or reputation value is transferred? Does it simply keep the original value from the witnesses? Or does its value attenuate a long distance as same as what

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

happens in the transitivity of human reputation? The solutions to these questions are keys to implement a practical system. Theories about social network analysis in social science make sense to build algorithms for searching witnesses in the network of intelligent agents. The transitivity or propagation algorithm can also benefit from studying the social network analysis. Some researchers have already noticed the transitive trust and proposed their ideas (Josang & Pope, 2005).

None of the reviewed models introduce risk management to control the uncertainty due to the environmental influence or accidental incidents, although they do define reliability to counteract the uncertainty due to instability of individual agent. It is necessary to use risk evaluation to evaluate the risk associated with the prospective transaction under specific environmental facts. The risk here mainly means the risk derivate from the environmental influence or accidental incidents. For example, agent A trust a reliable agent B, but B's environment is instable (B's system often breaks down due to an accidental power off). Such type of risk can be called as environmental reliability. If agent A has assessed the risk of B's unstable environment additional to its reliability and trust, it will be more careful when making decision about whether to trust B or not.

An important part is not listed in the comparison table but deserves discussions. Most models stay as a theoretical model without performing a strict experiment in the real system. A few of them do have simple test but are not complete. A complete experiment to assess the trust model for MAS should at least pass the functionality test, the performance test and the security test. The functionality test focuses on whether the model supports heterogeneous agents to effectively cooperate with each other in different scenarios. The performance test focuses on measuring the efficiency of the model in the form of comparison with other models. The security test needs to measure whether the trust model enhances or does harm to the traditional security like authorization and authentication. The test environment should firstly be a simple but complete multi-agent system. Some tools of building multi-agent system can be applied to the experiment like JADE (Java Agent Development Environment), FIPA-OS, zeous, etc. The details about different tools can be seen in technical report from Gerstner Laboratory (Laboratory, 2005).

JADE (Bellifemine, Caire & Greenwood, 2007), FIPA-OS (coordinating team, 2012), and ZEUS (Nwana, Ndumu, Lee & Collis., 1999) are the most prominent and prevalent platforms that support MAS development. The similarities among the three platforms are: all of them are based on Java programming language, all of them are open-source project, and all of them claim to strictly follow the FIPA (Foundation for Intelligent Physical Agents) specification. Through comparing these tools, the following observations and conclusions can be made:

- JADE platform is a better choice for MAS development than FIPA-OS and ZEUS in FIPA compliancy, platform maturity, and non-technical concerns.
- Adopting XML as the specification message encoding, representation, and content language can be a good choice to alleviate interaction problems.
- It is necessary to extend the range of applicability of the agent mobility to low-end devices.
- Web Service enhanced agent can freely integrate with any system that supports Web Services; The combination of agent paradigm and the Web Service paradigm may form an intelligent service provision network to complement the traditional static service oriented architectures (SOA).
- It is necessary to introduce agent oriented software engineering process and agent based modeling tools such as AUML (Luke, 2006) into MAS platforms.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The continuity of development efforts is important for the maturity of a successful agent platform. For the MAS research community, it is possible to develop better tools and platforms to support the agent development through three routes. The first route is to update the FIPA specification with continuously absorbing new findings and new mechanisms, to incorporate with the existing widely adopted specifications such as XML and UML; the second route is to build more powerful application platform through strengthening the administration, monitoring, debugging and logging functionality, to exploit the successful development tools like eclipse, to incorporate Web Service and agent oriented software engineering; the third route is to maintain and enlarge the open source communities and let more researchers take part in and contribute to the development of platforms. A framework is designed for the computational trust and reputation. In the mental space, agents carry out trust learning through observing the results of actions. Such learning leads to the generation, increasing or decreasing of trustworthiness. The learning process is discussed later in details. Next to the mental space is the decision space where agents use trustworthiness derived in the learning process to make delegation decisions. It is combinatory decision making process in that the risk and utility evaluation are also included. The outcomes of specific transaction are constantly observed by the other agents. The observation will become the input of next round trustworthiness learning. Apart from the agents own experience, the trust from the other agents, reputation, is also part of the input of trust learning. The reputation from an organized social network and its propagation mechanism is also a research topic in future work. Most aforementioned trust and reputation models are mathematically based upon simplistic algebraic summation or explicit statistical deduction through counting the success or failure of historical transactions toward the target agents. Surely these methods can produce reasonable numerical results and then translate them into thresholds which help agents make decisions. However, the assumption that the agents always repeat the same transactions with the same target opponents is impossible in an ever changing complex multi-agent environment.

- First, the semantics are various among objects and it does not make sense to hold some invariable elements as formula to calculate the trustworthy or not trustworthy.
- Second, the count of success or failure of transaction is non-sense when the target of discussion is different. It is impossible to draw an equivalence between a successful coke transaction and a successful airplane transaction.
- Third, there is no sure clear border between trustworthiness and distrust-worthiness, that is, trustworthiness or its opposite should be a pattern generated from the repeated transactions instead of a simplistic value.

The above analysis leads the modeling of trust and reputation to the area of computational intelligence. Actually, trustworthiness is one kind of belief. The research on generation of trustworthiness is to model a specialized kind of belief for the computer agent. In the study of computational intelligence, several streams are popular and focused in recent years: neural networks, evolutionary computation, swarm intelligence and fuzzy systems. Since 1970s, neural network becomes one of the main research streams of computational intelligence. It is widely applied in machine learning, pattern recognition, and biological science. The reasons of choosing neural network as the basic calculating mechanism are listed as below:

- Instances are represented by many attribute-value pairs. The target function to be learned is defined over instances that can be described by a vector of predefined features.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- The target function output of the computational trust can be discrete-valued, real-valued or a vector of several real or discrete-valued attributes.
- Neural network learning is good choice for fast evaluation of the learned target function. Usually, the recognition of trustworthiness should be finished several times per second by the agents.
- The ability of humans to understand the learned target function is not important. The weighted learned by neural networks are often difficult for humans to interpret.

The input layer is composed by dynamically organized element. The elements are extracted from the semantic library according to the target object. For example, if the target object is digital camera, then the elements extracted maybe price, quality, guarantee and delivery. The neural network will gradually adjust its weights in a way of unsupervised learning.

5. SUMMARY

Here the background knowledge of intelligent agent and multi-agent systems has been reviewed. The difference between DPS and DAI has been explained with a detailed description. Then the intelligent agent is reviewed from its concept, structure and communication. The explanations cover the concepts and meanings of agent in both broad and narrow sense. The popular theoretical models are listed. The differences of these models are compared. The discussions are focused on rational agent and Belief, Desire, Intention model. The intelligent agents are the fundamentals of multi agents. The investigation also includes the coordination and cooperation mechanisms. The advantages and disadvantages of different mechanisms are analyzed. The agent learning topics is also covered. The rational of learning is explained and the common algorithms such as active learning and CBR are mentioned. Finally, some interest topics in the field are listed and stated how the related results are located in the background.

Plenty of interests have been attracted to the construction of computational trust from various research communities. Through analyzing and comprising these models, the research proposes that a complete computational trust model should at least have seven fundamental elements. Some conclusions are drawn as follows:

- The current trust dimensions are not enough to represent trust in multi-agent systems and new dimension can be modeled through introducing concepts in sociology.
- Agents from different domains must fill their semantic gaps through constructing ontology with XML and RDF.
- The searching algorithm for trust network and the propagation mechanism for trust network can be progressed through introducing techniques in social network analysis.
- Except for the reliability of target agents, agents also need to manage the risk (or environmental reliability) due to environmental changes or accidental incidents.
- A complete experimental platform which used to test the functionality, performance and security of computational trust model is a necessity.
- The future work of the research domain can also be naturally derived from the above conclusions.
- The dimension of trust can be extended based on the study of trust in sociology and psychology. It is necessary to categorize the scenarios where the dimension is appropriately used.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- The results in the research of semantic web can be used in the computational trust model. It is necessary to create an ontology built upon XML and RDF that allows systems to provide machine-readable semantic annotations for trust of specific domains.
- Theories about social analysis in social science can be used to build algorithms for searching witnesses in the network of intelligent agents. It is necessary to develop a trust transitivity or propagation mechanism which fulfills the requirement of different situations.
- There is still a need for a simplified, effective mathematical model which can be evaluated through appropriately constructing experiments.

REFERENCES

Cruces, A., & Arriaga, F. D. (2000). Reactive agent design for intelligent tutoring systems. *Cybernetics and Systems: An International Journal*, *31*(1), 1–47. doi:10.1080/019697200124900

Alonso, E., Dinverno, M., Kudenko, D., Luck, M., & Noble, J. (2001). Learning in multi-agent systems. *The Knowledge Engineering Review*, *16*(03), 277–284. doi:10.1017/S0269888901000170

Bates, J. (1994). The role of emotion in believable agents. *Communications of the ACM*, *37*(7), 122–125. doi:10.1145/176789.176803

Bates, J., Loyall, A. B., & Reilly, W. S. (1992). *An architecture for action, emotion, and social behavior.* Technical Report CMU-CS-92-144. School of Computer Science, CMU.

Bellifemine, F., Caire, G., & Greenwood, D. (2007). *Developming multi-agent systems with JADE*. John Wiley and Sons. doi:10.1002/9780470058411

Bresciani, P., Perini, A., Giorgini, P., Giunchiglia, F., & Mylopoulos, J. (2004). Tropos an agent-oriented software de- velopment methodology. *Autonomous Agents and Multi-Agent Systems*, *8*(3), 203–236. doi:10.1023/B:AGNT.0000018806.20944.ef

Briot, J. P., & Gasser, L. (1998). Agents and concurrent objects. *IEEE Concurrency*, *6*(4), 74–77. doi:10.1109/4434.736431

Castelfranchi, C., Dignum, F., Jonker, C. M., & Treur, J. (1999). Deliberate normative agents: Principles and architecture. *Lecture Notes in Computer Science, 1757*, 364–378. doi:10.1007/10719619_27

Cohen, P. R. & Levesque, H. J. (1990). Intention is choice with commitment. *Artificial Intelligence, 42*(3).

Coordinating Team. F.-O. (2012). *Fipa-os developers guide*. Retrieved from http://citeseer.ist.psu.edu/477218.html

Coulouris, G., Dollimore, J., & Kindberg, T. (2000). *Distributed systems: Concepts and design*. Addison-Wesley Longman Publishing Co., Inc.

Dempster, A. P. (1968). A generalization of bayesian inference. *Journal of the Royal Statistical Society. Series A (General)*, *30*, 205–247.

DeGroot, M., & Schervish, M. (2002). *Probability and statistics* (3rd ed.). Addison Wesley.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Eberhart, R. C., & Shi, Y. (2007). *Computational intelligence: concepts to implementations*. Morgan Kaufmann. doi:10.1016/B978-155860759-0/50002-0

Ebay. (2012). Retrieved from http://www.ebay.com

Hewitt, C. (1985). The challenge of open systems. *Byte*, *10*(4), 223–242.

Huberman, B. A., & Hogg, T. (1995). Communities of practice: Performance and evolution. *Computational & Mathematical Organization Theory*, *1*(1), 73–92. doi:10.1007/BF01307829

Huynh, T. D. (2006). *Trust and Reputation in Open Multi-agent Systems* (PhD thesis). Schools of Electronics and Computer Science, University of Southampton.

Huynh, T. D., Jennings, N. R., & Shadbolt, N. R. (2006). An integrated trust and reputation model for open multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, *13*(2), 119–154. doi:10.100710458-005-6825-4

Ferber, J. (1999). *Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence*. Addison-Wesley.

Josang, A., Ismail, R., & Boyd, C. (2007). A survey of trust and reputation systems for online service provision. *Decision Support Systems*, *43*(2), 618–644. doi:10.1016/j.dss.2005.05.019

Rosenschein, J. S., & Zlotkin, G. (1994). *Rules of Encounter: Designing Conventions for Automated Negotiation among Computers*. The MIT Press.

Josang, A., & Pope, R. S. (2005). *Semantic constraints for trust transitivity. APCCM 2005*. University of Newcastle.

Grosz, B., & Kraus, S. (1996). Collaborative plans for complex group actions. *Artificial Intelligence*, *86*(2), 269–357. doi:10.1016/0004-3702(95)00103-4

Drexler, K.E. (1988). Incentive engineering for computational resource management. In *The Ecology of Computation* (pp. 231-266). Academic Press.

Kinny, D., George, M., & Rao, A. (1996). A methodology and modelling technique for systems of bdi agents. *Lecture Notes in Computer Science*, *1038*, 56–71. doi:10.1007/BFb0031846

Kraus, S. (1997). Negotiation and cooperation in multi-agent environments. *Artificial Intelligence Journal*.

Labrou, Y., & Finin, T. (1997). *A proposal for a new kqml specification. TR-CS-9703*. UMBC.

Labrou, Y., Finin, T., & Peng, Y. (1999). IEEE intelligent systems and their applications. *IEEE Intelligent Systems & their Applications*, *14*(2), 45–52. doi:10.1109/5254.757631

Laboratory, G. (2005). Agent system development hands-on exercise, easss-05 (7th European agent systems summer school). Technical report, Utrecht.

Lenat, D. B., & Brown, J. S. (1984). Why am and eurisko appear to work. *Artificial Intelligence*, *23*(3), 269–294. doi:10.1016/0004-3702(84)90016-X

Levesque, H. J. (1990). All i know: A study in autoepistemic logic. *Artificial Intelligence*, *42*, 263–309.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Luke, T. W. (2006). *Agent-based trust and reputation in the context of inaccurate information sources* (PhD thesis). Schools of Electronics and Computer Science, University of Southampton.

Teacy L, W. T., Patel, J., Jennings, N. R. & Luck, M. (2005). Coping with inaccurate reputation sources: Experimental analysis of a probabilistic trust model. *AAMAS05*.

Panait, L., & Luke, S. (2005). Cooperative multi-agent learning: The state of the art. *Autonomous Agents and Multi-Agent Systems*, *11*(3), 387–434. doi:10.100710458-005-2631-2

Nwana, H., Ndumu, D., Lee, L., & Collis, J. (1999). Zeus: a toolkit and approach for building distributed multi-agent systems. *Proceedings of the Third International Conference on Autonomous Agents*, 360-361 10.1145/301136.301234

Marsh, S. (1994a). *Formalising Trust as a Computational Concept* (PhD thesis). Department of Mathematics and Computer Science, University of Stirling.

Marsh, S. (1994b). Trust in distributed artificial intelligence. *Lecture Notes in Computer Science*, *830*, 94–112. doi:10.1007/3-540-58266-5_6

Minsky, N. H., & Murata, T. (2004). On manageability and robustness of open multi-agent systems. *Lecture Notes in Computer Science*, *2940*, 189–206. doi:10.1007/978-3-540-24625-1_11

Moore, R. (1990). *A formal theory of knowledge and action', In formalizing common sense*. Ablex Publishing Corporation.

Ourston, D. & Mooney, R. J. (1994). Theory refinement combining analytical and empirical methods. *Artificial Intelligence, 66*(2), 273-309.

Pazzani, M., & Kibler, D. (1991). The utility of knowledge in inductive learning. *Machine Learning*, *9*(1), 57–94. doi:10.1007/BF00993254

Pazzani, M., Merz, C., Murphy, P., Ali, K., Hume, T., & Brunk, C. (1994). Reducing misclassication costs. *Proceedings of the Eleventh International Conference on Machine Learning*, 217-225.

Rothkopf, M. H. (2003). The future of e-markets. *Journal of Economic Literature*, *41*(1), 214.

Sabater, J., & Sierra, C. (2001). Regret: areputation model for gregarious societies. *Fourth workshop on deception fraud and trust in agent societies*, 61-70.

Singh, M. P. (1994). Multiagent systems: A theoretical framework for intention as, know how, and communications. *Lecture Notes in Artificial Intelligence, 799*.

Shoham, Y. (1992). Agent-oriented programming: An overview and summary of recent research. *Proc. of Artificial Intelligence*.

Shoham, Y. (1993). Agent-oriented programming. *Artificial Intelligence*, *60*, 51–92.

Simon, H. A. (1996). *The Science of Artificial* (3rd ed.). The MIT Press.

Smith, R. G. (1980). The contract net protocol: High-level communication and control in a distributed problem solver. *IEEE Transactions on Computers*, *C-29*(12), 1104–1113. doi:10.1109/TC.1980.1675516

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Simon, H. A. (1996). *The Science of Artificial* (3rd ed.). The MIT Press.

Simon Parsons, M. W. (2002). Game theory and decision theory in multi-agent systems. *Autonomous Agents and Multi-Agent Systems*, *5*(3), 243–254. doi:10.1023/A:1015575522401

Subrahmanian, V. S., Bonatti, P., Dix, J., Eiter, T., Kraus, S., Ozcan, F., & Ross, R. (2000). *Heterogeneous Agent Systems*. The MIT Press.

Werger, B. B. (1999). Cooperation without deliberation: A minimal behavior-based approach to multi-robot teams. *Artificial Intelligence*, *110*(2), 293–320. doi:10.1016/S0004-3702(99)00023-5

Wooldridge, M. J. (2002). *An introduction to multiagent systems*. John Wiley and Sons.

Wooldridge, M. J. (2000). *Reasoning about rational agents*. The MIT Press.

Wooldridge, M., & Jennings., N. R. (1994). Towards a theory of cooperative problem solving. *Proc. of Modelling Autonomous Agent in a Multi-Agent World*.

Zacharia, G., & Maes, P. (1999). Collaborative Reputation Mechanisms in Electronic Marketplaces. *Proc. 32nd Hawaii International Conf on System Sciences*.

Yu, B., & Singh, M. P. (2000). *A social mechanism of reputation management in electronic communities*. Cooperative Information Agents IV - The Future of Information. doi:10.1007/978-3-540-45012-2_15

This research was previously published in Examining Information Retrieval and Image Processing Paradigms in Multidisciplinary Contexts edited by Joan Lu and Qiang Xu; pages 274-295; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 2
Applications of AI in Financial System

Santosh Kumar
IMS UNISON University, India

Roopali Sharma
Birla Institute of Technology, India

ABSTRACT

Role of computers are widely accepted and well known in the domain of Finance. Artificial Intelligence(AI) methods are extensively used in field of computer science for providing solution of unpredictable event in a frequent changing environment with utilization of neural network. Professionals are using AI framework into every field for reducing human interference to get better result from few decades. The main objective of the chapter is to point out the techniques of AI utilized in field of finance in broader perspective. The purpose of this chapter is to analyze the background of AI in finance and its role in Finance Market mainly as investment decision analysis tool.

OBJECTIVES

Role of computers are widely accepted and well known in the domain of Finance. Artificial Intelligence(AI) methods are extensively used in field of computer science for providing solution of unpredicted event in a frequent changing environment with utilization of neural network. Professionals are using AI framework into every field for reducing human interference to get better result from few decades. The main objective of the chapter is to point out the techniques of AI utilized in field of finance in broader perspective. The purpose of this chapter is to analyze the background of AI in finance and its role in Finance Market, mainly in process of investment decision making from evolution and use of AI techniques in this respect.

DOI: 10.4018/978-1-7998-0951-7.ch002

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION OF ARTIFICIAL INTELLIGENCE

Few decades before, a computer program was only used for a numerical computation and computing the path of a bullet. Now a days, a computer programs are used in decision making in important decisions area, supported by a big database across global. As the tasks that computers perform are getting complex and intertwined with the daily life decisions, such behavior of decision making from the computer programs associate with intelligence. So these programs labeled as intelligence incorporated without human neural system, named Artificial Intelligence (AI). However experts are using AI and its concepts in every area of decision making still this branch is unexplored and not known to masses. AI as an academic discipline started from Dartmouth conference organized by John McCarthy from Stanford University and Marvin Minsky from MIT in the 1955 as "the science and engineering of making intelligent machines". Many researchers define AI, as branch of computer science concerned with making computers behave like humans. While some define AI as extension of human intelligence through use of computers.

Alan Turing undertook a test of a machine's ability to demonstrate intelligence, Turing Test, by judgment to distinguish between human and AI machine on natural language conversation .The AI approach can further divided into two major areas -Conventional AI and Computational AI .Conventional AI is based out on logic and rules to make decisions. Computational AI takes biological mechanisms. Computational AI are more advanced form and broadly used in decision making in economy and social sciences. Artificial Neural network (ANN), Expert system (ES), Fuzzy systems (FS), Genetic algorithms (GA) are examples of Computational AI used in field of Financial and Investment decision making.

We use AI every day, without knowing the facts, a simple conversational navigator's application like Google Voice for Android. Our smart phones are getting better day by day; remind us on simple things like birthday of friends, time to go to work, hands free searching and messaging and even telling us if we forget attachment in an email. Such automations are based on concept and application of AI. Currently IBM along with Google is working on AI research for creating a Super computer, WATSON that could change our future significantly. WATSON is learning everything related to human behavior, nature and other complexities in order to forecast and recommend on our future actions.

Concepts of AI has been used in every industry, from computers to pharmaceutical .After gain of trust in these field, professional are start using AI in financial services decision making as well.

HISTORY OF AI IN FIELD OF FINANCE

History of AI in Financial services domain can be traced back from Pacific National Bank of USA in 1987, set-up a Fraud Prevention Security Task force to check the unauthorized use of cards at ATM machines and vendor counters. The bank build a prototype on a regular office computer to perform as warning system for fraud prevention, modeled around the weather warning systems that warned before possible natural disasters like cyclones and earthquakes. This prototype was successful and paved the way for AI-based technology into consumer interactive products and services. Thus,AI has found a new home in financial industries and used as a valuable tool to many business applications for decision making.

In 2001, Wednesday, 8 August, BBC has flashed a new "Robots beat humans in trading battle". Technology giant IBM pitted robotic commodity trading agents, known as "bots", in opposition to humans in trading Commodities, pork bellies and gold. The bots made 7% more return than the human traders, reported in New Scientist magazine.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

A book titled "Foundation of Investment Systems Using Artificial Intelligence and the Web" authored by Robert R. Trippi (2002), identify AI as Wall Street's most promising technologies. AI enabled systems can handle more information and react more quickly for consistent decisions than a group of human. Authors methodically explain how AI systems can help in higher investment returns, with real-life examples.

Sycara, Decker, Pannu, Williamson, & Zeng,(1996), researchers at Carnegie Mellon University, designed an integrated multi agent system that considers data of asset valuation, and risk management, named WARREN, which refers to the first name of the famous investor Warren Buffet .

The recent literature by Tsakonas, Dounias, Doumpos, & Zopounidis(2006) highlighted the role of neural network in investment by combination of knowledge based techniques, neural network and techniques of genetic algorithm .Bhattacharyya, Pictet, & Zumbach,(2002) have used knowledge-rich operators for investing in foreign exchange markets.

Many literatures supporting benefits of AI in investment and financial market has been published in last 2 decades. Thus different concept and use AI in different segments of finance market has been evolved. Kun Chang Lee, Namho Chung & Inwon Kang (2008), M. Uther & H. Haley (2008), Shun-Yao Tseng (2012), Rustam Vahidov and Gregory E. Kersten etc are the major contributors in the research of this field.

APPLICATION OF AI IN FINANCIAL MARKET

Accurate and effective decisions are required in constantly changing financial market condition. There is a massive increase in use of information technologies for accurate decisions making in finance field. Conventional statistical methods are regularly complemented by application of machine learning in recent times. The area of applied machine learning also helps in discovering hidden facts in huge amounts of data. The major techniques of AI utilized in finance Markets are as follows.

Expert (Knowledge) Systems (ES)

An expert system (ES) is a AI enabled computer-based system that emulates the logic process of an expert within a particular domain of knowledge. ES are designed on explicitly formulated specific knowledge gathered from experts for decision making as explained by Feigenbaum et al.(1988). The aim of an ES is not to copy the mental processes of experts during a decision making process but attainment of high quality decision. ES have capacity to give recommendation even the sufficient required data is un- available by using database of alternative inferences.

The research by Armstrong and Collopy (1992), developed a role model for making an annual extrapolation forecasts for demographic and economic time series on data.

This rule based ES produced more accurate forecasts than the traditional random walk and statistical tools combined. They performed better, especially in long periods for series with major trends, high uncertainty and low stability.

The comparative study by Korczak and Lipinsky (2004) between two real time trading systems by use of Stochastic oscillators based on relative Strength index & Ease of Movement. The first trading system was based on 350 trading rules and other on 150 trading rules created in linear combination on data of Paris Stock Exchange . They concluded that reduction in trading rules decreased the computation time without change on expertise quality. This highlights the main advantages of ES as they combined knowledge from different sources for decision making process.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Artificial Neural Networks (ANNs)

Artificial neural networks (ANNs) are combination of simple neural elements in parallel arrangement and mesh like a human nervous system likened data source with explicit knowledge about target values . The functioning of ANNs is largely depend on values in the network, widely used to solve problems of prediction, classification, and control by adjusting the weights between connected elements

The major advantage of ANNs is their capabilities to capture nonlinearity without predefine information about relationships between variables. ANNs operate as "universal approximation systems" with the skill to imitate almost any function related to variables. In comparison to traditionally used econometric models ANNs provide results in very short period. The drawback in ANNs based applications is the non existences of a standard pattern to design the network .ANNs have been successfully applied to solve problems of generalization in prediction of corporate bond rating. ANNs are used by central banks across globe to forecast the interest rates and the monetary policy.

The result of an empirical research shown in Table 1, depicts the association between forecasting accuracy of ANN and ARIMA model using Mean Square Error(MSE) on data of the USA GDP deflator, 1st quarter 1960 to 3rdquarter 2003. The performance of ANN better than AIRMA models in the first two quarters on test set, while the advantage disappeared for longer horizon due to early stopping approach .

There has been abundance of studies attempting to forecast the price levels in exchange by using indicators of technical analysis as input variables in ANNs framework (Egeli et al., 2003); (Gençay, r.; Stengos, t., 1998); (Safer, 2003). (Chen et al., 2008); (Mizuno et al., 1998); (Nagarajan et al., 2005); (Yao et al., 2000). The results show that indicators as inputs into ANN framework improve the forecasting accuracy of stock prices. Figure 1 shows the data transformation process in ANN framework applied on Japanese stock exchange.

Table 1. Ratio of ANN and AIRMA

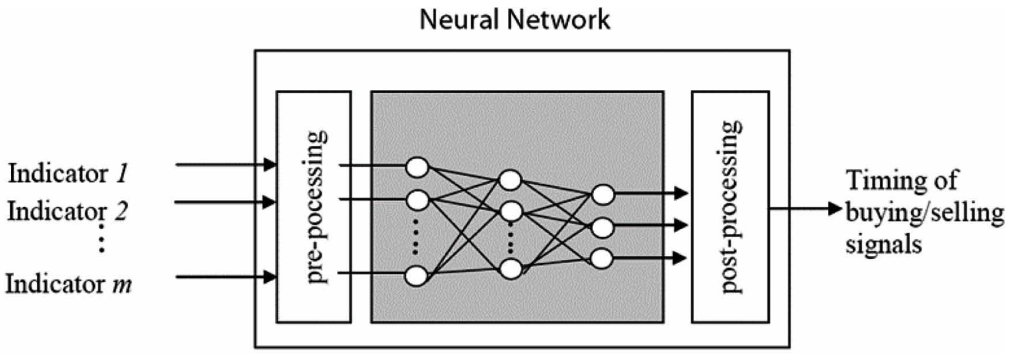
	Forecast Horizon (quarters)					Forecast Horizon (quarters)			
	1	2	3	4		1	2	3	4
Ar1	0.84	0.77	0.98	1.20		0.85	0.77	0.77	0.72
Ar2	0.89	0.84	1.04	1.18		0.86	0.81	0.81	0.73
Ar3	0.90	0.83	1.01	1.15		0.90	0.85	0.81	0.73
Ar4	0.90	0.82	0.96	1.12		0.90	0.85	0.83	0.72
Ar5	0.93	0.80	0.94	1.07		0.90	0.87	0.82	0.72
Ar6	0.86	0.79	0.91	1.04		0.92	0.86	0.82	0.72
Ar7	0.77	0.75	0.84	0.98		0.94	0.87	0.84	0.74
Ar8	0.77	0.73	0.83	0.91		0.93	0.88	0.84	0.76

Note: Each cell shows the ratio of the Mean Square Error of the Neural Network model to the Mean Squared Error to the Auto Regression model.

Source: Nakamura, 2005.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. System composed of ANNs and indicators of technical analysis as inputs into network



Source: Mizuno et al., 1998.

Fuzzy Systems

Fuzzy systems are a knowledge based expert system designed on Fuzzy logic concept. Fuzzy Logic calculates uncertainty from membership values in between the range of 0.0 to 1.0 with 0.0 indicates absolute failure and 1.0 represents absolute Success.

The fuzzy system is made up of three steps - Fuzzification, fuzzy inference and defuzzification . Fuzzification is the conversion of real to fuzzy data, and defuzzification is transformation of fuzzy data to real data after inference by fuzzy system, usually based on IF…THEN rules.

A fuzzy rule based neural networks have been used to predict stock market returns many times in various research. A stock trading system proposed by Ang and Quek (2006) on moving average rules achieved higher return than a traditional random trading system. Fuzzy neural network (FNN) based on general stock trading rules developed by Wong et.al achieved better return single ANNs and other traditional trading system in Taiwan stock Exchange.

Agent-Based Computational Economics (ACE)

Agent based Computational Economics (ACE) is an example of AI enable agent (variables) based model, suitable for economy modeled based forecasting . Economics can be study as a complex adaptive system to adapt changes in the external agents (environment). Bruun (2006), ACE is a bottom up system forming models of agents and interaction rules which lead to setting up the whole system using simulations within this modeling approach the structure evolves without forcing any presumptions upon it. The resulting ACE models must be dynamically create a economic system solely on the basis of agents' interactions without further intervention from the other modeler.

Early researches of artificial agents are based on zero intelligent agents. Le Baron (2000) investigated the efficiency of zero intelligence traders on artificial foreign exchange markets and he found that budget allocation of these agents was similar to human agents.

In other studies simulations with more sophisticated agents have evolved. Raberto et al. (2003) set up a model where trading strategies of agents were based on technical trading rules and fundamental values. Kendall and Su (2003) utilized agent based approach on five selected stocks from London Stock Exchange, trading behavior of artificial agents was based on indicators of technical analysis. This research

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

demonstrated stable and satisfactory learning abilities of artificial traders, different learning behavior related to the different stock price patterns was discovered, an important finding for further research in portfolio selection. Kumar and Bhattacharya (2009) used multi agent approach for portfolio selection and they achieved higher average returns across one month, two months and three months out-of-sample period than the FTSE 100 index.

The relative disadvantage of ACE modeling is that it requires detailed specification for agent data and methods determining structural attributes, behavioral dispositions and institutional arrangements. If the agent interactions induce sufficiently strong positive feedbacks, small changes in initial specifications could radically affect the types of outcomes that result.

COMPARATIVE STUDY IN PORTFOLIO PERFORMANCE

The investment and portfolio domain is changing in stochastic and random environment. Take the stock exchange as an example; there are more than five thousand securities available for selection by portfolio manager or individual investors. This creates a problem of filtering all those stocks to find the ones that are good for investment. There is also availability of enormous information that may affect the performance in market to some extent, as stock market is efficient. All above is making tremendously difficult for a portfolio manager to create the portfolio without relying on any tools.

An experiment has been conducted by department of commerce, Texas A&M University. In the experiments, commonly used AI systems - Bayesian network system, and a feed forward neural network system as shown in figure 2. The data of eight financial ratio mentioned below of the S&P 500 companies as the input variables these systems, from the period of 1987 to 1996.

- Debt to equity
- Market /book value
- EBIT
- Return on Equity (ROE)
- Pre-tax profit
- Total asset turn over (TAT)
- Financial leverage ratio (FLR) and
- Beta (β)

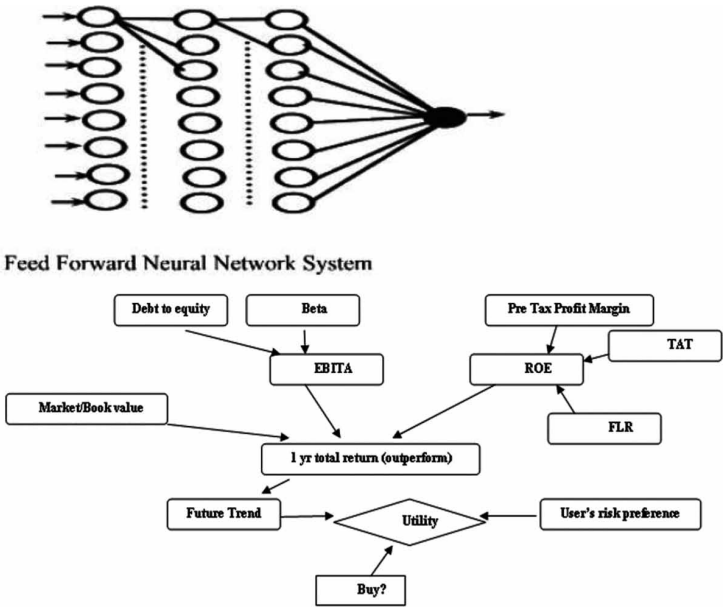
To test the performance, all these systems made the decision proposal on which of the S&P 500 stocks to be included in the portfolio. Compare the return output from selected portfolio by these AI system with return obtained from mutual funds, human portfolio selection, in same period.

Return from both AI systems outperforms the leading mutual fund by a large margin in 1997. They also ran these AI systems with the data from a negative return year (2000); the S&P 500 and the primary mutual fund houses produced negative return for the year while both AI systems produced a positive return.

By above experiment on real time data, we could conclude that the use of AI system for security selection in portfolio outperformed over the human portfolio manager.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Artificial Intelligence systems for portfolio selection



SUMMARY

The chapter offers supportive evidence for suitability of use AI enabled systems in various financial applications according to their capacity to process non-linear variable relationships, ability to learn and evolve in time and also ability to make decisions at expert level.

AI algorithms are used for optimization problems - optimization of stock market timing and portfolio creation. These tools can be used for creation of quantitative tools for short-time prediction of exchange, inflation and interest rates. Expert systems try to make decisions on the expert level and they are applied in analysis of securities as well as in company assessment. If we have an uncertain notion of the methods for a decision (what is typical for human decision makers), fuzzy systems could be the appropriate solution. In last decades the Agent-based Computational Economics has emerged. Current ACE concentrates on financial market modeling. The approach of artificial neural networks is put into effect due to their ability to study the nonlinear relation between variables and their ability to work with uncertainty. They are frequently used for solving prediction problems, forecasting macroeconomic indicators and time series prediction on financial markets.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Algoet, P. H., & Cover, T. M. (1991). Asymptotic optimality and asymptotic equip-partition properties of log-optimum investment. *Annals of Probability*, 876–898.

Chen, A., Hsu, Y., & Hu, K. A (2008). Hybrid Forecasting Model for Foreign Exchange rate Based on a Multi-neural network. *ICNC´08: Fourth International Conference on Natural Computation*. Jinan: IEEE.

Cover, T. M. (1991). Universal portfolios. *Mathematical Finance*, *1*(1), 1–29. doi:10.1111/j.1467-9965.1991.tb00002.x

Trippi, R. R., & Lee, J. K. (1996). *Artificial Intelligence in Finance and Investing*. Irwin.

This research was previously published in Maximizing Business Performance and Efficiency Through Intelligent Systems edited by Anukrati Sharma and Om Prakash Rishi; pages 101-109; copyright year 2017 by Business Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 3
Aspect-Based Sentiment Analysis of Online Product Reviews

Vinod Kumar Mishra
Bipin Tripathi Kumaon Institute of Technology, India

Himanshu Tiruwa
Bipin Tripathi Kumaon Institute of Technology, India

ABSTRACT

Sentiment analysis is a part of computational linguistics concerned with extracting sentiment and emotion from text. It is also considered as a task of natural language processing and data mining. Sentiment analysis mainly concentrate on identifying whether a given text is subjective or objective and if it is subjective, then whether it is negative, positive or neutral. This chapter provide an overview of aspect based sentiment analysis with current and future trend of research on aspect based sentiment analysis. This chapter also provide a aspect based sentiment analysis of online customer reviews of Nokia 6600. To perform aspect based classification we are using lexical approach on eclipse platform which classify the review as a positive, negative or neutral on the basis of features of product. The Sentiwordnet is used as a lexical resource to calculate the overall sentiment score of each sentence, pos tagger is used for part of speech tagging, frequency based method is used for extraction of the aspects/features and used negation handling for improving the accuracy of the system.

INTRODUCTION

Sentiment analysis is an area of text classification that starts in early of the last decade and has recently been receiving a lot of attention from the researchers and academicians. Sentiment analysis includes an analysis of data sets, such as online review, social media, blogs, and discussion groups, which contain opinions and opinion can be classify as positive, negative, or neutral. Opinion plays an important role in our information-collection behavior before taking a decision with applications in many fields, including recommender, customer intelligence, advertising systems and information retrieval.

DOI: 10.4018/978-1-7998-0951-7.ch003

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sentiment analysis can be also considered as the computational study of human opinions and sentiments, likes/dislikes, good\bad toward an entity such as products, services, organization, individual, issues, event, topics and their attributes. It is basically combination of natural language and data mining process that aims to extract user opinion and sentiment, expressed in positive, negative and neutral comments, by analyzing a large amount of unstructured user generated data. It is a task of automatically extracting the user opinion and sentiment from a plain text over the web. Sentiment analysis or opinion mining refers to the application of natural language processing, computational linguistics and text analysis to identify and extract subjective information from large amount of data.

Sentiment analysis also refers to identification and extraction of subjective information from text sources. It aims to determine the opinion of a person with respect to something in particular or the overall contextual polarity of a document. Sentiment analysis and its application in various disciplines; in information extraction, it is used to discard subjective information, in question-answering, it identifies opinion-oriented questions. Sentiment analysis is also used in the business intelligent and politics. Application areas of sentiment analysis are very vast and challenging. Sentiment analysis has focused on a great deal of attention as its very wide application in opinions detection, customer review summary, and other systems which is required to extract public opinions and sentiments. Some application areas of sentiment analysis are given in the next sub section.

APPLICATIONS OF SENTIMENT ANALYSIS

Sentiment occurs in almost all human activity because they describe the human behaviors. Whenever the human want to take a decision they want to know about others opinion. In real words, all business organization and all politicians want another opinion about their product and political condition even individual consumer also wants to know the opinion of the other user before purchasing any new product. In the past, whenever we want to take any decision we discuss it with our family member and friends for personel decisions and for taking any decision for an organizing we conducted the surveys. These traditional method are very time consuming and always does not give a good result.Application as a sub-component technology:Sentiment analysis plays an important role for enabling one technology for other system e.g. it cannot recommend those systems to be acquired with other if the systems contain more negative review. Detection of flame in the email and other type of communication is another use of sentiment classification. An online system, it helpful to take only the positive material in their system avoids all negative ads and content to be added to the online system and also helpful to the improve the online system and also helpful in the citation analysis, resulted increases the human computer interaction.

Application in Business and Government

Sentiment analysis is well suited in various intelligence applications. Business intelligence is one of the main application of sentiment analysis e.g. for any organization it is important to know what is the current position of their organization and for a businessman it is important to know the current position of their product and material in the market and due to the rapidly increase of web uses it is easy to analyze it with the help of user generated review and sentiment analysis is the best tool for it. Government intel-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ligence is the another application in which we can analyze the user opinion on the new policy started by the government allowing the automatic analysis of the opinions that people submit about pending policy or government-regulation proposals.

In this chapter we provide an overview of aspect based sentiment analysis with current and future trend of research. We gave a prototype algorithm for aspect based sentiment analysis and for better understanding of the topic we provide an experimentation for the aspect based sentiment analysis in which we use sentiwordnet as a lexical resource to calculate the overall sentiment score of each sentence, pos tagger is used for part of speech tagging and frequency based method is used for extraction of the aspects/ features. We also use negation handling for improving the accuracy of our system.

BACKGROUND

Now days the internet and web is used in every field of life. Due to the rise of internet and web its popularity among the peoples as a communication medium and business platform is increasing. People can very easily express their views on internet which generates very much amount of user data. Exponential growth of social networking has increased the amount of information generated over the internet on daily basis. The use of online shopping is also increasing as the vast amount of users is using internet to do shopping. The generated data of the views of customers is in unstructured form over the internet. The user generated content is the important source of extracting sentiment and opinions about the products. To extract the user opinion and sentiment from a large amount of data is very typical task. Many researchers have worked in this field of sentiment analysis in the last decade and still it's a very active research area. In last few years due to the rapid increase in the use of internet, plenty of work has been done in the field of sentiment analysis with special focus on aspect based sentiment analysis. Some important contributions are as follows:

- **Ghosh and Kar (2013)** has proposed a technique for sentence level sentiment classification of the online product review using the sentiwordnet a lexical resource. They semantically classify sentiment using rule based method and check the subjectivity and objectivity of an individual sentence and decided the semantic orientation using lexical resource.
- **Bhadane, et al. (2015)** gaves an overview of various type of method used for sentiment classification and techniques for measuring the opinions.
- **Varghese and Jayasree (2013)** has classified the unstructured user opinion text. They used sentiwordnet for classification, pos tagger for feature extraction, SVM classifier for quantitative classification and co reference resolution to resolve all the anaphora that appears in the opinion sentences. For the evaluation they checked the precision, recall and accuracy of a sentence and they achieved the 77.98% accuracy.
- **Chinsha and Shibily (2014)** have developed an aspect based opinion mining system for restaurant reviews, which automatically finds important aspects and opinions of a restaurant by analyzing the reviews. They used linguistic rule, sentiwordnet and two word phrases for sentiment detection with automatically extracted the aspect.
- **Singh, et al. (2013)** performed sentiment analysis using machine learning approach, unsupervised approach and two naive bayes approach and analytically showed that naive bayes approach gives better result in comparison to other two approaches. They also showed that semantic orientation

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

approach achieved good result but it need to calculate lots of PMI values which is very time consuming while sentiwordnet approach achieved lowest accuracy but are easy to use and does not require any training data set.

- **Nithya and Maheswari (2014)** performed a sentiment analysis of unstructured reviews for polarity and subjectivity detection. They used naive bayes classifier for identifying the feature and check their polarity. For feature extraction they used the pos tagger and for finding the polarity of a sentence they used a SentiStrength a lexical based classifier.
- **Taylor, et al. (2014)** extended the work of Bing Liu's aspect-based opinion mining approach. Bing Liu's approach is mainly focused on the physical product. They purposed a new NLP based rule for aspect level sentiment analysis which helps users to digest the vast availability of opinions in an easy manner.
- **Suchdev, et al. (2014)** classified the sentiment of the tweets using machine learning techniques and to increase the accuracy they used the symbolic technique with machine learning. They also used knowledge based approach and using backtracking method to measured the slag word frequency count.
- **Abulaish, et al. (2009)** presented an opinion mining system to classify the opinion and product feature from reviews. They used the semantic and linguistic approach for classifying the feature and used the sentiwordnet for classifying the polarity. They refine the rule based method to achieve the good accuracy and generate the feature based summary.
- **Gautam and Yadav (2014)** proposed a set of machine learning techniques with semantic analysis for classifying product reviews based on twitter data. They classified a large amount of reviews by using twitter dataset which are already labeled. They used naïve byes technique to improve the accuracy over maximum entropy and SVM technique and shows that the accuracy has improved from 88.2% to 89.9%.
- **Jo and Oh (2011)** proposed two models to discover aspects and sentiment in reviews. There models assume that one sentence contains exactly one aspect, they split sentences not only by punctuations but also by conjunctions. They use a part of-speech tagger for more accurate negation detection to make each aspect clear, filter out words that are common across many senti-aspects by adding a background language model. They show that the senti-aspects found by aspect based sentiment unification model (ASUM) can be used to illustrate direct sentiments toward the same aspect, in the evaluation of sentiment classification.

MAIN FOCUS

Sentiment Analysis is basically process of identifying whether a given text is subjective or objective and if it is subjective, then classify whether it is negative, positive or neutral. Sentiment analysis is basically study in three levels: document level, sentence level and entity\aspect level. In document level we identify the whole document as a positive, negative or neutral, in sentence level we analyze a particular sentence as a positive, negative and neutral and in entity\aspect level, which is also known as feature level, we analyze the feature of a particular product or item. Aspects\features can be understood with an example of restaurant review in which we can extract the aspects like its food, service, location, charges etc which described whole restaurant easily. Aspect based opinion mining deals with the extraction of opinion aspects by applying various techniques. Aspects are either in hidden form in the text or itself

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

appears in the text accordingly known as implicit and explicit aspects. Aspect level is very useful for taking fast decisions. In document and sentence level classification, each document/ sentence is classified as a single entity, a negative opinion document/sentence about an entity does not mean that the customer has negative opinions about all aspects of the entity. Likewise, a positive opinion document does not mean that the author is positive about everything. For effective and correct opinion about an entity, we need to classify the sentiment on the basis of different aspect of an entity.

Types of Opinions

The opinions can be categorizes as regular opinion and comparative opinion. A regular opinion is a simple opinion expressed in any document and it is mainly divided in two sub task namely direct opinion and indirect opinion. Direct opinion is expressed directly on an entity or entity aspect e.g. "the sound quality is great" describe directly the sound quality and give positive opinion. Indirect opinion is an opinion that is expressed indirectly on any product or document or aspect of product or document and based on its effects on some other entities e.g." after seeing the environment of restaurant my mood going worse" describe the bad environment of a restaurant which indirectly give the negative opinion. Most of the study focuses on direct opinion. Indirect opinion needs more analysis and complex task. Comparative opinion expresses a relation of similarities or differences between two or more entities and a preference of the opinion holder based on some shared aspects of the entities like nokia is better than micromax. Nokia is better shows two comparative opinions.

The opinions can be also classified on the basis of how they express in the text as explicit and implicit opinion. An explicit opinion is an opinionated statement that gives a regular or comparative opinion e.g. "nokia perform great" and "nokia is better than micromax". An implicit opinion is an objective statement that implies a regular or comparative opinion. Objective statement is the statements which cannot directly contain any opinion it only contain facts e.g. "The battery life of Nokia phones is longer than micromax phones".

Opinion and Target Relations

The opinions expressed in sentence have certain targets and they are clearly related to each other. Their relationships can be abused to find aspects, which are opinion targets, because opinion words are generally known for us e.g. "The battery is good". In this sentence the good is the opinion word and with the help of this it is easy to find an aspect of battery. This type of relation is also useful for finding the infrequent aspect. Several methods are used for better understanding of the relation between the opinion word and target e.g. dependency parsing and rule based method. The same opinion word can be used to classify or customize the different aspects. If a sentence does not have a frequent aspect but has some opinion words, the nearest noun/noun phrase to each opinion word is extracted. This method turns out to be very useful in practice even when it is applied alone. This is very useful relation in sentiment analysis because an aspect is rare to be important if nobody expresses any opinion or sentiment about it.

Aspect Level Classification

We can classify the sentiment at the aspect level, which we called aspect based sentiment analysis, it is also called feature based sentiment analysis. It covers both the entity and aspect to achieve the goal in aspect level; it discovers every quintuple in a given document. Aspect can be categories as an explicit

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

aspect or implicit aspect. Explicit aspects are directly present in the document e.g. “the size of the phone is not good” it directly shows the feature of a particular domain. Explicit aspects are only noun and noun phrase. Implicit aspect are hidden aspect in the document e.g. “the phone is not fit in my pocket” this show the indirect aspect size and it indirectly shows the feature of a particular domain.

- **Aspect Extraction:** Aspect is basically the noun and noun phrase so first we use the pos tagger to extract all nouns and noun phrase then we apply different approach to extract the aspect. Some of these approaches are:
- **Frequency Based Approach:** In this approach to extract the aspect first we used a pos tagger to extract the noun and noun phrases and check their occurrence frequencies and only the frequent ones are kept. Frequent one can be decided on the basis of a threshold value of the frequency and frequency threshold can be decided experimentally. This method work well in the medium size corpus while in large size corpus they extract a large number of frequent noun and noun phrases as a product aspect which may be not a genuine aspect.
- **Propagation Approach:** In this method the opinion lexicon expansion and target extraction tasks iteratively based on propagation using the dependency relations between the opinion word and target (aspect/feature). In this process initially we require only seed opinion lexicon and using the seed opinion lexicon we extracted new opinion words and targets for further target and opinion word extraction. This method adopts a rule based strategy, it follow a dependency rules to extract the opinion words and targets (aspect/feature) The propagation ends until no more new opinion words or targets can be identified.
- **Aspect Ranking Based Approach:** In aspect rankling approach, we arrange all aspects in the decreasing order of their frequency of occurrence and take only highest frequent aspect. The basic idea for aspect ranking is that, those aspects which most frequently talked and consider in the review has higher importance in the analysis and subsequently considered as high ranking aspect. We can also maintain a list of opinion words which are generally used to modify a particular feature to compute feature relevance. Relevant features are generally modified by various opinion words.
- **Other Approaches:** In 2007 a language model for aspect extraction came into existence and in this model it is assumed that the important aspect are used more often in the review in the same year and it is used as a syntactic pattern learned via pattern mining to extract the aspect. In 2008 author used a supervised technique to extract the aspect they treated target extraction as a topic co reference resolution problem. The key of their approach is use of cluster opinions, sharing the same target together. Another approach to extract aspect is the use of topic modeling. Topic modeling is to model the generations of a document set and mine the implied topics in the documents.

Opinion Aspect Pruning

During the extraction some incorrect opinion target and word are extracted so opinion aspect pruning is remove the noise beside the genuine target and opinion word, this is not essential only desirable. Currently no effective method for pruning is available. There are two types of noise occuring during the aspect extraction, one is the ordinary nouns that are not targets but are extracted as targets due to parsing errors and second is names of other competing products on which the reviewers also expressed opinions. The problem of extraction of incorrect opinion target can be avoided by doing pruning based

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

on clauses and pruning of other products and dealers. In pruning based on clauses method the sentence contains only one target unless there are any conjunction like “and” and “or”. If the sentence does not contain any conjunction and contain two aspects then we remove one of them. Filter non-aspect based on frequency and which is less frequent are removed and in pruning of other products and dealers method is used when comparing one aspect of a review product with the other.

Popular Techniques/Approaches for Sentiment Analysis

Machine learning approach, lexical based approach and hybrid approach are most popular approach/ techniques for sentiment analysis.

Machine Learning Approach

This approach basically works on machine learning algorithm to classify the sentiment analysis and uses the syntactic and linguistic feature (Medhat et al., 2014). In this approach there is a set of training records and each record is defined by a class and the classification model is exhibit one of the class label. This approach can be further classified as supervised learning approach and unsupervised learning approach. Supervised learning approach depends on the existing training data set and in supervised learning we need a labeled training and test data set for classification. Test and training data are usually product review. The major disadvantages of supervised machine learning approach is use of large labeled data for training which is very time consuming as well as this is also very difficult to create this large amount of labeled data. Unsupervised learning approach does not use the labeled data set and obviously the collection of unlabelled data is easier in comparison to the labeled data.

Lexicon-Based Approach

Lexicon based approach used a lexical resource for the classification process. In this approach we use the opinion lexicons and they are the collection of positive and negative opinion word express in the document. These are also called opinion phrase and idioms and together are called opinion lexicon. This approach is further classified as manual approach or automatic approaches. Dictionary based and corpus based approach are widely used automatic approaches. In dictionary based approach first we manually select the small set of opinion words and then these word are search for its synonyms and antonyms from WordNet or other corpa and repeat the process until no new opinion word are found and last manually check the error. Disadvantage of this approach is the inability to find opinion words with domain and context specific orientations. Corpus-based approach depends on syntactic pattern in which one set of opinion word is used to find the other opinion word in the large corpus. It also overcomes with the problem of finding opinion words with context specific orientations.

Features for Sentiment Analysis

Feature selection is the very essential task for the sentiment analysis in data driven approach. To convert a sentence in a feature vector is very important step for sentiment analysis.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Term Presence and Term Frequency

Term frequency is very essential part in the text classification and information retrieval task. Term presence is more important than term frequency for a sentiment analysis task. For indicating the occurrence of word we used a binary-valued feature vector, 1 for occurring and value 0 for not occurring, but for numerous sentences it countered as we seen many more frequent occurring sentence contain less information than the rare occurring sentence or word.

Term Position

Word appearing elsewhere contain less sentiment and weightage than the word contain in certain position. This is similar than information retrieval. In the review, the word contains negative sentiment throughout but the presence of positive word at the starting of the review play a very important role for determines the sentiment. Thus the sentence appearing in the first and last in sentence has more weightage than it appear elsewhere.

N-Gram Features

N-grams are able to capture the context up to some degree and are mostly applied to natural language processing task. Generally N-gram are of three types 1-unigram 2-bigram 3-trigram. In the movie review unigram perform better than bigram but doing some changes bigram and trigram give better performance.

Part of Speech

In most of the NLP task, part of speech information is most commonly used. They provide a raw form of word sense disambiguation; this is the most important reason of using part of speech. Pos tagger will be used for part of speech and with the help of part of speech we can extract all noun, adjective and adverb from a text document. Adjective and adverb are basically the opinion word and noun and noun phrases are feature and aspect of given product review. There is a strong correlation between the adjective and subjective. It achieved an accuracy of around 82.8% in movie review domains using only adjectives in movie review domains.

Topic-Oriented Features

Phrases and bag-of-words are broadly used as features. But in different domain a specific phrase value carry a small relation with whole text sentiment. Analyze the text which feat those aspects and which represent whole text is a challenging task in the sentiment analysis. With the use of phrases and bag-of–word will cannot differentiate the locally and globally meant of text in review and all of which are used in affluence in many converse domain.

Prototype Algorithm for Aspect Based Sentiment Analysis

Input: Labeled Data Set
Output: Positive and negative polarity of a particular feature in the data set

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Step 1: Pre-processing the review
- a. Remove unwanted text
- b. Remove stop words
- c. Stemming

Step 2: Find the feature \Aspects and opinion words
- a. Find noun, adjective, adverb and verb using pos tagger
- b. Noun\Noun phrases = Aspect\Feature
- c. Adjectives = opinion words

Step 3: Apply frequency based method for finding feature\aspect
- a. Check the frequency of noun in the review data set
- b. Take the most frequent noun as feature\aspect
- c. Ignore unfrequented nouns

Step 4: Check the polarity of the opinion word with respect to feature using sentiwordnet

Step 5: Match the words with review dataset
- a. If (positive polarity score > negative polarity score)
- b. Print ("the feature is positive")
- c. If (positive polarity score < negative polarity score)
- d. Print ("the feature is negative")

Step 6: Apply negation handling.
- a. Input – Negation word file which contain all negative words
- b. If (opinion word found in the sentence)
- c. If (opinion word is positive)
  - i. If (negation word found in the sentence)
  - ii. Print (polarity of a sentence should be negative)
  - iii. Else
  - iv. Print (polarity of a sentence should be positive)

End
- a. If (opinion word found in the sentence)
  - i. If (opinion word is negative)
  - ii. If (negation word found in the sentence)
  - iii. Print (polarity of a sentence should be positive)
  - iv. Else
  - v. Print (polarity of a sentence should be negative)

Detail Description of Terminology used in Algorithms

Review Sentence Tagging

A POS tagger parses a string of words and tags each word with its part of speech. The use of Stanford pos tagger for POS tagging is beneficial because of its good performance. It helps to tag the texts of the reviews without any need for training data. This tagger used by splitting text into sentences and to produce the part-of-speech tag for each word (whether the word is a noun, verb, adjective, etc) e.g. of pos tag sentence of the sentence ….this is really big step with quality of camera…. is….. this/DT is/VBZ really/RB big/JJ step/NN with/IN quality/NN of/DT camera/NN.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Aspect Extraction

The process of aspect extraction is described in following steps:

Step 1: From all tag sentences by pos tagger we collect all nouns and find the unique noun among the nouns.

Step 2: Retrieved the entire frequent nouns and ignore the unfrequented nouns. Then set the threshold value for frequent nouns and extract all nouns as an aspect/feature which is above the threshold value.

Sentiment Classification using Sentiwordnet

As we know sentiwordnet is an open lexical resource for sentiment classification in which all the score of the opinion word are store in the positive and negative way in the synset term. Each score are from 0.0 to 1.0 and there sum is 1.0 for each synset. To calculate the score of each sentence with the help of these score we calculate the score of all sentences which consist an aspect term and finally combine the score and if the positive score is grater then negative score then we consider the polarity of a feature is positive otherwise negative.

Negation Handling

Negation handling is an important task in the sentiment analysis. Without the explicit use of any negative word, negation can be expressed in different ways. In the sentence "i do not like this move" the negation handling can be done by reverse the polarity of all words appearing after negative word which is 'do not' in this sentence but it cannot work in the sentences like "I do not like the movie but I like the action". Therefore for negation handling we need to consider also the scope of negation as well. So in this case the thing that can be done is to change polarity of all words appearing after a negation word till another negation word appears. But still there can be problems, for example, in the sentence "Not only did I like the acting, but also the direction", the polarity is not reversed after "not" due to the presence of "only". So this type of combinations of "not" with other words e.g. "only" has to be kept in mind while designing the algorithm.

Performance Measuring Parameters

The performance of aspect based sentiment analysis (Figure 1) can be measured as the precision, recall and accuracy which are defined as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + FP + FN} \tag{3}$$

Where:

- TP is true positive. True positive means predict the sentence positive and it also really positive
- TN is true negative. True negative means predict the sentence negative and it also really negative
- FP is false positive. False positive means predict the sentence positive but it is not actually positive it is either negative or neutral
- FN is false negative. False negative means predict the sentence negative but it is not negative actually it is either positive or neutral.

Experimentation

We collected the reviews for the product (Mobile Nokia 6600) from the internet domain and we generated the data set from those reviews. Table 1 shows the data set for the product reviews of Nokia 6600. It contains number of total sentences from reviews and number of positive and negative sentences in the reviews. This is obtained after the process of finding the subjectivity of the sentences and splitting the sentence into positive and negative sentences.

Figure 1. Graphical representation of aspect based sentiment analysis

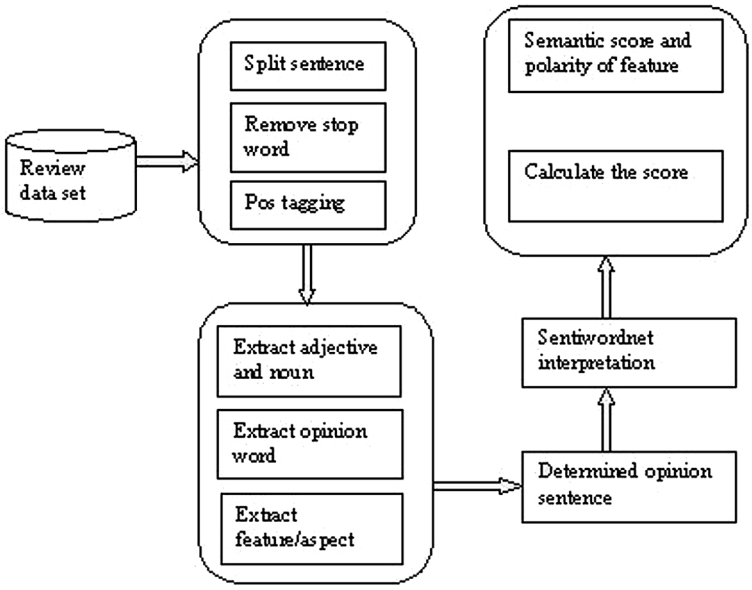


Table 1. Data set for Nokia 6600

	Total Sentence	Positive Sentence	Negative Sentence
Nokia 6600	546	250	118

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Scores and polarity of aspects/features

Aspect/Feature	Positive Score	Negative Score	Polarity
Battery	0.08	0.02	Positive
Feature	0.11	0.28	Negative
Nokia	0.12	0.18	Negative
Phone	0.02	0.05	Negative
Quality	0.00	0.00	Neutral
Radio	0.00	0.00	Neutral
Service	0.00	0.00	Neutral
Use	0.00	0.00	Neutral

We calculated the positive score, negative score and polarity of each aspect of products using sentiwordnet. The Table 2 shows the positive score, negative score and the polarity of the features.

From the Table 2 it is clear that the polarity of the product is negative on the basis of average scores of the aspects. To check the accuracy we calculate the precision and recall. For this task we firstly categorised the true positive, true negative, false positive and false negative from the review sentences then calculated the precision, recall and accuracy from the Equations 1, 2 and 3.

Table 3 shows the scores of the precision, recall and accuracy of the aspects.

Basic Requirement for Experimentation

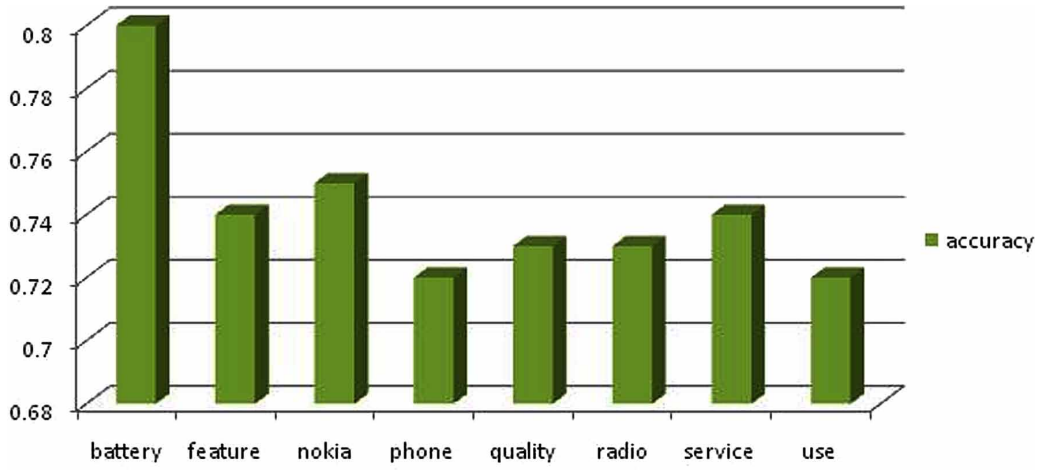
This experimentation has done on eclipse or Netbean platform which have a development interface for java programs. We use a Stanford parsing tool for finding adjective and noun/noun phrase and an open lexical resource sentiwordnet for obtaining the positive or negative scores of all opinion words in the form of synset terms. We also use the list of all positive and negative words which is very useful for splitting the review data in the positive and negative reviews.

Table 3. Precision, recall and accuracy of the aspects

Aspect/Feature	Precision	Recall	Accuracy
Battery	1.0000000	0.80000000	0.8000000
Feature	0.9736842	0.75510204	0.7400000
Nokia	0.9846154	0.76190480	0.7529412
Phone	0.9891892	0.72332020	0.7176471
Quality	0.9900498	0.73431736	0.7289377
Radio	0.9907407	0.74564460	0.7343174
Service	0.9911111	0.74581940	0.7408638
Use	0.9919028	0.72700300	0.7227139

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Accuracy of aspects



After calculating the accuracies of each aspect (Figure 2), we calculate the average accuracy for the product by applying the following formula:

$$\text{Average accuracy} = \frac{\sum_{n=1}^{n} score\,of\,Aspects}{total\,no\,of\,aspects} \tag{4}$$

The final accuracy of the aspects of the product Nokia 6600 is calculated as 74%.

FUTURE AND CURRENT TRENDS

Sentiment analysis is basically a classification task in three different levels: aspect level, document level, and sentence level. Sentiment analysis mainly concentrate on identifying whether a given text is subjective or objective and if it is subjective, then whether it is negative positive or neutral. Sentiment analysis is also considered as a subfield of computational linguistics concerned with extracting sentiment and emotion from text. The data set is the main issue in sentiment analysis and the main source of the data set is product review. Sentiment analysis is not important for only product, review it is also used in many other fields like the stock market, politics etc. The major challenges and scope for the researcher in the sentiment analysis are as follows.

Comparative Sentence Identification

Comparative sentence means where the comparison between two object is done e.g. “product a is better than product b” it completely change the polarity of a sentence. To clearly understand these comparative sentence, a well effective sequential information, modelling and structure are needed. Product review are mainly contain subjective information but comparative sentence contain both subjective and objective items. Thus to identify the comparative sentence is very difficult task because only few indicator will help to identify such sentence and those indicator also present in the sentence which are not comparative.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Leveraging Domain-Dependency

Almost all sentiment analysis and opinion mining research the domain dependency is an important issue. All approaches perform differently in different domain e.g. any approach can perform well in the product domain but may be not suitable for the politics feedback analysis. The one major and obvious reason is the difference in vocabularies across different domains. So designing the domain dependent model is also a very challenging task. Lots of work can be done in this field. In 2006 a reasearcher build a domain dependent classifier using a machine learning technique. In 2009 they build a query-specific subjective lexicon to improve the result but it also exploit some domain independent feature. Yang et al. (2006) propose a set of domain in dependent features and performed in the movie and product domain. There are several avenues open for the research in this field using the machine learning techniques that achieves good accuracy.

Association of Different Opinion

A document contains much opinionated information about many different topics in a single document. In this situation separation of opinionated information from factual information and look for opinion–topic associations in the documents need a very effective approach. To finding the association between the topic and corresponding opinion, various techniques can be used. Some of these techniques are: Natural language processing (NLP) techniques (like POSTagging and Syntactic Parsing) are used to identify opinion words and syntactic relation between the sentiment expression and subjective terms. Pos tagger is helpful to identify the opinion word as they identify all noun, adjective, verbs in the given piece of a document and the adjective is considered as a opinion word, Proximity-based techniques used to find the topic-association in textual document for example it say the proximity of the query terms to the opinionated sentences in the document helps to find that level of opinion–topic association necessary for opinion finding task. They just check the occurrence opinionated word around the query word.

Challenges in Natural Language Processing (NLP)

It is not forget that sentiment analysis is a NLP problem. It touches every aspect of NLP, e.g., negation handling, coreference resolution and word sense disambiguation, which combines more difficulties, since these are not solved problems in NLP. However, it is also useful to understand that sentiment analysis is a highly restricted NLP problem because the system does not need to fully understand the semantics of each sentence or document but only needs to understand some aspects of it, i.e., positive or negative sentiments and their target entities or topics. In this sense, sentiment analysis offers a great platform for NLP researchers to make tangible progresses on all fronts of NLP with the potential of making a huge practical impact.

Contextual Polarity of Words

An accurate polarity identification of a opinion word is still need a deep analysis. We cannot assign a polarity with respect to the prior polarity of a word. A polarity of a word is also affected by its surrounded word for example if the word “good” show the positive polarity the word “not good” it show the negative polarity because it contains a negation word. The new polarity define by its context is called contextual

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

polarity the context can be defined by the negation or by other ways. In 2003, Yi et al. used a lexicon and manually developed high-quality patterns to classify contextual polarity. Popescu and Etzioni (2005) use a relaxation labelling of an unsupervised classification technique to recognize the contextual polarity of words. They use the bootstrapping approach to classify the polarity of tuples. Negations (only not) has also taken into account when identifying contextual polarities

CONCLUSION

In this chapter first we provide an overview of sentiment analysis with special emphasis on aspect based sentiment analysis. A prototype algorithm has introduced for aspect based sentiment analysis and a case study has carried out for better understanding of aspect based sentiment analysis. For the aspect based sentiment analysis we have used the sentiwordnet 3.0 and for implementation of proposed system we used java programming language on eclipse platform. For the simulation and analysis of the system we used mobile (Nokia-6600) review data. The pos tagger, which does not need of trained data, has been used to tag the sentences. To remove the stop word and unwanted data, we pre-process the data and not using stemming because it put the distortion in review data. We checked the subjectivity of the sentence using the list of positive and negative word and used frequency based method for aspect extraction in which we checked the frequency of noun which is considered as aspect and take most frequent noun as a product aspect. We used sentiwordnet for semantic analysis of data and calculate the final score for classifying the polarity of each aspect/feature in terms of positive and negative. By using these techniques we obtain a good result in terms of precision, recall and accuracy. These techniques are very easy to use with respect to supervised machine learning techniques because it cannot need any training and testing data. This technique is also taking less time for computation in large data set because it cannot need to train a data. This system extract only the explicit features therefore in order to upgrade the system in future, we plan to develop some extraction technique which also extract the implicit feature of a product. The other possible extension of this system is hybrid approach i.e. by combining the sentiwordnet lexical resource method with the supervised machine learning method to improve the accuracy of this system.

REFERENCES

Abulaish, M., Jahiruddin, Doja, N. M., & Ahmad, T. (2009). Feature and opinion mining for customer review summarization. In Pattern Recognition and Machine Intelligence (LNCS), (vol. 5909, pp. 219-224). Springer-Verlag Berlin Heidelberg.

Bhadane, C., Dalal, H., & Doshi, H. (2015). Sentiment analysis: Measuring opinions. *Procedia Computer Science*, *45*, 808–814. doi:10.1016/j.procs.2015.03.159

Chinsha, T. C., & Shibily, J. (2014). Aspect based opinion mining from restaurant reviews. In *Proceedings of International Conference on Advanced Computing and Communication Techniques for High Performance Applications* (Vol. 1, pp. 1-4).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Gautam, G., & Yadav, D. (2014). Sentiment analysis of twitter data using machine learning approaches and semantic analysis. In *Proceedings of IEEE Seventh International Conference on Contemporary Computing* (pp. 437-442). 10.1109/IC3.2014.6897213

Ghosh, M., & Kar, A. (2013). Unsupervised linguistic approach for sentiment classification from online reviews using sentiwordnet 3.0. *International Journal of Engineering Research & Technology*, *2*(9), 55–60.

Jo, Y., & Oh, A. H. (2011). Aspect and sentiment unification model for online review analysis. In *Proceedings of the fourth ACM international conference on Web search and data mining* (pp. 815-824). 10.1145/1935826.1935932

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal.*, *5*(4), 1093–1113. doi:10.1016/j.asej.2014.04.011

Nithya, R., & Maheswari, D. (2014). Sentiment analysis on unstructured review. In *Proceeding of IEEE International Conference on Intelligent Computing Applications* (pp.367-371).

Popescu, A.-M., & Etzioni, O. (n.d.). Extracting product features and opinions from reviews. In *Proceedings of Conference on Empirical Methods in Natural Language Processing* (pp. 339-346). 10.3115/1220575.1220618

Singh, V. K., Piryani, R., Uddin, A., & Waila, P. (2013). Sentiment analysis of textual reviews. In *Proceedings of IEEE 5th International Conference on Knowledge and Smart Technology* (pp. 122-127).

Suchdev, R., Kotkar, P., Ravindran, R., & Swamy, S. (2014). Twitter sentiment analysis using machine learning and knowledge-based approach. *International Journal of Computers and Applications*, *103*(4), 36–40. doi:10.5120/18066-9006

Taylor, M. E., & Velásquez, J. D. (2014). A novel deterministic approach for aspect-based opinion mining in tourism products reviews. *Expert Systems with Applications*, *41*(17), 7764–7775. doi:10.1016/j.eswa.2014.05.045

Varghese, R., & Jayasree, M. (2013). Aspect based sentiment analysis using support vector machine classifier. In *Proceedings of IEEE International Conference on Advances in Computing, Communications and Informatics* (pp.1581-1586). 10.1109/ICACCI.2013.6637416

Yang, H., Si, L., & Callan, J. (2006). Knowledge transfer and opinion detection in the TREC2006 blog track. In *Proceedings of TREC*.

Yi, J., Nasukawa, T., Bunescu, R., & Niblack, W. (2003). Sentiment analyzer: Extracting sentiments about a given topic using natural language processing techniques. In *Proceedings of IEEE International Conference on Data Mining* (pp. 427-434).

ADDITIONAL READING

Liu, B. (2012). *Sentiment analysis and opinion mining, Synthesis Lectures on Human Language Technologies*. Morgan and Claypool Publishers.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

Bayesian Network (BN): Bayesian network iis the directed acyclic graph in which the nodes represent the random variable and the edge represent the conditional dependencies. it specified the joint probability distribution over all the variable.

Decision Tree Classifiers: In decision tree classifiers we provides the training data in the hierarchical form and for dividing the data we used the conditions on attributes and to perform the spilt we use the similarity based multi-attribute split.

Maximum Entropy Classifier: This classifier determine the most likely class for a document set it convert the labelled document set into a vector using encoding and with the help of encoded vector we calculate the weight of a document and combine to get the result.

Naive Bayes Classifier: It is a simple probability based classifier based on bayes theorem with strong independence assumption it compute the posterior probability of a class based on the distribution of the word in the document it ignore the position of the word in the document.

Neural Network (NN): Neural network is combination of neuron. Neural network is used for both classification and prediction. Neural network contain three layers for computation first is input layer second is hidden layer and last is output layer.

Rule-Based Classifiers: Rule based classifier is a sets of rule for classification and uses term absence and presence for the classification. In this classifier we define some criteria to generate the rule and theses rule are generates at the training time.

Support Vector Machine (SVM): It determine the linear separator by constructing hyperlanes that separates the cases that belong to different categories. Hyperlanes provide best separation between the class and it takes the maximum margin for separation because the normal distance between any data point is largest.

This research was previously published in the Handbook of Research on Advanced Data Mining Techniques and Applications for Business Intelligence edited by Shubhamoy Dey, Shrawan Kumar Trivedi, Anil Kumar, and Tapan Kumar Panda; pages 175-191; copyright year 2017 by Business Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Section 2

Development and Design Methodologies

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 4
Compliance with International Soft Law:
Is the Adoption of Soft Law Predictable?

Michael D'Rosario
Deakin University, Australia

John Zeleznikow
Victoria University, Australia

ABSTRACT

The present article considers the importance of legal system origin in compliance with 'international soft law,' or normative provisions contained in non-binding texts. The study considers key economic and governance metrics on national acceptance an implementation of the first Basle accord. Employing a data set of 70 countries, the present study considers the role of market forces and bilateral and multi-lateral pressures on implementation of soft law. There is little known about the role of legal system structure-related variables as factors moderating the implementation of multi-lateral agreements and international soft law, such as the 1988 accord. The present study extends upon research within the extant literature by employing a novel estimation method, a neural network modelling technique, with multi-layer perceptron artificial neural network (MPANN). Consistent with earlier studies, the article identifies a significant and positive effect associated with democratic systems and the implementation of the Basle accord. However, extending upon traditional estimation techniques, the study identifies the significance of savings rates and government effectiveness in determining implementation. Notably, the method is able to achieve a superior goodness of fit and predictive accuracy in determining implementation.

INTRODUCTION

The decision to implement international soft law maybe motivated by a number of different factors as noted by Ho (2002) the implementation of economic soft law, specifically in the case of the Basle Accord of 1998, appears to be driven by the strength of the underlying democratic system of governance in place within a particular jurisdiction affording credence to the democratic legalists theories of international

DOI: 10.4018/978-1-7998-0951-7.ch004

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

law. While the extant research does appear to support this proposition, there appears to be little, if any, credence afforded to the interaction between democracy and government effectiveness and their impact on the implementation of international soft law. Artificial intelligence methods have been largely absent from discussions of international soft law implementation. This appears problematic and potentially inapt given the plausible interaction effects that exist between government effectiveness democracy and other macro-economic variables in determining the likelihood of soft law implementation.

The concept of soft law, pertains to quasi-legal instruments which have no formal legally binding capability, or where the enforcement capability is assumed to be weaker relatively speaking than the force that underlines traditional law, sometimes in this context referred to as hard law. While the term soft law is largely synonymous with discussions with international law, it may also be referred to when commenting on domestic legal and doctrinal matters. Herein, our focus remains on international law. An apt example of what may be described as soft law would include resolutions and declarations of the United Nations general assembly components of bilateral and multi-lateral treaties, such as codes of practice and other non-treaty obligations. It is nonetheless problematic identifying a consistent definition of soft law given the fervent debate that exists between those denying that such a legal paradigm exists and those that deem it an additional sphere of international law. Shelton (2000), offers the following guiding definition that soft law is in essence the normative provisions contained in non-binding texts. Arguably, the works of Baxter (1980), and Weil (1983), are seminal within the sphere of research. Baxter (1980), contends that soft law is representative of the different intensity of agreement that exists within the expression of international law. Weil (1983), responds to this assertion and argues for caution when seeking to derive concepts of relative normativity in international law. Shelton (2000), and Abbott & Snidal (2000), offer a unifying set of theses that serve as the conceptual framework for the present study. Specifically, Shelton (2000), outlines compliance with soft law through the analysis of a wide variety of non-binding legal instruments within a variety of domains. Abbott & Snidal (2000), contend that soft law exists by virtue of the weakening of a legal arrangement with respect to the obligations delegation and precision of said arrangement. Having accepted the existence of soft law, it is pertinent to understand what factors motivate the enactment of international soft law.

Guzman & Meyer (2010), offer a worthwhile summation of why soft law is employed by states. The authors advance for complementary explanations for why states employ soft law that seek to describe a much broader range of state behaviour than has been previously explained.

First, and least significantly, states may use soft law to solve straightforward coordination games in which the existence of a focal point is enough to generate compliance.

Second, under what we term the loss avoidance theory, moving from soft law to hard law generates higher sanctions that both deter more violations and, because sanctions in the international system are negative sum, increase the net loss to the parties. States will choose soft law when the marginal costs in terms of the expected loss from violations exceed the marginal benefits in terms of deterred violations.

Third, under the delegation theory, states choose soft law when they are uncertain about whether the rules they adopt today will be desirable tomorrow and when it is advantageous to allow a particular state or group of states to adjust expectations in the event of changed circumstances. Moving from hard law to soft law makes it easier for such states to renounce existing rules or interpretations of rules and drive the evolution of soft law rules in a way that may be more efficient than formal renegotiation.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Fourth, we introduce the concept of international common law (ICL), which we define as a nonbinding gloss that international institutions, such as international tribunals, put on binding legal rules. The theory of ICL is based on the observation that, except occasionally with respect to the facts and parties to the dispute before it, the decisions of international tribunals are nonbinding interpretations of binding legal rules. States grant institutions the authority to make ICL as a way around the requirement that states must consent in order to be bound by legal rules. ICL affects all states subject to the underlying rule, regardless of whether they have consented to the creation of the ICL. As such, ICL provides cooperation minded states with the opportunity to deepen cooperation in exchange for surrendering some measure of control over legal rules.

Ho (2002), offers a worthwhile account of the economic and institutional determinants of the implementation of the first Basle accord employing traditional empirical method. The study of 107 countries explored the importance of factors such as banking sector concentration, democracy, macro-economic conditions and savings rates on the likelihood of implementation of the accord. As noted, the study identified both a positive and significant effect associated with the nature of the jurisdictions degree of democracy, as measured by the author using an ordinal sale variable.

The noted study employs logistic regression modelling techniques. This is consistent with much of the extant research when exploring relationships involving dichotomous dependent variables. The most commonly method employed in empirical legal research is logistic regression, largely because it is well suited to the dichotomous dependent variables that serve as the subject of many studies within the legal sciences. While the contributions to empirical legal research have been apt, many complex non-linear problems have not been considered empirically. The present article proffers an alternative method asserted to be parsimonious for research in the legal domain, representing a viable alternative to the commonly employed logit model.

An Introduction to the Multilayer Perceptron

While there remains the need for further acknowledgement of the benefits associated with parametric estimation methodology within the law and economics, the so called second wave of law and economics (Richardson, 1989), it is apparent that novel non-parametric methodologies made possible by modern computing technologies represent an equally important sphere of inquiry. In particular, the advantages of multi-layer perceptron models in comparison to logit models when modelling complex nonlinear processes are pertinent to consider a number of recent studies have sought to advance the role of artificial intelligence methods within legal research and legal practice. Hanke (2017), considers the role of artificial methods intelligence in response to dispute resolution. The author contends that artificial intelligence applications will augment the capabilities of lawyers and as a consequence should make them more productive and reduce the cost associated with legal advice provision. Nissan (2017), offers a precis on recent advancements within law and A.I. as it relates to the legal profession. The author explains how machine learning technologies and non-machine learning based applications are supplementing and augmenting legal service provision and legal practice. Artificial intelligence methods and the role of such technologies in advancing justice outcomes has been the focus of much recent research (see inter alia Bench-Capon & Modgil, 2017; Branting, 2017; Verheij, 2017; D'Rosario, 2017a, 2017b). This notable interest extends upon the substantial interest in the law and A.I. field of research, deriving from recent advancements in autonomous vehicle technologies and their associated law and artificial

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

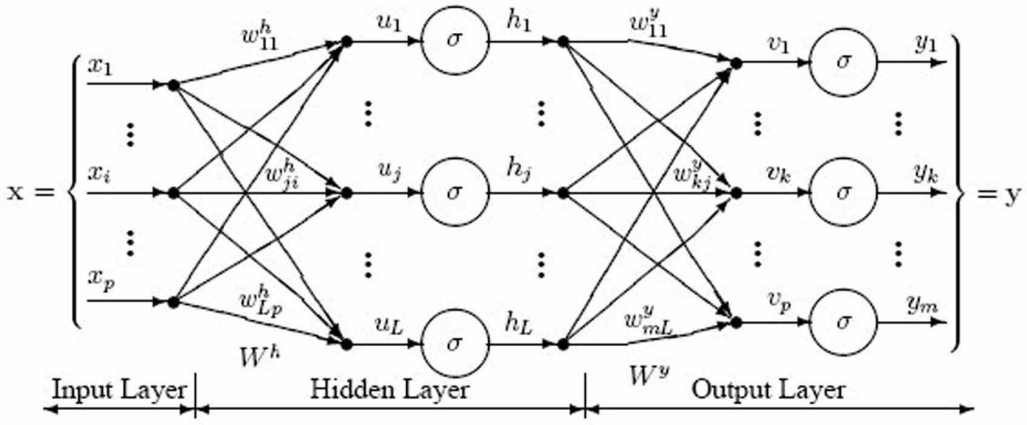
intelligence ethics questions. There remains however, only a small number of studies employing artificial intelligence methods when addressing empirical legal questions pertaining to economic soft law. The multilayer perceptron model is simply a form of artificial neural network model (herewith MP-ANN), a non-parametric predictive model that seeks to replicate the structure of a biological neuron as it occurs in nature. More plainly it is a series of weighted, aggregative, non-linear values. Neural network models emerged out of the Artificial Intelligence body of research that should to attempt to model the human process of learning through the development of a modelling framework resembling the structure of the human brain (Patterson, 1996). The MP-ANN model has the potential to provide more accurate predictive outcomes than traditional parametric estimation techniques, it is therefore well suited to the present research questions (D'Rosario, 2017). Rosenblatt (1958) posited the perceptron model, framing the concept of the retina layer, distributing input values. The work of Minsky & Papert (1969) asserted a number of the limitations of the perceptron model, the work brought about a diminution of interest in perceptron modelling. The key challenge evidenced in the perceptron model as posited by Rosenblatt was the step function that made problematic to train (D'Rosario, 2017). A revival of sorts occurred as a result of the work of Rumelhart, Hinton and Williams whose seminal paper "Learning Internal Representations by Error Propagation", offered resolution. Their work enabled the effective training or a multilayer neural network. The backpropagation training method and viable algorithm was first asserted in Rumelhart and McClelland (1986); it was the first practical method for training neural networks (D'Rosario, 2017).

Figure 1 illustrates conceptually the study's posited three-layer perceptron network.

This visual depiction is of a three-layer perceptron model, though more layers are possible, notwithstanding the significant increase in resources required for multiple hidden layers being employed. The three-layer perceptron is constituted by an input layer, hidden layer and an output layer. The first layer is the input layer. The input layer processes and standardises a vector of the predictor variables ($x_1 ... x_p$). These variables then take the range of -1 to 1. These standardised values are then passed through to the hidden layer. This process loosely resembles the process of "saltatory conduction" with a biological neuron.

The standardised vector values of the predictor variables, for example standardised vectors of amici curie and/or attitudinal variables, reach the hidden layer. At the hidden layer, the standardised values are multiplied by a weight (w_{ji}), with the weighted values aggregated (u_j) and passed forward to the transfer function. The values from the transfer function are passed to the output layer (h_j).

Figure 1. Three-layer perception network



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The output layer accepts the values from the hidden layer, multiplying the values by weights (w_{kj}), with the weighted values then aggregated resulting in a composite value v_j . The v_j value is then passed forward to the transfer function. The values from the transfer function are the model outputs y_k. Should the target variable be dichotomous, then there will be *2* neurons in the output layer producing *2* values, one for each of the categories of the target variable. Where the output value takes greater than to values (y_k = n), then the output layer will be constituted by n values.

Neural Models as an Alternative to Logistic Regression in Legal Science Research

The use of such methods is not uncommon in the social sciences but remains relatively underemployed within interdisciplinary legal research. Buchanan & Headrick (1970), offers an early speculative account in regard to the role of artificial intelligence methods within the legal profession asserting the role of artificial intelligence in modelling legal reasoning and argumentation. Rissland (1989), provided an account of the potential role of artificial intelligence methods in modelling legal process. The author provided practical insight into how reasoning could be captured and implemented within an artificial intelligence framework. This study extended upon the earlier work of McCarthy & Haze (1981), who addressed many of the issues associated with subsuming complex logic into an artificial intelligence framework. Bench-Capon (1997), provided a formal explication on the manner with which results-based frameworks and rule based legal information could be incorporated into artificial intelligence models of legal argumentation. The study outlined four areas where legal argument has been applied. Specifically, legal reasoning based on cases, explaining results based on rules of law and rule based legal information in resolution of normative conflict and problems on non-monotonicity and the basis of dialogue games to support the modelling of legal argument. Walton (2005), offer some refinement to these methods in the authors text, Walton (2005) outlines methods of argumentation and their application in artificial intelligence methods research relating to law incorporating practical and legal profession relevant guidance. However, it is noteworthy to acknowledge that artificial intelligence methods remain underutilised within legal research in comparison with other disciplines.

Importantly, there remains a dearth of research considering key empirical questions, such as the impact of legal experience of lawyers and the role of amicus curie on legal outcomes; and an even greater dearth of literature positing practically framed deterministic models of legal outcomes utilising the posited non-parametric models. Notable parametric studies within this area of research include those of Matthew Sag, Tonya Jacobi and Barton Beebe. There is little if any non-parametric research within this field. The use of neural networks is not however absent within the broader body of legal research. Warner (1990, 1992, 1993), was amongst the first to posit the benefits of logic based legal expert systems and one of the first legal realists to employ neural networks within legal expert systems. Warner (1990, 1992, 1993), contends that neural networks have the unique capacity to replicate deontic logic. However, the use of neural networks has not been limited to legal realists alone. Lothar Philipps (1991), emphasis the benefit of neural networks within the legal domain. Hobson & Slee (1994), emphasis the benefits of neural networks in their application to case analysis and case based reasoning. Rose & Belew (1989, 1991), claim that neural networks can retrieve appropriate data while employing a legal realistic perception of the law.

The first significant practical use of artificial neural networks in law was first developed by Stranieri et al (1999). They used a combination of rules and artificial neural networks and rules to determine what percentage of the common pool of assets a husband would receive following divorce in Australia. The

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

argument structure of Toulmin (1958) was used to justify conclusions made by their Split-Up system. In Stranieri and Zeleznikow (2005) the authors provided a detailed discussion on how knowledge discovery from legal databases could be gainfully used to assist legal decision making.

While some of these assertions may possibly overstate the benefits of neural networks within the legal domain to an extent, it is undeniable that neural networks possess greater pattern identification and are more consistent with logic and rule base systems of case classification than logistic regression techniques. The current study responds to both the noted dearth of deterministic research and the dearth of empirical work considering these matters.

The Multi-layer perceptron model is asserted to be a viable alternative to the logit model because it offers a more accurate framework for prediction. Consider some of the challenges associated with alternative parametric techniques. Firstly, there are obvious issues associated with predictive analysis where the process is non-linear. Logit and Probit models are relatively less capable of modelling non-linear decision processes. MP-ANN overcomes this challenge through the application of weighted, aggregative non-linear values based estimation. Neural networks have been advocated as an alternative modelling technique to logistic regression within the applied sciences, however there is a dearth of research considering the viability of such methods within law and economics research.

Within law and economics research, logistic regression is by far the most commonly employed method, given its suitability to dichotomous dependent variables. Within the applied sciences many more complex alternatives are being adopted. These methods present as viable alternatives to logit and probit modelling.

As asserted by Tu (1996) neural networks offer a number of advantages over logistic regression, including the ability to implicitly detect complex nonlinear relationships between dependent and independent variables, ability to detect all possible interactions between predictor variables, and the availability of multiple training algorithms. Driesietl and Ohno-Machado (2002) offer a formal methodological review of each modelling technique, noting the formal association between the techniques. The multi-layer perceptron model can be seen as an abstraction of the logistic model, employing a sigmoid function. A more apt description however would be that the logistic model is a simple single layer perceptron model, that does not employ a second (layer) weighted aggregate input set. [1] Indeed, there are a number of benefits associated with neural networks that render the approaches advantageous in comparison to traditional logistic estimation. Firstly, the methods appear to evidence greater consistency with legal decision making. If it is assumed that all decision making is linear or sigmoidal, and therein all decisions can be estimated with a linear or stigmoidal function then neural networks and logistic regression offer similar estimates.

Neural models appear more consistent with legal decision making (see Kannai et al 2007, D'Rosario, 2017). Members of the legal system carefully weigh alternative criteria, balancing competing criteria and implicitly employ these weighted considerations in their decision making. While the process may not be explicit it is nonetheless essential to legal decision making.

Moreover, the perceptron may be more accommodating of non-linear decision processes. The perceptron model and back propagation method were developed to respond to complex non-linear processes and problems (n.b. Rumelhart, Hinton and Williams). This is because of the existence of the hidden layer that accepts input values from vectors of the predictors. This allows for the combination of vector inputs, rather than the direct pass through of the predictor weights directly to the output/dependent variable. The logistic regression method involves the direct pass through of the predictor weights. This is largely because the logistic regression method can be seen as a simple form of single layer feed forward neural network. It is this assertion that may result in greater consideration being afforded to non-parametric models of estimation.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Neural models have been shown to provide more accurate prediction rates in legal studies such as those of Stranieri et al. (1999), Kannai et al (2007), D'Rosario (2017) and D'Rosario (2017). As such where outcome prediction is assigned the highest priority it is a superior estimation methodology. Where there is complex interaction between predictor variables the neural model is likely to provide a better rate of prediction. Many legal decision processes involve complex interactions between predictor variables. As such these legal decision processes are not well suited to conventional regression based estimation, and in particular linear regression. Perceptron models appear to be well suited to modelling legal decision process because of the formal consistency of the modelling technique with the process of legal decision making.

The method is not without its challenges. The most obvious challenge associated with the use of neural network models are the computing resources required for the employment of such modelling techniques. Multilayer perceptron models are more resource intensive than traditional parametric techniques, though this is of little consequence given the ubiquity of modern computing technology.

The most significant challenge of neural network-based modelling and machine learning generally is the perception of such methods as black boxes that are difficult to understand. This is a misnomer as there are now many techniques available to understand the relative importance of model variables and model parameter impacts. Given the techniques developed by Garson (1991) and Goh (1995), as well as more contemporary methods, this is certainly no longer an accurate representation of the methodology. Normalised importance analysis allows for the interpretation of neural network connection weights. Thus, providing a set of compelling metrics for the interpretation and analysis of predictor variables, and determinations about the relative importance of predictor variables.

The Logit model is the obvious choice when a dependent variable is non- continuous. It is also useful when addressing non- liner associations. However, the technique fails to properly measure threshold events and specification effects. This results in model specifications incapable of analysing and measuring complex interactions between regressors. A number of non-parametric techniques are more capable of accommodating complex relationships with threshold and interaction effects.

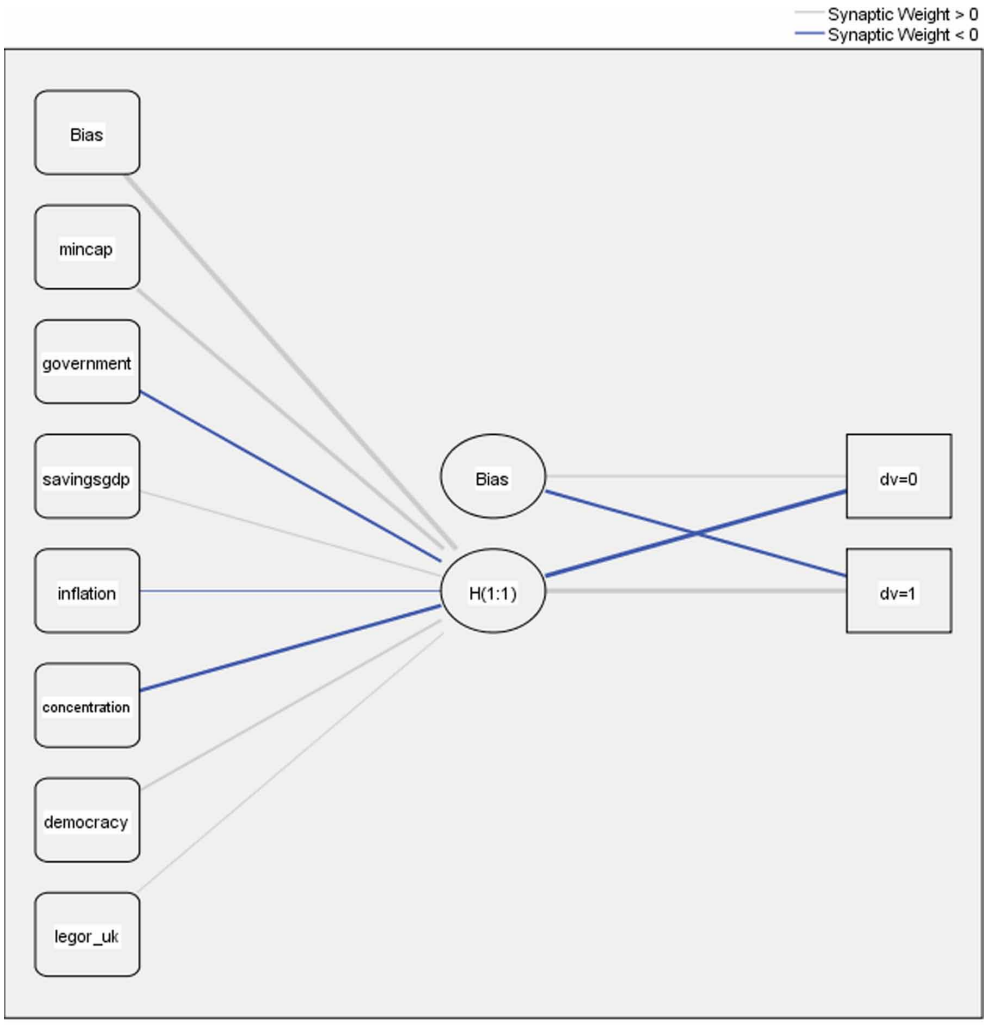
MODEL STRUCTURE AND PARAMETERS

The model seeks to determine the likelihood of soft law (Basle Accord) implementation given a series of macroeconomic, legal and demographic variables. Specifically, the specification models the implementation of the Basel accord based on a number of viable estimators. Figure 2 conceptually illustrates the study's posited three-layer perceptron network. The model is a three-layer network where Basle Accord implementation, the dependant variable is a binary variable taking the values of implement or no implementation. (This model denotes the layer and node structure of the model. It does not define weights and connections, which are purely illustrative):

The ANN model employed — the MP-ANN model — is a three-layer model with an input layer deriving from the variables provided by Ho (2002). Accord implementation is the variable denoting the implementation of the first Basle accord denoted as a binary dummy variable. Government is a scale variable denoting the effectiveness of government. SavingsGDP is a continuous variable denoting the level of savings relative gross domestic product measured in U.S. dollars (U.S. constant dollars base year 2000). Inflation denotes the rate of inflation within each jurisdiction in average terms during the preceding year. Concentration is a measure of the level of banking sector concentration. Democracy is

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Conceptual model diagram



a proxy variable quantifying strength of democratic practices within the jurisdiction. Corruption (denoted in the model as corrupti80-89, as per Ho (2002)) is a measure of corruption deriving from World Bank corruption statistics. RL is a measure of the rule of law within each jurisdiction. Lngdpcap is the log value of GDP per capita. Signatory id a variable denoting whether the nation is a signatory to the Basle accord. Oppositionul is a variable denoting division within government. Debtwldperc is a variable denoting the current level of debt. Lndebt is the log value of the current level of debt. Lninswld is a measure of international insurance and financial exposure. Supervise is a measure of banking sector supervision (see Table 1).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Variable List

Variable	Obs	Mean	Std. Dev.	Min	Max
concentration	95	68.04189	22.77217	12	100
democracy	153	4.241486	3.840977	0	10
corrupti80_89	140	0.085714	0.487037	0	4
Rl	155	-0.0008	0.906071	-1.56681	2.129017
lngdpcap	177	7.572594	1.551932	4.61494	10.70118
signatory	217	0.059908	0.237865	0	1
oppositionul	153	0.281261	0.458103	0	1.571429
debtwldperc	140	0.714286	1.696409	0	11.58676
lndebt	135	23.58345	2.310281	18.00163	28.44928
lninswld	155	19.05335	2.258084	13.72721	24.46665
supervise	107	10.25234	2.588412	4	14
transfers	175	-0.28275	4.030983	-32.3727	6.9032
reglnorm	87	0.574702	0.060534	0.4772	0.6667
capacctsm	206	0.083658	0.592008	-3.17104	5.841527
insfinwldsm	160	1.827934	5.58895	0	42.2405
totradesm	161	5.35267	18.14126	0.002116	124.5998

Note: This table summarises the data employed within the study. All data was sourced from Ho, (2001).

THE LAYERS OF A NEURAL NETWORK

The first layer is the input layer. The input layer processes and standardises a vector of the predictor variables, herein the macroeconomic, governance and stability variables. These variables then take the range of -1 to 1. These standardised values are then passed through to the hidden layer. This process loosely resembles the process of "saltatory conduction" with a biological neuron.

The standardised vector values of the predictor variables, government, savingsgdp, inflation, concentration, democracy, corruption, rl, lngdpcap, signatory, oppositionul, debtwldperc, lndebt, lninswld, Supervise, reach the hidden layer. At the hidden layer, their standardised values are multiplied by a weight, with the weighted values aggregated and passed forward to the transfer function. The values from the transfer function are passed to the output layer.

The output layer accepts the values from the hidden layer, multiplying the values by weights, with the weighted values aggregated and passed forward to the transfer function. The values from the transfer function are the model outputs. As the target variable is the outcome for the petitioner, there are *2* neurons in the output layer producing *2* values, one for each of the categories of the target variable.

The hidden layer size was determined based on Heaton's work, noting optimality generally exists between the number of input layers and output layers (Heaton, 2011). Moreover, two hidden layers may be employed when modelling data that evidences discontinuities, such as a saw-tooth wave pattern. Given the nature of the underlying dataset, this is not the case. The use of an additional hidden layer reduces estimation efficiency and offers no improvement to the model. Herein, a single hidden layer model is employed to avoid a greater risk of converging to an incorrect local minima. One potential shortcoming

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

of the selected methodology relates to the efficiency of ANN models in accommodating non-numeric variables, commonly referred to as dummy variables in economics. This is, however, only the case where non-numeric variables can take a large number of values, which is not the case in the present study. The total number of units employed in the model was 16, and determined through optimisation, again indicating interaction effects.

Data pruning techniques were not employed to improve model resolution, and the model performed efficiently and robustly absent of any node reduction. As such, the input values remain the same as those employed in Ho (2002). This is pertinent as it ensures that each predictive modelling framework can be compared on equal terms employing the same parameter set (see Table 2).

Table 2. Network Information

Input Layer	Factors	1	government
		2	savingsgdp
		3	inflation
		4	concentration
		5	democracy
		6	corrupti80_89
		7	rl
		8	lngdpcap
		9	signatory
		10	oppositionul
		11	debtwldperc
		12	lndebt
		13	lninswld
		14	supervise
	Number of Units[a]		563
Hidden Layer(s)	Number of Hidden Layers		1
	Number of Units in Hidden Layer 1[a]		16
	Activation Function		Hyperbolic tangent
Output Layer	Dependent Variables	1	Dv
	Number of Units		2
	Activation Function		Softmax
	Error Function		Cross-entropy

a. Excluding the bias unit

Note: This table summarises the structure of the model presented within this study. The model is a multi-layer perceptron employing hyperbolic tangent and soft max activation functions. The dependant variable is the implementation of the Basel Accord.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

MODEL FINDINGS

The MP-ANN is able to robustly predict the decision to implement the Basel Accord 100 per cent of the time (with no converging minima) with accuracy in prediction and goodness of fit far greater than the traditional logistic regression. This suggests that the underlying variables are useful in the determination of the outcome of a soft law implementation process. The model is capable of robustly predicting outcomes more consistently than any of the models in the extant literature. The size of the relative weights employed in the model is detailed in Table 3.

The goodness-of-fit measures suggest that the MP-ANN is superior to the logistic model in the Ho study (Ho, 2002). While the Ho model remains conceptually and empirically sound, it does not achieve the same level of predictive accuracy. The findings suggest that this model is more predictively capable, with an improvement in predictive accuracy of a minimum of14 per cent. This is amongst the most pertinent contributions of the present study. The study conveys the benefits associated with non-parametric methods of prediction in economic soft law research and the benefit of MP-ANN modelling in the development of predictive models in the economics law.

Employing the methods of Garson (1991) and Goh (1995), the normalised importance analysis suggests that the democracy (the proxy of the strength of democratic practices) ranks as the most important determinant of litigant success, followed by the government effectiveness measure and the strength of the rule of law measure. The data is presented graphically in Figure 3. The variation in the hidden layer structure yields comparable results and robustly achieves a 100 percent predictive accuracy. Thought that it is notable that the relative ranked importance of government effectiveness and democracy was reversed in some model specifications, lending credence to the assertion that government effectiveness may be of equal importance to the strength of democratic practices when determining the likelihood of soft law implementation. These findings are broadly consistent with the findings of the Ho study, with a few notable points of difference (Ho, 2002).

Firstly, unlike the Ho study, the current study indicates that government effectiveness and the rule of law may rank as equally important to the strength of democracy in determining the likelihood of soft law implementation. Additionally, the network results suggest that threshold and interaction effects may exist between key independent variables democracy rule of law government effectiveness and savings GDP. Again, noting that substantial improvement in pseudo R-Squared values ranging between 0.65 – 0.81 in contrast of the present study of 0.99 to 1.0.

CONCLUSION

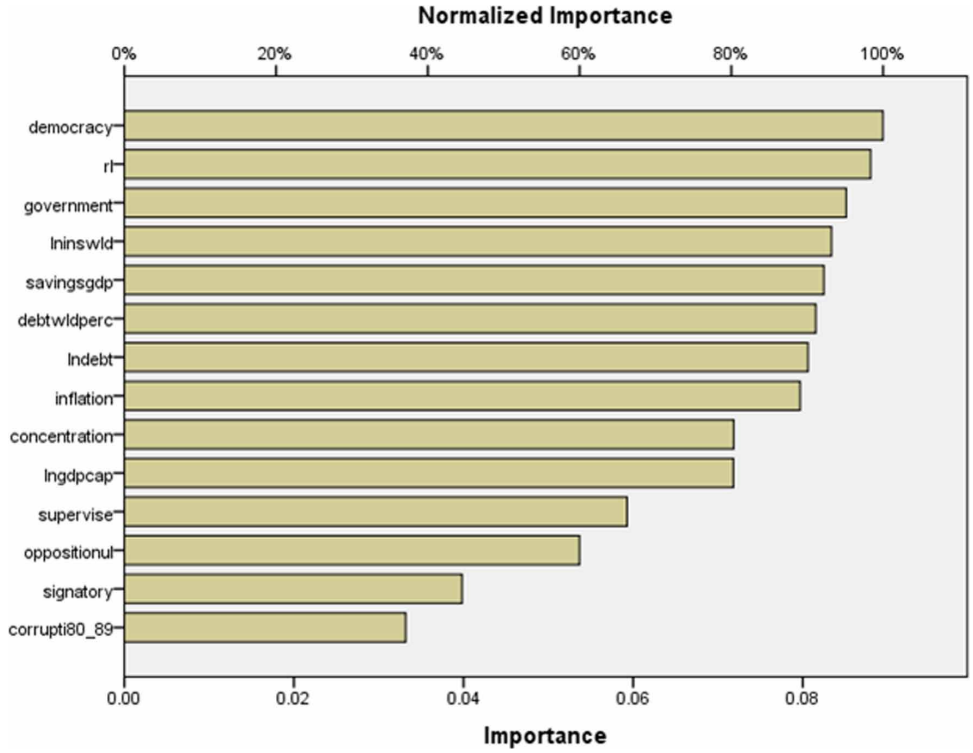
The study has evidenced the benefits associated with non-parametric predictive modelling and specifically the benefits of artificial neural networks in legal research. Unlike a number of recent studies employing time series and panel data in non-parametric modelling in legal research, specifically employing artificial neural networks, the present study makes evident the efficacy and efficiency of neural networks in legal research employing cross sectional data. The findings indicate the employed predictive methods may be of greater utility than logistic models when emphasis is placed on outcome prediction rather than casual association, which is particularly important when seeking to determine the likelihood of economic soft-law adoption. The study indicates that while logit and probit models are worthwhile exploratory tools when engaging in empirical work in law and economics and can accommodate non-linear problems;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Normalized Variable Importance

Independent Variable Importance		
	Importance	**Normalized Importance**
Government	.085	95.1%
Savingsgdp	.082	92.2%
Inflation	.080	89.1%
concentration	.072	80.3%
Democracy	.089	100.0%
corrupti80_89	.033	37.1%
Rl	.088	98.4%
Lngdpcap	.072	80.3%
Signatory	.040	44.5%
oppositionul	.054	60.0%
Debtwldperc	.082	91.1%
Lndebt	.081	90.1%
Lninswld	.083	93.2%
Supervise	.059	66.3%

Figure 3. Independent variable importance analysis



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

they are less capable of accommodating non-linear problems with complex interactions as efficiently as multilayer perceptron models, at least when the desire is to engage in a predictive exercise. Logit and probit models cannot capture threshold and interaction effects unless such effects are known and understood and can be specified in model prior to estimation. Additional research is nonetheless necessary to determine the potential generalizability of this claim but the claim appears to be highly plausible in light of recent studies. The notion that perceptron modelling may be beneficial and potentially superior in soft law and social sciences research represents the first contribution of the study. A further contribution of the study is its response to the dearth of studies employing empirical methods to legal research pertaining to economic soft-law, and in particular methods beyond logistic regression.

Importantly the current research displays the potential benefits associated with legal system and legal structure data collection efforts when employed with modern parametric estimation techniques. The MP-ANN models may be useful tools in determining the relative importance of independent variables and the model of estimation to employ when engaging in casual analysis. The modelling methodology employed herein evidences the benefit of the presented empirical technique in predicting soft law outcomes. Critically, the role of MP-ANN models in empirical legal research is advanced based on the notion that such models require heuristic knowledge pertaining to data preparation and cataloguing, and sufficient technical knowledge to employ the appropriate ANN for the estimation. It is hoped that the present study shall inform future legal researchers engaging in empirical legal research employing non-parametric estimation technique such as those employed herein. The ANN models are more conceptually simple than linear modelling techniques (albeit more mathematically complex); not to mention more appropriate for modelling complex decision models, thus supportive of greater potential adoption by the legal fraternity. It would be apt for future research to consider the potential application of probabilistic neural networks to such empirical legal analysis of soft law variables.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and the editor for their insightful comments and suggestions.

REFERENCES

Abbott, K. W., & Snidal, D. (2000). Hard and soft law in international governance. *International Organization*, *54*(3), 421–456. doi:10.1162/002081800551280

Ashley, K. D. (2017). *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*. Cambridge University Press. doi:10.1017/9781316761380

Baker, L. A. (1996). Interdisciplinary Due Diligence: The Case for Common Sense in the Search for the Swing Justice. *Southern California Law Review*, *70*, 187.

Baxter, R. R. (1980). International Law in “Her Infinite Variety.” *The International and Comparative Law Quarterly*, *29*(4), 549–566. doi:10.1093/iclqaj/29.4.549

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bench-Capon, T., & Modgil, S. (2017). Norms and value based reasoning: Justifying compliance and violation. *Artificial Intelligence and Law*, *25*(1), 29–64. doi:10.100710506-017-9194-9

Bex, F., Prakken, H., van Engers, T., & Verheij, B. (2017). Introduction to the special issue on Artificial Intelligence for Justice (AI4J). *Artificial Intelligence and Law*, *25*(1), 1–3. doi:10.100710506-017-9198-5

Bex, F., Prakken, H., van Engers, T., & Verheij, B. (2017). Introduction to the special issue on Artificial Intelligence for Justice (AI4J). *Artificial Intelligence and Law*, *25*(1), 1–3. doi:10.100710506-017-9198-5

Branting, L. K. (2017). Data-centric and logic-based models for automated legal problem solving. *Artificial Intelligence and Law*, *25*(1), 5–27. doi:10.100710506-017-9193-x

D'Rosario, M. (2017). The Impact of Legal Advocacy Experience Within the US Supreme Court on Trial Decision Outcomes. *International Journal of Strategic Decision Sciences*, *8*(1), 65–76. doi:10.4018/IJSDS.2017010105

D'Rosario, M. (2017). Fair Use Defences During Copyright Litigation: Is the Success of a Fair Use Defence Strategy Predictable? *International Journal of Strategic Decision Sciences*, *8*(2), 31–51. doi:10.4018/IJSDS.2017040103

Enns, K. M. (1998). Can a California Litigant Prevail in an Action for Legal Malpractice Based on an Attorney's Oral Argument before the United States Supreme Court? *Duke Law Journal*, *48*(1), 111–146.

Epstein, L., Ho, D. E., King, G., & Segal, J. A. (2005). The Supreme Court during crisis: How war affects only non-war cases. *NYUL review, 80*, 1.

Epstein, L., & Knight, J. (1997). The choices justices make. *Sage (Atlanta, Ga.)*.

Garson, G. D. (1991). Interpreting neural-network connection weights. *AI Expert*, *6*(4), 46–51.

Goh, A. T. C. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, *9*(3), 143–151.

Gressman, E., & Stern, R. (2002). *Supreme Court Practice: For Practice in the Supreme Court of the United States* (8th ed.). Bureau of National Affairs.

Guzman, A. T., & Meyer, T. L. (2010). International soft law. *The Journal of Legal Analysis*, *2*(1), 171–225. doi:10.1093/jla/2.1.171

Hanke, P. C. (2017). Computers with Law Degrees? The Role of Artificial Intelligence in Transnational Dispute Resolution, and its Implications for the Legal Profession. *Transnational Dispute Management, 14*(2).

Ho, D. E. (2002). Compliance and International Soft Law: Why Do Countries Implement the Basle Accord? *Journal of International Economic Law*, *5*(3), 647–688. doi:10.1093/jiel/5.3.647

Hunter, D. (1994). Looking for law in all the wrong places: Legal theory and legal neural networks. In Legal knowledge based systems: The relation with legal theory (pp. 55-64).

Hunter, D. (1999). Out of their minds: Legal theory in neural networks. *Artificial Intelligence and Law*, *7*(2-3), 129–151.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kannai, R., Schild, U., & Zeleznikow, J. (2007). Modeling the evolution of legal discretion – an Artificial Intelligence Approach. *Ratio Juris*, *20*(4), 530–558. doi:10.1111/j.1467-9337.2007.00374.x

Karnin, E. D. (1990). A simple procedure for pruning back-propagation trained neural networks. *IEEE Transactions on Neural Networks*, *1*(2), 239–242.

Makridakis, S. (2017). The forthcoming artificial intelligence (AI) revolution: Its impact on society and firms. *Futures*, *90*, 46–60. doi:10.1016/j.futures.2017.03.006

Martin, A. D., & Quinn, K. M. (2002). Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999. *Political Analysis*, *10*(2), 134–153.

Martin, A. D., Quinn, K. M., & Epstein, L. (2004). The median justice on the united states supreme court. *NCL review, 83*, 1275.

Maskus, K. (2002). Intellectual Property Issues for the United States and Japan: Disputes and Common Interests. In Issues and Options for US–Japan Trade Policies (pp. 165-190).

McAtee, A., & McGuire, K. T. (2007). Lawyers, justices, and issue salience: When and how do legal arguments affect the US Supreme Court? *Law & Society Review*, *41*(2), 259–278.

McGuire, K. T. (1995). Repeat players in the Supreme Court: The role of experienced lawyers in litigation success. *The Journal of Politics*, *57*(1), 187–196.

Minsky, M., & Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry*. Cambridge, Mass.: MIT Press.

Nissan, E. (2017). Digital technologies and artificial intelligence's present and foreseeable impact on lawyering, judging, policing and law enforcement. *AI & Society*, *32*(3), 441–464. doi:10.100700146-015-0596-5

Olden, J. D., & Jackson, D. A. (2002). Illuminating the "black box": A randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling*, *154*(1-2), 135–150.

Patterson, D. W. (1998). *Artificial neural networks: theory and applications*. Prentice Hall PTR.

Prakken, H. (2017). On the problem of making autonomous vehicles conform to traffic law. *Artificial Intelligence and Law*, *25*(3), 341–363. doi:10.100710506-017-9210-0

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Learning Internal Representations by Error Propagation. In *Readings in Cognitive Science* (pp. 399–421). Science Direct.

Sankar, A., & Mammone, R. J. (1991, July). Optimal pruning of neural tree networks for improved generalization. In *Seattle International Joint Conference on Neural Networks IJCNN-91* (Vol. 2, pp. 219-224). IEEE.

Sheehan, R. S., Mishler, W., & Songer, D. R. (1992). Ideology, status, and the differential success of direct parties before the Supreme Court. *The American Political Science Review*, *86*(2), 464–471.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Shelton, D. (Ed.). (2000). *Commitment and Compliance: The Role of Non-binding Norms in the International Legal System*. Oxford: Oxford University Press.

Songer, D. R., & Sheehan, R. S. (1992). Who wins on appeal? Upperdogs and underdogs in the United States Courts of Appeals. *American Journal of Political Science*, 235–258.

Stranieri, A., & Zeleznikow, J. (2005). Knowledge Discovery from Legal Databases (Vol. 69). Springer Science & Business Media.

Stranieri, A., Zeleznikow, J., Gawler, M., & Lewis, B. (1999). A hybrid—neural approach to the automation of legal reasoning in the discretionary domain of family law in Australia. *Artificial Intelligence and Law*, *7*(2-3), 153–183. doi:10.1023/A:1008325826599

Toulmin, S. (1958). *The Uses of Argument*. Cambridge: Cambridge University Press.

Verheij, B. (2017). Proof with and without probabilities. *Artificial Intelligence and Law*, *25*(1), 127–154. doi:10.100710506-017-9199-4

Weigend, A. S., Rumelhart, D. E., & Huberman, B. A. (1991). Back-propagation, weight-elimination and time series prediction. In Connectionist models (pp. 105-116).

Weil, P. (1983). Towards relative normativity in international law? *The American Journal of International Law*, *77*(3), 413–442. doi:10.2307/2201073

ENDNOTE

[1] For a full methodological review see Dreiseitl, S. & Ohno-Machado, L. (2002). Logistic regression and artificial neural network classification models: a methodology review. *Journal of Biomedical Informatics*, *35*(5–6), 352-359. ISSN 1532-0464

This research was previously published in the International Journal of Strategic Decision Sciences (IJSDS), 9(3); edited by Madjid Tavana; pages 1-15; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 5
Predicting Credit Rating Migration Employing Neural Network Models

Michael D'Rosario
Deakin University, Australia

Calvin Hsieh
Deakin University, Australia

ABSTRACT

Credit rating migration ranks amongst the most pertinent issues concerning institutional lenders and investors alike. There are a number of studies that have employed both parametric and non-parametric methodologies to forecast credit rating migration, employing machine learning methods; and notably, artificial intelligence methods becoming increasingly popular. The present study extends upon research within the extant literature employing a novel estimation method, a neural network modelling technique, herewith the MPANN (multi-layer neural network). Consistent with the extant literature, the present article identifies that the legal framework and system of taxation enacted within a polity are pertinent to predicting rating migration. However, extending upon traditional estimation techniques the study identifies that a number of different model calibrations achieve greater predictive accuracy than traditional parametric regression. Notably, the method is able to achieve superior goodness of fit and predictive accuracy in determining credit rating migration than models employed within the extant literature.

INTRODUCTION

Credit rating migration is one of the most significant concerns for institutional lenders and investors alike. Credit rating migration is a risk that individuals or financial institutions are both exposed to. In essence, rating migration is the worsening or upgrading in credit rankings, assigned by an independent ratings agency. There are a number of studies that have employed both parametric and non-parametric methodologies to forecast financial outcomes, employing machine learning methods; and notably,

DOI: 10.4018/978-1-7998-0951-7.ch005

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

artificial intelligence methods becoming increasingly popular, (see inter alia, Martinez et al., 2016). There are only a handful of studies on employing artificial intelligence methodologies to determine their ability to predict changes in credit status. Bhattacharya (2006), found that artificial neural network models had outperformed linear discriminant models; the study gave rise to further studies employing neural network models (see inter alia D'Rosario, 2017). Given the advancements achieved in artificial intelligence methods, such methods can now be employed more efficiently and robustly in application to financial research. The present study seeks to predict credit migration of sovereign ratings in Latin America, employing a multi-layer perceptron, artificial neural network model.

The present study is timely in light of changes to the Basel Accord. The Basel accord has generated interest in credit ratings methodology and prediction since banks and financial institutions employ the prescribed risk weights to determine the probability of default and estimate capital adequacy requirements (Behr and Guttler, 2008). There is an increasing number of users interested to predicting credit migration and the probability of credit rating changes of the borrowers (Jarrow et al., 1997). A few studies have tried employing credit transition matrices (Jarrow et al., 1997). However due to variations in the credit quality of borrowers, various credit rating agencies quote variations in the rating of the same borrower, mainly due to the timing of the changes of credit rating, the current rating grade and various other factors (Hill, Brooks & Faff, 2010). It is noteworthy that the afore stated studies employ traditional parametric methods, rather than the contemporary artificial intelligence methods employed herein.

The critical aspect of credit rating of sovereign bonds is that they act as a signal in the market, which helps government issuers and corporations issuing securities alike to attract both domestic and foreign investment. As noted, there have been a number of studies predicting credit rating migration using credit transition matrices, while others employing financial time series data and associated time series estimation techniques, such as GARCH and related statistical tools in credit rating migration analysis. There is a genuine dearth of studies employing artificial neural network models to predict credit migration of sovereign bonds. Moreover, the majority of credit migration studies focus on individual credit ratings prediction and ratings migration. The present study addresses the dearth of research employing artificial intelligence methods and specifically multi-layer perceptron artificial neural network models to predict the migration of sovereign ratings.

LITERATURE REVIEW

Over the past few years Latin America has seen a change in how foreign capital is acquired. In the 1960's the alliance program provided assistance, the 1970's was when the commercial banks had an influx of funding due to the petroleum boom, which led foreign countries to lend aggressively to developing countries. In recent times these nations have become reliant on the IMF and other international agencies and other private sources (Sinclair, 2003). Apart from private funding, sovereign bonds and debt securities issued by government agencies are increasing in popularity in Latin America. Due to the rise in popularity, investors usually look at the credit ratings of major credit rating agencies, such as Moody's Investor Services, Standard & Poor's Rating service to assess the likelihood of default. As credit rating agencies have grown in importance and become a key intermediary between investors and borrowers and are nothing less than guardians of capital assigned to emerging markets (Sinclair, 2005).

In recent times due to the increasing complexity of financial markets, it becomes increasingly difficult to assess or predict the credit rating of governments/ sovereign bonds. The more effectively credit migra-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

tion is estimated, the more efficiently capital can be allocated. The complexity of securities in facilitating credit risk transfers, securities such as credit derivatives and collateralized debt obligations (CDO), has led to increased complexity in how credit risk is shared with market participants (Ferguson, 2001). As such, recent sovereign defaults on foreign denominated debt have increased, the 1994 Mexican Peso Crisis, the 1997 Asian financial crisis, the 2001 credit failure in Argentina and the recent European debt crisis (Greece) have raised concerns within financial institution about their risk exposures in emerging markets. Financial institution & regulators use VaR models or the expected shortfall to price risky loans (J.P Morgan's Credit Metrics, 2018). Additionally, with recent acceptance of the Basel Accords, capital adequacy requirements have been set in place for financial institutions by regulators.

The amount (in terms of currency) of debt written off in a sovereign crisis has exceed any corporate default. In this respect, it is necessary that credit migration warrants particular attention when assessing risk. Sovereign credit migration is difficult, mainly due to lack of data, with samples being small in both time series and cross-sectional terms. To estimate credit migration with high accuracy becomes challenging when it comes to emerging markets since it requires sovereign bonds to transition from one category to another. It becomes difficult since the data on emerging markets goes back only a few years. One plausible solution to this issue is to look at corporate bonds within a given country (Jackson & Perraudin, 2000), but this approach inherently assumes that sovereign bonds credit migration process follows corporate bonds, while empirical evidence suggests otherwise. Fundamentally, sovereign bonds have much smaller spreads when compared with corporate bonds. Additionally, sovereign bonds have evidenced greater variation in ratings among credit rating agencies as compared to corporate bonds.

The use of neural network models to predict credit migration has only been a source of focus and enquiry in recent times. Linear discriminant models (herewith LDA) and the Artificial neural network models (ANN) were used to predict credit migration, where the ANN employed back propagation with 3 layers in Bhattacharya et al, (2006). In Bhattacharya et al, (2006) where the ANN was set against the linear model, the ANN outperformed the linear model comprehensively (see inter alia. Bhattacharya et al., (2006), Zhang et al. (2004)). In the context of credit migration, when governments default on their obligations, this has serious ramifications within the financial market. Beyond LDA, there are several studies that employ neural networks to classify corporations based on nine popular financial ratios (see inter alia, Odom & Sharda 1990, Alam et al. 2005, Kim & Kang 2010), and have evidenced prediction accuracies of up to 73%. Other studies in financial services research beyond credit migration also employ ANN models in financial data mining (n.b. Zhiang et al (2004) and Chen et al, 2003). The noted studies which employed Probabilistic Neural Network, identified Neural networks achieved higher returns when compared to traditional parametric investment selection methodologies.

It is notable that, the usual methodologies used to predict credit rating migrations use a discrete homogeneous Markov Process. This approach disregards, the timing of the ratings changes and the duration of each rating. Typically, discrete multinomial estimators were employed by the credit rating agencies (Carty & Forns, 1993) and produces zero probability for unobserved transitions (e.g. Aa to a default). Basel II supports having a 0.03% for such rare events (events of Aa to default). Additionally, Basel has a positive weight on sovereign bonds which moves from Aa to Baa, but this implies that the migration risk is constant over time which clearly simplifies their model (Carty & Forns, 1993). The duration of credit ratings affects the probability of the transition (credit migration). Lando & Skodeberg (2002) and Kavvathas (2001) both confirm earlier findings of Carty & Fons (1993) which states that the duration of the credit rating effects the Corporate ratings but evidence on the sign of the ratings are mixed.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

There are two notable studies which look at sovereign bonds credit migration risk, namely Wei (2003) proposes Markov Model which accommodates time heterogeneity, but the sample size of the study was relatively small. Hu et al. employ a substantially larger data set and employ a probit analysis offering a more robust set of findings. Both studies achieved modest predictive accuracies, but are pertinent and serve to inform the present study. In this regard, much of the extant of the studies use linear statistical model to estimate sovereign credit migration. Other proximate studies that use Artificial neural nets study corporate bankruptcy, personal insolvency among other areas of study, while omitting credit rating migration. Our research seeks to employ artificial neural networks to estimate sovereign credit rating migration risk with the intention of providing robust and accurate predictive models.

An Introduction to the Multilayer Perceptron

While traditional statistical methods have some restrictive assumptions such as linearity, normality and independence among variables, there remains the need for further acknowledgement of the benefits associated with parametric estimation methodology within the law and economics, the so called second wave of law and economics (Richardson, 1989). Considering these assumptions of the statistical methods, financial data constantly violate these assumptions. Hence the methods can have limitations in validity and accuracy (Elliot & Kennedy, 1998). It is apparent that novel non- parametric methodologies are made possible by modern computing technologies which represent an important sphere of inquiry. In particular, the advantages of multi-layer perceptron models & artificial neural networks, in comparison to logit models when modelling complex nonlinear processes, a number of recent pertinent studies have sought to advance the role of artificial intelligence methods within finance & financial research studies. There remains however, only a small number of studies employing artificial intelligence methods when addressing credit migration of sovereign bonds. With the advancement of computing power and artificial intelligence, soft computing methods specifically when applied to predicting credit migration, include backward propagation and probabilistic model's neural networks, these models unlike traditional statistical methods offer qualitative tools which statistical methods can't estimate (Bhattacharya et al., 2006). The biological structure of a neuron is complex to model, but a non-parametric predictive model, i.e., the multilayer perceptron model is a form of artificial neural network model (herewith MP-ANN), that seeks to replicate it, essentially it is a series of weighted, aggregative, non-linear values.

Artificial Intelligence being the main body of research led to the emergence of neural network models. Given that neural network models is a form of AI, it therefore attempts to model the process of learning which in part resemble the human brain (Patterson 1996). Since the MP-ANN model (Figure 1) has the potential to estimate more accurate outcomes than traditional statistical models, it is well suited to shed light onto the present research study (D'Rosario, 2016). The work of Minsky & Papert (1969) asserted a number of the limitations of the perceptron model which led to a fall in interest in perceptron modelling. Although the perceptron model has its advantages over the traditional parametric estimation models, the key area of challenge in the perceptron model as posited by Rosenblatt was the step function that made it problematic to train the neural network (D'Rosario 2016). Rumelhart, Hinton and Williams research work on "Learning Internal Representations by Error Propagation", offered resolution to the problem of training the neural network. Their work enabled the effective training or a multilayer neural network through backpropagation training method and viable algorithm which was first asserted in Rumelhart and McClelland (1986); it was the first practical method for training neural networks (D'Rosario, 2016).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The ANN gets more efficient with each training iteration by adjusting positive weights when it predicts the right outcome. Although the work of Rumelhart, Hinton & Williams citation needed had posited a practical solution for training neural networks, which makes the ANN useful when a model is unknown since it can learn from the data set.

ANN are good at grouping data for classification, which has made such methods more efficient than other traditional parametric models in predicting outcomes. This discover has made neural networks become of significant interest in the financial industry, since vast amounts of historical financial data are available to train ANN models. Increasingly, the use of such methods has not only made it easier to accurately predict financial outcomes, but it also outperformed traditional statistical approaches (n.b. Schoneburg (1990) & Francis et al. (1994)).

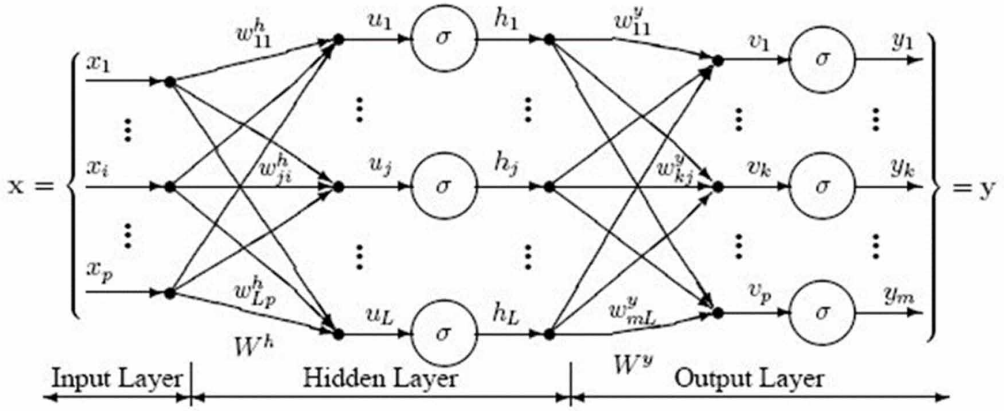
The image below is a visual depiction of a Multi-layer perceptron model:

This visual depiction is of a multi-layer perceptron model, a popular ANN structure (though more layers are possible, nonetheless there stands a significant increase in resources required for multiple hidden layers being employed). The multi-layer perceptron is constituted by an input layer, hidden layer and an output layer. The first layer being the input layer. The input layer processes and standardises a vector of the predictor variables ($x_1...x_p$). These variables then take the range of -1 to 1. These standardised values are then passed through to the hidden layer. This process loosely resembles the process of "saltatory conduction" with a biological neuron.

The standardised vector values of the predictor variables, for example standardised vectors of macroeconomic parameters such as inflation or GDP, reach the hidden layer. At the hidden layer, the standardised values are multiplied by a weight (w_{ji}), with the weighted values aggregated (u_j) and passed forward to the transfer function. The values from the transfer function are passed to the output layer (h_j) (D'Rosario, 2017b).

The output layer accepts the values from the hidden layer, multiplying the values by weights (w_{kj}), with the weighted values then aggregated resulting in a composite value v_j. The v_j value is then passed forward to the transfer function. The values from the transfer function are the model outputs y_k. Should the target variable be dichotomous, then there will be *2* neurons in the output layer producing *2* values,

Figure 1. Conceptual diagram of MPANN



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

one for each of the categories of the target variable. Where the output value takes greater than to values ($y_k = n$), then the output layer will be constituted by n values. The present MP-ANN outputs are denoted by the ordinal scales of the ratings agency ratings for the selected countries.

Beyond Parametric Regression, Can Neural Models Predict Credit Migration with Improved Accuracy

The use of such methods is not uncommon in financial studies but remains relatively underemployed within sovereign bonds ratings migration. There are several studies which employ OLS estimators to determine credit migration. Cantor & Packer (1996), Horrigon (1966) & Vaaler & McNamara (2004) are a few papers which have used OLS estimators to estimate credit migration. Given the simplicity of this approach, the model in some aspects is easy to interpret, but the main drawback of the model, if any, in setting a panel due to panel heteroskedasticity of the errors, which leads to inconsistent standard errors for the panel data. Due to such errors, panel corrected standard errors were used to accommodate for such variation. Beck & Katz (1995) had employed this approach to solve for the issue, but this method assumes that heteroskedastic errors/ disturbances are correlated across the panel. This is an issue since it fundamentally assumes correlation across the panel, which in part may not be true, i.e., for instance, tax reform, recession, GDP per capita a correlated and have similar effect across the panel, but may not be true when it comes to other variables in a model. The studies show that while quantitative data can be accounted for in the estimations, qualitative data such as, tax reform & liberalisation of trade can't be accurately accounted for in the model.

A related area of study by Elena et al. (2007) which uses transition matrices & hazard rate estimators to determine the risk of sovereign credit migration. However, the study tested the model for a Markov-Process, with spectral analysis and logit models for the transition matrix, both of which implied that it followed a non-Markovian process. But the Markovian process assumes that transition probability remains constant over time, which inherently means that the model can't accommodate for momentum and duration effects.

The use of Artificial Neural Nets in recent studies has increased partly due to the advancement of technology and deeper understanding of neural nets and lastly, not to ignore the fact that neural networks are more accurate than traditional statistical models. Zhang et al. (2007) which used neural networks in financial data mining, showed that it could effectively forecast financial time series data, although the accuracy was about 56.3%, nonetheless it still resulted in the neural network could achieve a 3-fold gain in profit as compared to buy and hold investors. It can also be noted that the model did better in bear market conditions than bull market, which suggests that the data mining system was better at managing risk.

Furthermore, studies carried out by Francisco Venegas-Martinez et al. (2013) which studied the effectiveness of ANN models with respect to financial time series data. In the paper, the authors posited that, the Multi-layer perceptron was the best ANN model when it came to forecasting financial time series data, which is the main reason it is widely used in the financial industry. Mainly due to the MLP being a feedforward model and the model inherently, unlike traditional statistical models, have no underlying assumptions.

Bhattacharya et al. (2006) had looked into the effectiveness of linear discriminant analysis (LDA) models versus ANN, which indicated that the ANN had outperformed the LDA by 79% to 33.33% when it came to a blind set test, although there were large unexplained variations in the models, it was mainly due to the fact that the data used were financial ratios and it suggests that CRA's don't just use financial

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ratios when it comes to corporate credit rating. This is not to say that the model has no technical variation, but it shows the effectiveness of ANN over LDA models. Nonetheless the study concluded by suggesting that the ANN can deal with more complex models and doesn't require inherent assumptions such as normality & linearity. Given that most of these studies that use ANN models, study financial time series forecasting & corporate credit migration, not many have used it to predict sovereign credit rating migration and this study in particular would shed light on that area of finance.

Model Structure and Parameters

The model seeks to predict credit migration on sovereign bonds, given a series of macroeconomic, legal and demographic variables. Specifically, the specification models the movement of sovereign bonds, either them being upgraded or downgraded. The model is a single layer network where Risk of credit migration, the dependant variable is a multinomial variable taking the values of the ratings scale converted into numeric form.

The ANN model employed — the MP-ANN model — is a multi-layer model with an input layer deriving from the variables provided by. The dependent carriable credit rating is the variable denoting the credit rating assigned by each of Moody's and S&P respectively. The rating agency is a variable used to denote which credit rating agency appraises the bond. Rating (lagged) is the latest or the most recent credit rating given to the bond. Inflation denotes the rate of inflation within each jurisdiction in average terms during the preceding year. A number of other macroeconomic variables are included in the specification, listed in the table summary below. All variables derive from the noteworthy study of Biglaiser & DeRouen (2007).

The first layer is the input layer. The input layer processes and standardises a vector of the predictor variables, herein the macroeconomic, governance and stability variables. These variables then take the range of -1 to 1. These standardised values are then passed through to the hidden layer. This process loosely resembles the process of "saltatory conduction" with a biological neuron.

The research data is of 16 Latin American countries for 1992-2003, these years were used since market reform had impacted nearly all the Latin American countries during this period. The table below (Table 1) illustrates summary statistics of the data set.

The Dependent Variables in the Model

The sovereign bond ratings are sourced from S&P & Moody's, was used since these rating agencies are the most commonly employed and utilised ratings service providers in the world and were used extensively during this period. Additionally, Bloomberg international (2005) was used to collect data on ratings migration changes of ratings of the above stated bonds by the credit rating agencies. Since the credit agencies followed a lettered system of rating bonds, it was normalized to a 17-point scale, with 0 being the lowest credit ("C") score and 16 being the highest ("AAA").

The Independent Variables in the Model

While a number of specifications were employed as part of the initial study estimations, The standardised vector values of the predictor variables, Lagged rating, tax rates, Rating Agency, COM (commercial index), FRI (financial reform index), PRI (privatisation index), CPI (unlagged), Current account balance

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Summary statistics (Moody's)

Variable	Observed	Mean	Standard Deviation	Minimum	Maximum
Moody's Bond rating	133	4.180451	2.551988	0	9
Trade reform	133	0.9377963	0.0270923	0.8112959	0.9978791
Financial reform	133	0.782716	0.1643583	0.5364308	0.9876156
Privatization	127	0.7986653	0.1877783	0.1432927	1
Tax reform	133	0.5805916	0.0998359	0.1186177	0.7736412
Capital opening	133	0.8510428	0.1157079	0.48	1
Polity (Democracy)	125	7.6	2.225222	0	10
Election year	125	0.232	0.4238076	0	1
Honeymoon	125	3.536	2.270228	1	10
President's ideology	124	0.2822581	0.4519242	0	1
IMF agreement	124	0.4596774	0.5003932	0	1
Inflation	125	170.0392	156.5154	0.0129808	1,022.401
Current account balance	124	2.568732	3.839382	14.22086	12.63721
GDP growth	125	2.343222	3.908493	10.89448	12.66971
GDP per Capita	125	3,616.058	1,955.649	933.8787	8,685.504
Bond default	133	0.2030075	0.4037588	0	1
External debt/GDP	117	0.3945228	0.1451298	0.18904221	0.129269

(% of GDP) (lagged), External Debt (lagged), GDP (Current), Budget Balance, GDP (constant 1995 US$, lagged), GDP growth (lagged) were found to be sound parameters for inclusion in the representative specification. These variables are consistent with the extant literature and the comparative study of

The hidden layer size was determined based on Heaton's work, noting optimality generally exists between the number of input layers and output layers (Heaton, 2011). Moreover, two hidden layers may be employed when modelling data that evidences discontinuities, such as a saw-tooth wave pattern. Given the nature of the underlying dataset, this is not the case. The use of an additional hidden layer reduces estimation efficiency and offers no improvement to the model. Herein, a single hidden layer model is employed to avoid a greater risk of converging to an incorrect local minima. One potential shortcoming of the selected methodology relates to the efficiency of ANN models in accommodating non-numeric variables, commonly referred to as dummy variables in economics. This is, however, only the case where non-numeric variables can take a large number of values, which is not the case in the present study. The total number of units employed in the model was 16, and determined through optimisation, again indicating interaction effects.

Data pruning techniques were not employed to improve model resolution, and the model performed efficiently and robustly absent of any node reduction. As such, the input values remain the same as those employed in Biglaiser & DeRouen, (2007). This is pertinent as it ensures that each predictive modelling framework can be compared on equal terms employing the same parameter set.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Variable descriptions

Variable Name	Abbreviation	Definition
Moody's Bond rating/ S&P Bond rating	Rating	These variables denote the score given to a (default risk) sovereign nation/government.
Financial reform	FRI	Financial reform being the policies/laws that influence investments and other capital flows into a nations' financial system, where liberal policy positively impacts the bond rating and protectionist policy have a negative impact on it.
Privatization	PRI	Denotes privatisation status of key utilities. Privatisation is the purchase of all outstanding shares of a publicly traded firm/ the sale of a state-owned firm into the private sector.
Tax reform	TAX	Changes in tax laws in a country.
Capital Account Liberalisation	CALI	A measure of international financial liberalisation.
Honeymoon	Honeymoon	Honeymoon referring to the newly elected officials have more freedom implement reform.
Polity (Democracy)	Polity	The political environment which consists of Political position of a nation (democratic/ Authoritarian), the president's ideology, IMF agreement & election cycles
Election year	Election	Elected year is election cycle of a nation, which represents the change of leadership.
President's ideology	Ideology	A scale measure of Presidential ideology. President's ideology also influences economic reform since leftist presidents oppose such reform which lead to job losses/price hikes.
IMF agreement	IMF	Variable denoting whether IMF agreement is in place. IMF agreements positively influence bond ratings since it signifies economic reform
Inflation	CPI	A sustained increase in general price levels measured by a basket of goods and services over several months or a fiscal year.
Current account balance	Current Account Balance	Consisting of the balance of payments, net transfers and net income.
GDP growth	GDP	Rate of change in economic production. Economic growth is an increase in the capacity of an economy to produce goods and services compared from one period to another.
GDP per Capita	GDP Per Capita	This measure compares the GDP on a purchasing power parity basis, which is divided by a country's population.
Bond default	Bond Default	Bond default measure whether a nation has previously defaulted or the percentage of bonds which have been written off as a loss.
External debt/GDP	External Debt	External debt/ GDP is the amount of outstanding debt that a nation borrows from the public/foreign investors and divided by the GDP of the nation.
Commercial Index	COM	This variable indexes measures trade liberalisation based on tariff policy.
Budget Balance	Budget Balance	This variable captures the government budgetary balance capturing whether it is in deficit or surplus.

MODEL FINDINGS

The MP-ANN is able to robustly predict the risk of credit migrating risk at 98.5% accuracy, while the model was trained, which showed that the increased accuracy and far greater goodness of fir than traditional Regression models. This finding suggests that the variables in the model are useful in determining the outcome/risk of credit migration. The model is capable of robustly and more consistently determining the outcome of credit migration than traditional regression model previously used in this area of study. Though it needs to be noted that when the model came to testing phase, its accuracy dropped slightly to 75.5%, which suggests that it still is better at predicting credit migration risk than traditional regression models. Table 2 shows the prediction accuracy of the model during training and testing phases.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Classification

Sample	Observed	Predicted											
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Percent Correct**
Training	0	3	0	0	0	0	0	0	0	0	0	0	100.0%
	1	0	3	0	0	0	0	0	0	0	0	0	100.0%
	2	0	0	18	1	0	0	0	0	0	0	0	94.7%
	3	0	0	0	17	1	0	0	0	0	0	0	94.4%
	4	0	0	0	0	16	0	0	0	0	0	0	100.0%
	5	0	0	0	0	0	22	0	0	0	0	0	100.0%
	6	0	0	0	0	0	0	21	0	0	0	0	100.0%
	7	0	0	0	0	0	0	0	20	0	0	0	100.0%
	8	0	0	0	0	0	0	0	0	4	0	0	100.0%
	9	0	0	0	0	0	0	0	0	0	2	0	100.0%
	10	0	0	0	0	0	0	0	0	0	0	9	100.0%
	Overall Percent	2.2%	2.2%	13.1%	13.1%	12.4%	16.1%	15.3%	14.6%	2.9%	1.5%	6.6%	98.5%
Testing	0	1	0	0	0	0	0	0	0	0	0	0	100.0%
	1	0	0	2	0	0	0	0	0	0	0	0	0.0%
	2	0	0	3	0	1	0	0	0	0	0	0	75.0%
	3	0	0	1	5	0	0	0	0	0	0	0	83.3%
	4	0	0	0	1	3	1	0	0	0	0	0	60.0%
	5	0	0	0	2	1	9	0	0	0	0	0	75.0%
	6	0	0	0	0	0	0	4	0	0	0	0	100.0%
	7	0	0	0	0	0	0	0	9	0	0	0	100.0%
	8	0	0	0	0	0	0	0	0	0	0	0	0.0%
	9	0	0	0	0	0	0	1	0	0	0	2	0.0%
	10	0	0	0	0	0	0	0	0	0	0	3	100.0%
	Overall Percent	2.0%	0.0%	12.2%	16.3%	10.2%	20.4%	10.2%	18.4%	0.0%	0.0%	10.2%	75.5%
Dependent Variable: Rating													

In terms of training the model, 70% of the data set was used to train the neural net and 30% was used to test the model. In essence, a neural network is ideal when it can be used to predict outcomes which it have not seen before. This makes the neural net a lot more accurate when it comes to estimating non-parametric models than traditional regression models.

As seen in Table 3 the correlation or r-squared for the estimated model is at 0.956, which suggests that the MP-ANN explain 95% of the variation in the estimates. Generally, when this is compared to a traditional regression model, the r-squared values of multi nominal logit models employed in comparable specifications within the earlier citation study achieve only modest goodness of fit relatively speaking

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

suggesting that the neural network is capturing significant threshold and interaction effects not captured within the comparable multinomial logit specifications. this suggests that the MP-ANN model is better when it comes to a measure of goodness-of-fit.

The goodness-of-fit measures suggest that the MP-ANN is superior to the logistic model in the Biglaiser & DeRouen study (Biglaiser & DeRouen, 2007). While the former study model remains empirically sound, it does not achieve the same level of predictive accuracy. The findings suggest that this model is more predictively capable, with an improvement in predictive accuracy of a minimum of 14 per cent. This is amongst the most notable contributions of the present study. The study makes clear the benefits associated with non-parametric methods of prediction in credit migration research and the benefit of MP-ANN modelling in the development of predictive models in the judicial sciences.

Employing the methods of Garson (1991) and Goh (1995), the normalised importance analysis (Figure 2) suggests that the lagged rating variable ranks as the most important determinant of credit ratings, followed by the TAX measure and the GDP variable. The data is presented graphically below. The variation in the hidden layer structure yields comparable results and robustly achieves a 100 percent predictive accuracy. Thought that it is notable that the relative ranked importance of GDP and PRI was reversed in some model specifications, lending credence to the assertion that government effectiveness may be of equal importance to the importance to the privatisation program of an economy when determining the credit rating predication. These findings are broadly consistent with the findings of the Biglaiser & DeRouen, (2007) study, with a few notable points of difference (n.b. Biglaiser & DeRouen, 2007).

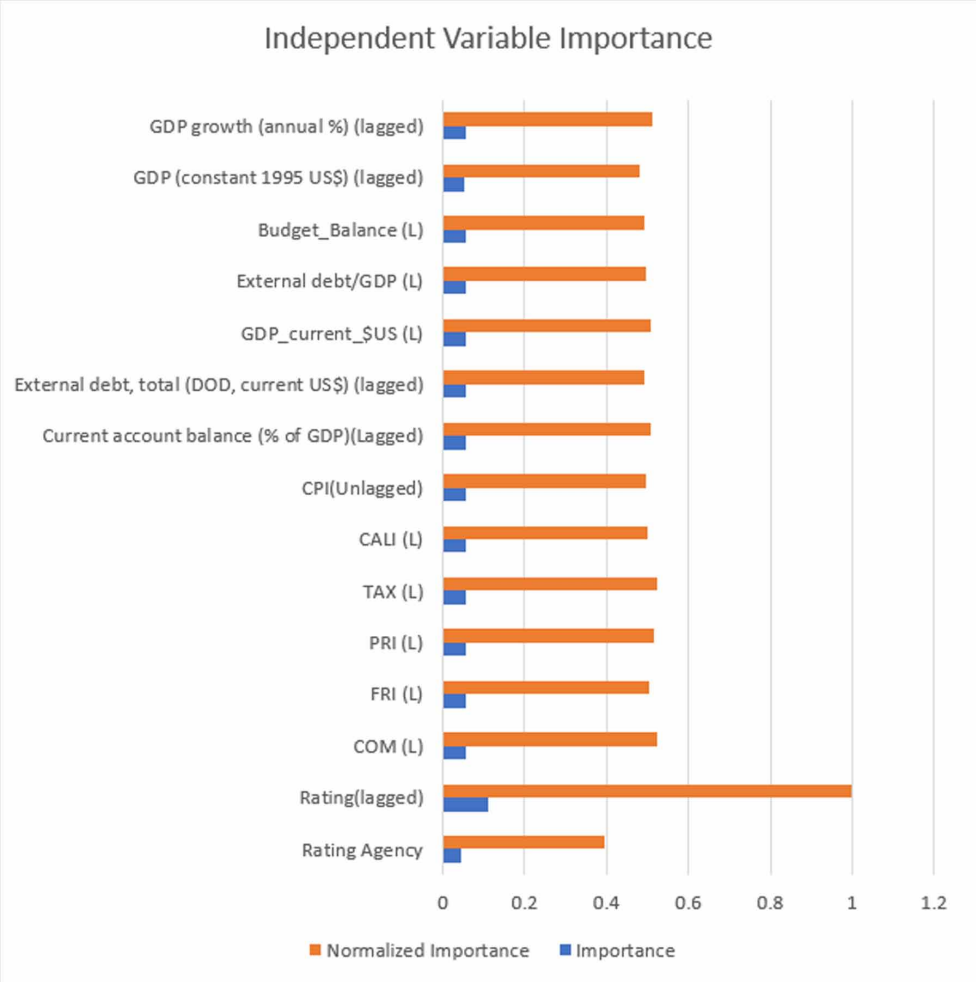
Firstly, unlike the Biglaiser & DeRouen study, the current study indicates that GDP, TAX, COM and as anticipated lagged rating and the ratings source (i.e. Standard & Poors, or Moody's, denoted by the variable ratings agency) rank as important to the ratings prediction. Additionally, the network results

Table 4. Normalized variable importance

	Importance	Normalized Importance
Rating Agency	.044	39.3%
Rating(lagged)	.112	100.0%
COM (L)	.058	52.2%
FRI (L)	.056	50.4%
PRI (L)	.058	51.5%
TAX (L)	.058	52.2%
CALI (L)	.056	49.9%
CPI(Unlagged)	.056	49.7%
Current account balance (% of GDP) (Lagged)	.056	50.6%
External debt, total (DOD, current US$) (lagged)	.055	49.4%
GDP_current_$US (L)	.057	50.7%
External debt/GDP (L)	.055	49.7%
Budget_Balance (L)	.055	49.4%
GDP (constant 1995 US$) (lagged)	.054	48.0%
GDP growth (annual %) (lagged)	.057	51.2%

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Independent variable importance analysis



suggest that threshold and interaction effects may exist between key independent variables GDP, COM and TAX. Again, noting that substantial improvement in pseudo R-Squared values ranging between 0.65 – 0.81 in contrast of the present study of 0.95 to 0.97.

CONCLUSION

The study has made clear the predictive gains associated with MPANN modelling and specifically non-parametric modelling in financial research. The findings indicate the predictive techniques utilised herein may be more beneficial than logistic models when emphasis is placed on outcome prediction rather than casual association which is pertinent to ratings migration. The study suggests that while logit and probit models are useful exploratory tools when engaging in empirical work and can accommodate non-linear problems; they appear to be less able to accommodate non-linear problems with complex interactions as efficiently as multilayer perceptron models. Moreover, they are incapable of accounting for thresh-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

old or interaction effects unless explicitly specified within the estimation particularly important when the desire is to engage in a predictive exercise. Further research is necessary to determine the potential generalizability of this claim. The notion that perceptron modelling may be beneficial and potentially superior in financial services and social sciences research when focused on predictive accuracy represents the first contribution of the study. A further contribution of the study is its response to the dearth of studies employing empirical methods to credit rating migration research, and in particular methods beyond logistic regression and linear discriminant analysis.

Importantly the current research displays the potential benefits associated with big data capture efforts when employed with modern parametric estimation techniques. The MP-ANN models may be highly effective tools in determining the relative importance of independent variables and the model of estimation to employ when engaging in casual analysis pertaining to rating migration. Moreover, the modelling methodology employed herein evidences the benefit of the presented empirical technique in predicting migration at high levels of accuracy with minimal variation in predictive outcomes relative to reality. Critically, the role of MP-ANN models in empirical financial research is advanced based on the notion that such models require heuristic knowledge pertaining to data preparation and cataloguing, and sufficient technical knowledge to employ the appropriate ANN for the estimation. The ANN models are more conceptually simple than linear modelling techniques (albeit more mathematically complex); not to mention seemingly more appropriate for modelling complex processes that replicate decision models, thus supportive of greater potential adoption by the financial services fraternity. It would be apt for future research to consider the potential application of probabilistic neural networks to such financial analysis of credit migration to discern its potential efficacy.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers and the editor for their insightful comments and suggestions.

REFERENCES

Ashley, K. D. (2017). *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age*. Cambridge University Press; . doi:10.1017/9781316761380

Baker, L. A. (1996). Interdisciplinary Due Diligence: The Case for Common Sense in the Search for the Swing Justice. *Southern California Law Review*.

Beck, N., & Katz, J. N. (1995). What to do (and not to do) with time-series cross-section data. *The American Political Science Review*, *89*(3), 634–647. doi:10.2307/2082979

Behr, P., & Güttler, A. (2008). The informational content of unsolicited ratings. *Journal of Banking & Finance*, *32*(4), 587–599. doi:10.1016/j.jbankfin.2007.04.021

Bex, F., Prakken, H., van Engers, T., & Verheij, B. (2017). Introduction to the special issue on Artificial Intelligence for Justice (AI4J). *Artificial Intelligence and Law*, *25*(1), 1–3. doi:10.100710506-017-9198-5

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Biglaiser, G., & DeRouen, K. Jr. (2007). Sovereign bond ratings and neoliberalism in Latin America. *International Studies Quarterly*, *51*(1), 121–138. doi:10.1111/j.1468-2478.2007.00442.x

Cantor, R., & Packer, F. (1996). Determinants and impact of sovereign credit ratings.

Elliott, J. A., & Kennedy, D. B. (1988). Estimation and prediction of categorical models in accounting research. *Journal of Accounting Literature*, *7*, 202–242.

Enns, K. M. (1998). Can a California Litigant Prevail in an Action for Legal Malpractice Based on an Attorney's Oral Argument before the United States Supreme Court? *Duke Law Journal*, *48*(1), 111–146.

Epstein, L., Ho, D. E., King, G., & Segal, J. A. (2005). The Supreme Court during crisis: How war affects only non-war cases. *New York University Law Review*, *80*(1).

Fei, F., Fuertes, A. M., & Kalotychou, E. (2012). Credit rating migration risk and business cycles. *Journal of Business Finance & Accounting*, *39*(1-2), 229–263. doi:10.1111/j.1468-5957.2011.02272.x

Fons, J. S., Kimball, A. R., & Kulakowski, S. M. (1970). Corporate bond defaults and default rates 1970. *Moody's Special Report*, January.

Garson, G. D. (1991). Interpreting neural-network connection weights. *AI Expert*, *6*(4), 46–51.

Goh, A. T. (1995). Back-propagation neural networks for modeling complex systems. *Artificial Intelligence in Engineering*, *9*(3), 143–151.

Gressman, E., & Stern, R. (2002). *Supreme Court Practice: For Practice in the Supreme Court of the United States* (8th ed.). Bureau of National Affairs.

Haldane, A. G., Hoggarth, G., Saporta, V., & Sinclair, P. (2005). Financial stability and bank solvency. In *Systemic Financial Crises* (pp. 83–113). Resolving Large Bank Insolvencies; . doi:10.1142/9789812569479_0007

Hanke, P. C. (2017). Computers with Law Degrees? The Role of Artificial Intelligence in Transnational Dispute Resolution, and its Implications for the Legal Profession. *Transnational Dispute Management, 14*(2).

Hill, P., Brooks, R., & Faff, R. (2010). Variations in sovereign credit quality assessments across rating agencies. *Journal of Banking & Finance*, *34*(6), 1327–1343. doi:10.1016/j.jbankfin.2009.11.028

Ho, D. E. (2002). Compliance and International Soft Law: Why do countries implement the basle accord? *Journal of International Economic Law*, *5*(3), 647–688. doi:10.1093/jiel/5.3.647

Hoggarth, G., Reidhill, J., & Sinclair, P. J. (2004). On the resolution of banking crises: theory and evidence.

Horrigan, J. O. (1966). The determination of long-term credit standing with financial ratios. *Journal of Accounting Research*, *4*, 44–62. doi:10.2307/2490168

Hunter, D. (1994). Looking for law in all the wrong places: Legal theory and legal neural networks. *Legal knowledge based systems: The relation with legal theory*, 55-64.

Hunter, D. (1999). Out of their minds: Legal theory in neural networks. *Artificial Intelligence and Law*, *7*(2-3), 129–151.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Jarrow, R. A., Lando, D., & Turnbull, S. M. (1997). A Markov model for the term structure of credit risk spreads. *Review of Financial Studies*, *10*(2), 481–523. doi:10.1093/rfs/10.2.481

Kannai, R., Schild, U., & Zeleznikow, J. (2007). Modeling the evolution of legal discretion – an Artificial Intelligence Approach. *Ratio Juris*, *20*(4), 530–558. doi:10.1111/j.1467-9337.2007.00374.x

Karnin, E. D. (1990). A simple procedure for pruning back-propagation trained neural networks. *IEEE Transactions on Neural Networks*, *1*(2), 239–242.

Kavvathas, D. (2001). Estimating credit rating transition probabilities for corporate bonds.

Kim, M. J., Kang, D. K., & Kim, H. B. (2015). Geometric mean based boosting algorithm with over-sampling to resolve data imbalance problem for bankruptcy prediction. *Expert Systems with Applications*, *42*(3), 1074–1082. doi:10.1016/j.eswa.2014.08.025

Knight, J. (1997). *The Choices Justices Make*. Congressional Quarterly Inc.

Kumar, K., & Bhattacharya, S. (2006). Artificial neural network vs linear discriminant analysis in credit ratings forecast: A comparative study of prediction performances. *Review of Accounting and Finance*, *5*(3), 216–227. doi:10.1108/14757700610686426

Lando, D., & Skødeberg, T. M. (2002). Analyzing rating transitions and rating drift with continuous observations. *Journal of Banking & Finance*, *26*(2-3), 423–444.

Lee, K., Booth, D., & Alam, P. (2005). A comparison of supervised and unsupervised neural networks in predicting bankruptcy of Korean firms. *Expert Systems with Applications*, *29*(1), 1–16. doi:10.1016/j.eswa.2005.01.004

Makridakis, S. (2017). The forthcoming artificial intelligence (AI) revolution: Its impact on society and firms. *Futures*, *90*, 46–60. doi:10.1016/j.futures.2017.03.006

Martin, A. D., & Quinn, K. M. (2002). Dynamic ideal point estimation via Markov chain Monte Carlo for the US Supreme Court, 1953–1999. *Political Analysis*, *10*(2), 134–153.

Martin, A. D., Quinn, K. M., & Epstein, L. (2004). The median justice on the united states supreme court. *North Carolina Law Review*, *83*, 1275.

Maskus, K. (2002). Intellectual Property Issues for the United States and Japan: Disputes and Common Interests. *Issues and Options for US–Japan Trade Policies*, 165-90.

McAtee, A., & McGuire, K. T. (2007). Lawyers, justices, and issue salience: When and how do legal arguments affect the US Supreme Court? *Law & Society Review*, *41*(2), 259–278.

McGreal, S., Adair, A., McBurney, D., & Patterson, D. (1998). Neural networks: The prediction of residential values. *Journal of Property Valuation and Investment*, *16*(1), 57–70. doi:10.1108/14635789810205128

McGuire, K. T. (1995). Repeat players in the Supreme Court: The role of experienced lawyers in litigation success. *The Journal of Politics*, *57*(1), 187–196.

Minsky, M., & Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry*. Cambridge, MA: MIT Press.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Nissan, E. (2017). Digital technologies and artificial intelligence's present and foreseeable impact on lawyering, judging, policing and law enforcement. *AI & Society*, *32*(3), 441–464. doi:10.100700146-015-0596-5

Odom, M. D., & Sharda, R. (1990, June). A neural network model for bankruptcy prediction. In *IJCNN International Joint Conference on Neural Networks* (pp. 163–168). IEEE; . doi:10.1109/IJCNN.1990.137710

Olden, J. D., & Jackson, D. A. (2002). Illuminating the "black box": A randomization approach for understanding variable contributions in artificial neural networks. *Ecological Modelling*, *154*(1-2), 135–150.

Patterson, D. W. (1998). *Artificial neural networks: theory and applications*. Prentice Hall PTR.

Prakken, H. (2017). On the problem of making autonomous vehicles conform to traffic law. *Artificial Intelligence and Law*, *25*(3), 341–363. doi:10.100710506-017-9210-0

Refenes, A. N., Zapranis, A., & Francis, G. (1994). Stock performance modeling using neural networks: A comparative study with regression models. *Neural Networks*, *7*(2), 375–388. doi:10.1016/0893-6080(94)90030-2

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). *Learning internal representations by error propagation (No. ICS-8506)*. California Univ San Diego La Jolla Inst for Cognitive Science.

Sankar, A., & Mammone, R. J. (1991, July). Optimal pruning of neural tree networks for improved generalization. In *IJCNN-91-Seattle International Joint Conference on Neural Networks* (Vol. 2, pp. 219-224). IEEE.

Schöneburg, E. (1990). Stock price prediction using neural networks: A project report. *Neurocomputing*, *2*(1), 17–27. doi:10.1016/0925-2312(90)90013-H

Sheehan, R. S., Mishler, W., & Songer, D. R. (1992). Ideology, status, and the differential success of direct parties before the Supreme Court. *The American Political Science Review*, *86*(2), 464–471.

Songer, D. R., & Sheehan, R. S. (1992). Who wins on appeal? Upperdogs and underdogs in the United States courts of appeals. *American Journal of Political Science*, 235–258.

Stranieri, A., & Zeleznikow, J. (2006). *Knowledge discovery from legal databases* (Vol. 69). Springer Science & Business Media.

Stranieri, A., Zeleznikow, J., Gawler, M., & Lewis, B. (1999). A hybrid—neural approach to the automation of legal reasoning in the discretionary domain of family law in Australia. *Artificial Intelligence and Law*, *7*(2-3), 153–183. doi:10.1023/A:1008325826599

Toulmin, S. (1958). *The Uses of Argument*. Cambridge: Cambridge University Press.

Vaaler, P. M., & McNamara, G. (2004). Crisis and competition in expert organizational decision making: Credit-rating agencies and their response to turbulence in emerging economies. *Organization Science*, *15*(6), 687–703. doi:10.1287/orsc.1040.0089

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Wei, J. Z. (2003). A multi-factor, credit migration model for sovereign and corporate debts. *Journal of International Money and Finance*, *22*(5), 709–735. doi:10.1016/S0261-5606(03)00052-4

Weigend, A. S., Rumelhart, D. E., & Huberman, B. A. (1991). Back-propagation, weight-elimination and time series prediction. In Connectionist models (pp. 105-116).

Wolff, T., & Richardson, M. (1989, January). Fault detection in structures from changes in their modal parameters. In *Proceedings of the 7th international modal analysis conference* (Vol. 1, pp. 87-94).

Zhang, D., Jiang, Q., & Li, X. (2004). Application of neural networks in financial data mining. *International Journal of Computational Intelligence*, *1*(2), 116–119.

This research was previously published in the International Journal of Strategic Decision Sciences (IJSDS), 9(4); edited by Madjid Tavana; pages 70-85; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX: ADDITIONAL SUMMARY STATISTICS

Table 5. S&P variable summary statistics

Bond rating	142	4.514085	2.477519	0	11
Trade reform	142	0.9415874	0.0240344	0.8786718	0.9978791
Financial reform	142	0.7873715	0.162292	0.5364308	0.9876156
Privatization	137	0.8056366	0.1793056	0.1432927	1
Tax reform	142	0.5870538	0.0985983	0.1186177	0.7736412
Capital opening	142	0.8497998	0.1082285	0.53	1
Polity (Democracy)	133	7.56391	2.064686	0	10
Election year	134	0.2164179	0.4133476	0	1
Honeymoon	134	3.492537	2.293326	1	10
President's ideology	133	0.2406015	0.4290648	0	1
IMF agreement	133	0.443609	0.4986882	0	1
Inflation	134	166.002	150.8112	2.76845	1,022.401
Current account balance	133	2.4868	3.603607	14.22086	12.63721
GDP growth	134	2.707314	3.964061	10.89448	12.27793
GDP per Capita	134	3,532.304	1,906.76	933.8787	8,685.504
Bond default	142	0.1830986	0.3881164	0	1
External debt/GDP	129	0.4038407	0.1658139	0.1962219	1.219388

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 6
A Thorough Insight into Theoretical and Practical Developments in MultiAgent Systems

Dimple Juneja
Dronacharya Institute of Management and Technology, India

Aarti Singh
Maharishi Markandeshwar University, India

Rashmi Singh
Banasthali University, India

Saurabh Mukherjee
Banasthali University, India

ABSTRACT

Multiagent systems have been a fascination for research community and are often seen as an intelligent solution to many complex real world problems. Researchers have been active in the domain since last three decades and many developments pertaining to theoretical design and practical developments of multiagent systems are worth appreciating. The growth in MAS is multidirectional ranging from conceptual ideas to practical implementations and from the wide range of applications; it appears that multiagent systems are proving to be universal. The paper presents a concise survey of developments in MAS highlighting the important contributions in the field and also questions the universal applicability of agents.

INTRODUCTION

With the transformation from individual computing technologies to self-motivated distributed systems, a lot many research opportunities and threats have evolved. Dynamic and open environments (Shakshuki, Ghenniwa & Kamel, 2003) extent across business boundaries where in heterogeneous systems (Sycara, 1998) not only cooperates but also operate reactively and proactively. Such systems are capable of adapting to rapidly changing environments and are also able to respond to complex queries in real time. However, such systems demand high degree of autonomy (Barber & Martin, 1999; Lange & Oshima, 1999; Suzanne

DOI: 10.4018/978-1-7998-0951-7.ch006

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Barber, Goel & Martin, 2000) which actually becomes a threat. For instance, web services (Alberti, Chesani, Daolio, Gavanelli, Lamma, Mello & Torroni, 2007) are the most sought after way of carrying out business activities as it offers service oriented architecture where in divergent and self-governing software components interact to offer automated solutions. Such automated systems are employing software entities more commonly known as software agents (Mařík & Lažanský, 2007; Pěchouček & Mařík, 2008) that offer powerful solutions to the emerging issues and are also able to manage the complexity of entire system. Since 1990, agent-based computing (Graesser, Jeon & Dufty, 2008; Nwana & Wooldridge, 1996; Nwana, 1996) is attracting community of researchers and lot many developments and changes can be seen in the technology. Simply speaking, a software agent is a software component capable of carrying out rational and self-directed actions in a typical multi-agent system (Panait & Luke, 2005; Hoek, Wiebe & Wooldridge, 2008). Agentsare finding applicability in a miscellaneous array of domains including intelligent manufacturing (Shen, Hao, Yoon & Norrie, 2006), traffic and transportation systems (Chen &Cheng2010), communications (Manvi & Venkataram, 2004), industrial process control (Mařík & Lažanský, 2007, Metzger & Polakow, 2011), Military applications (Cioppa, Lucas & Sanchez, 2004), geographical systems (Heppenstall, Crooks, See & Batty, 2011), optimization problems (Barbati, Giuseppe& Andrea, 2012), disaster management (Fiedrich & Burghardt, 2007)and many more.

AGENT BASED COMPUTING: IS IT SUPREME THAN ITS COUNTERPARTS?

The concept of software agents have evolved from Distributed Artificial Intelligence (DAI) (Ferber, 1999) research which was conducted about 3 decades ago.Carl Hewitt proposed an Actor system (Hewitt, Bishop & Steiger, 1973)and defined actors with clear internal definitions with the potential to communicate with peers in the system. The succeeding developments focused on making intelligent actors now known as software agents where the word agent has evolved from Greek word '*agein*', which implies to guide (Milojicic, 2000). The term is being explored by research fraternity continuously and many reports pertaining to the developments in agent oriented techniques that enable a more active role of the computers in knowledge acquisition have been contributed. In the domain of computer science, a software agent is a special software entity dedicated to perform an assigned task on behalf of a human or an object (Nwana, 1996).

Following definition has been adopted from (Smith, Cypher & Spohrer, 1994) and the agent is known as KidSim Agent, in particular.

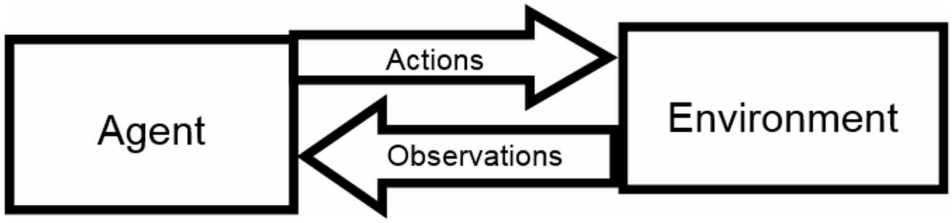
…an agent is a persistent software entity dedicated to a specific purpose. 'Persistent' distinguishes agents from subroutines; agents have their own ideas about how to accomplish tasks, their own agendas. Special purpose' distinguishes them from entire multifunction applications; agents are typically much smaller. (Smith, Cypher & Spohrer, 1994)

Franklin and co-authors define an agent as

… special software entities carry out some set of operations on behalf of a user or another program with some degree of independence or autonomy, and in so doing, employ some knowledge or representation of the user's goals or desires… (Franklin & Graesser, 1996)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Agent Interaction Environment



Defining it more precisely, A Software Agent is

… an object of the environment that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators… (Russell, Norvig & Intelligence, 1995)

Figure 1 represents the agent interaction environment.

Study in this domain reveals that the agent paradigm is an extension rather than a replacement of conventional systems that are either object-oriented or component-based. The software group at MIT[1] compared and contrasted software agents to conventional software and concluded that software agents have longer life, partially autonomous and these can be designed to behave proactive as well as adaptive.

Another popular definition of agent states that

… Autonomous agents are computational systems that inhabit some complex dynamic environment, sense and act autonomously in this environment and by doing so realize a set of goals or tasks for which they are designed… (Maes, 1995)

According to the definitions, software agents are good to be employed as subordinates, memorize important tasks, wisely recapitulate intricate data, acquire knowledge from user's actions and hence recommend on the topics of user's interest (Maes, 1994). Realm of software agents has fascinated the researchers to a great extent because of its significant properties (Bauer, Jörg & Odell, 2001) (see Table 1) as compared to object oriented and expert systems, which make it stand ahead of its counterparts.

Though the concept of software agent looks much like the conventional objects of OOP, it differs from later in terms of being autonomous or the self-initiating-property. However, an agent may or may not possess all the above-mentioned characteristics for acting in an agency. They can in fact, function efficiently even by possessing some of the features. The taxonomy of agents, depending upon the properties they possess is given in the next section. Since the origin of agent-based computing, researchers have shown non-consensus over defining agents and different research community has been defining them according to the situation in which these have been employed. In general, an agent situated in an environment should be autonomous and flexible (Franklin & Graesser, 1996).

Here, situated in an environment implies agent is proficient of accepting input from the environment and hence carry out the actions which may change the state of environment. The actions may be carried out autonomously without direct intervention of human user (Russell, Norvig & Intelligence, 1995). Most popular and common examples of such intelligent systems are simple thermostats (Jennings, Sycara &

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Comparison of agents, objects and expert systems

Entities / Properties	Agent	Object	Expert system
Social	Include social ability	Have social ability	Lack social ability
Autonomous	Can work independently	Need command to invoke	Can work on their own
Reactive	React to changes in their environment	React to change in input values	Work on input provided, only when commanded
Adaptable	Change as per preferences of the user	Work as per the input values	Work as per the input values
Intelligent	Can handle uncertain or new situations	Can work only for specified situations	Can handle uncertainty to some extent
Goal Oriented	Plan actions to achieve a goal	Works to solve a problem	Works towards specified problem
Mobile	Can move with its data and instructions	Can't port	Can migrate
Learning Ability	Can learn from user behavior	Don't have provision to learn or change on its own	Can learn from user behavior so can update itself
Pro-Activity	Can initiate actions to change the environment	Works only on command	Can initiate actions

Wooldridge,1998). However, these systems offer no flexibility so these are not pure agent based systems. A flexible system must be either reactive or proactive and should respond to change in state of the system well within time. In order to respond actively and accurately, the agent should be social and when required should interact with its environment including humans to resolve a problem. Some other desired facets of being agents are mobility, adaptability, learning and responsive. It is these properties that make agent-based computing more supreme than its counterparts such as object-oriented systems, distributed systems, and expert systems (Jennings, Sycara, & Wooldridge, 1998).

It is obvious from the above discussion that an agent based system comprises of many agents making it to be a multi-agent system (Vigna, 2004; Weiss, 1999; Wooldridge, 1998) (see Figure 2) where many agents cooperate, coordinate, negotiate and communicate to solve problems. Numerous heterogeneous agents offer variety of problem solving methods with different viewpoints. In fact, these flexible interactions make MAS more powerful than conventional software systems.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

MULTIAGENT SYSTEMS: THE THEORETICAL DEVELOPMENTS

The emergence of software agents has offered a different perspective of understanding and hence developing the complicated software systems. In fact, the technique offers new tools and methods to design the software applications ranging from simple to complex real time control systems. A multi-agent system encapsulates the complexity of the complex software systems and hence presents an abstract and a simple view to its users. Researchers (Laghari & Niazi, 2016; Sim, 2012) are finding the field interesting and available literature (Barbati, 2012; Barbati, Giuseppe & Andrea, 2012; Dam, Governatori, Ito & Xu, 2016) indicates that lot many efforts are being put to improve the computing. On contrary, few works (Brooks,1991a; Brooks, 1991b; Chapman, 1987; Jennings, Sycara & Wooldridge, 1998; Vigna, 2004) also reflect that with the level of activity going on, the computing may become frenzy and disjoint. The section talks about the developments in MAS restricting to physiognomies, categorization, agent communication and the concept of ontology.

Physiognomies of Agents

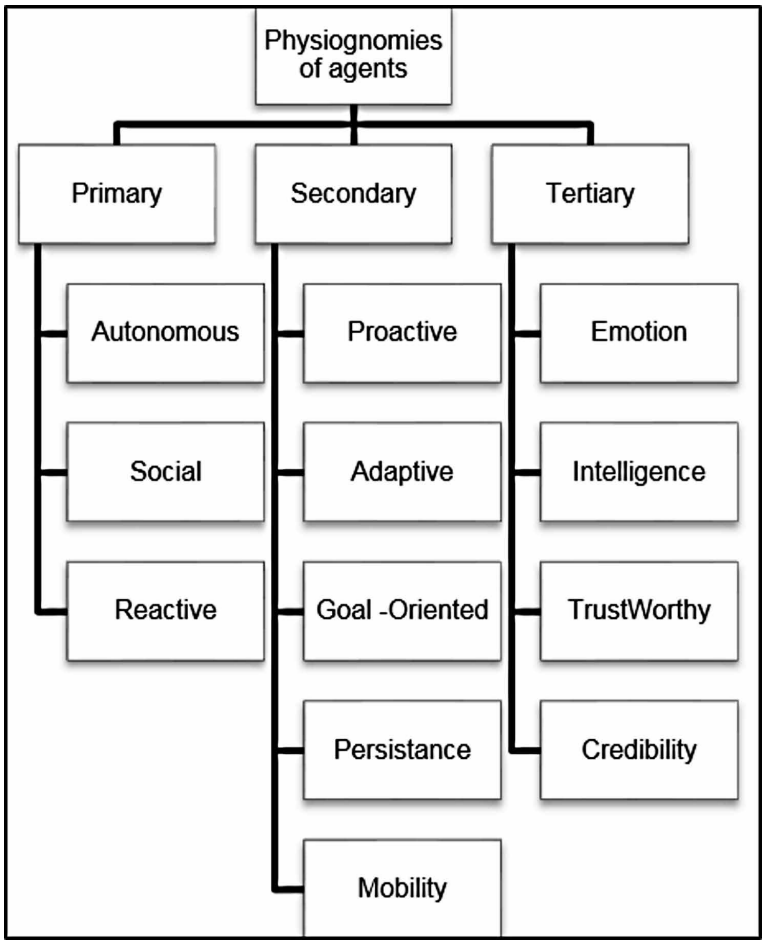
The penetrating agent based approaches demanded for a concrete definition and the classification of agents. The developments projected many independent and application specific definitions (Genesereth & Ketchpel, 1994). At present, several definitions exist however, each definition reflects that all agents are expected to possess some basic characteristics that define the behavior of every agent. For instance, BDI agents (Rao & Georgeff, 1995; Sharma & Juneja, 2006) must show explicit beliefs, desires and intentions (Jennings, 2000), while Woodridge and Jennings (Wooldridge & Jennings, 1995) desire agents to be independent, responsive and societal. In addition, Franklin & Graesser (1996) lists agents as self-directed, responsive, communicative, learning, portable, rational, goal-oriented, and unremitting and should also have human like emotions. Using such definitions from the literature, it has been possible to create a set of primary, secondary and tertiary agent characteristics. Primary characteristics are inherent to most of the popular agent definitions while secondary characteristics are the set of extended characteristics usually associated with agents. The third set of characteristics contains more abstract, desirable, human-like features. Figure 2 illustrates the physiognomies of agents thus summarized from literature and further discussion regarding the same is being overlapped in the next section.

Categorization of Software Agents

Scientists have suggested numeral contemporary categories and criterion required for an entity to be termed as software agent. For instance, (Nwana, 1996) identifies three key attributes namely autonomy, learning and cooperation. The author further classifies these agents as collaborative agents, reactive agents, interface agents, information agents, hybrid agents, mobile agents and ants. Wooldridge and Jennings (Wooldridge & Jennings, 1995) distinguish agents to be weak agents and strong agents. While a weak agent is self-directed, societal, and responsive and takes the initiative, the strong agent follows Belief-Desire-Intention (BDI) model. Further, (Jennings, 2000) define an agent as "a computer system in an environment that can take autonomous actions to meet the objectives." Magedanz, Rothermel & Krause (1996) divides agents as local agents, networked agents and mobile agents, DAI-based agents and MultiAgent Systems (MAS) respectively. In addition, there also exist advisory agents acting as personal assistants (Nangle, Cunningham & Evans, 1998) supporting web browsing, searching etc. Number of

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Physiognomies of Agents



researchers has classified agents as task-specific and performative agents (Etzioni & Weld, 1994), (Heger & Koen, 1991), (Sycara, Pannu, Williamson, Zeng & Decker, 1996) such as search agents, navigation agents, information agents, softbots, knowbots, infobots, autobots, webbots and so on. Depending upon the responsibilities to be carried out and aims to be achieved, agents with different attributes is designed.

A critical look at the literature highlights that agents are classified (Brenner, Zarnekow & Wittig, 2012; Franklin & Graesser, 1996; Nwana, 1996 & Wooldridge; Nwana, 1996; Singh, Gulati & Niranjan, 2012) according to the behavior they present within and outside a multi-agent system. For instance, a collaborative agent cooperates with its peers to perform an assigned task. They possess social and communicative abilities. Every agent has some knowledge and experience of a particular domain and can also seek services of other agents having specialization in other domains. An interface agent provides an interface to users without explicit instructions from the user. These are possessed with autonomy, responsiveness, skill and user-friendly characteristics. The interface agent accepts input from the user and may delegate the same to other agents changing the state of environment. The interface agent might be designed to have learning abilities and hence can directly take decisions by observing user input over

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

long durations. Information agents gather information from various resources including the environment in which these are situated and support the smooth conduct of task. To an extent such agents can be treated as intelligent and hence are designed to be proactive, societal and autonomous. Reactive Agents upholds the continuing communication with the environment and peer agents as well. These are required to reactwell within time. An agent does not reflect the long-term effects of its actions, and does not consider the coordination of activity with other agents. A mobile agent is a special agent which in addition to the basic properties of agents also has the special ability to migrate in a distributed network of nodes. It carries code and data with itself. On completing the task in the destination node, such agents are required to return to the originating node. Ants are another type of agents which are alike mobile agents i.e. are possessed with ability to travel from hop to hop but the former ones can stay at the destination node even after task completion and during the next event, the last node visited becomes the originating node. A new form of software agents is known as SMART (System for Managing Agents in Real Time) agents. Such agents are self-organizing and agree to be the part of an entire multiagent system. The intelligent agents are agents with learning capabilities and exhibit adaptive behavior. They may or may not have social abilities. In addition to intelligent agents, developers are now designing agents having cognitive behavior and have named such agents as cognitive agent. The cognitive agents are being designed with the aim to work with minimum user intervention and should be self-capable similar to human beings.

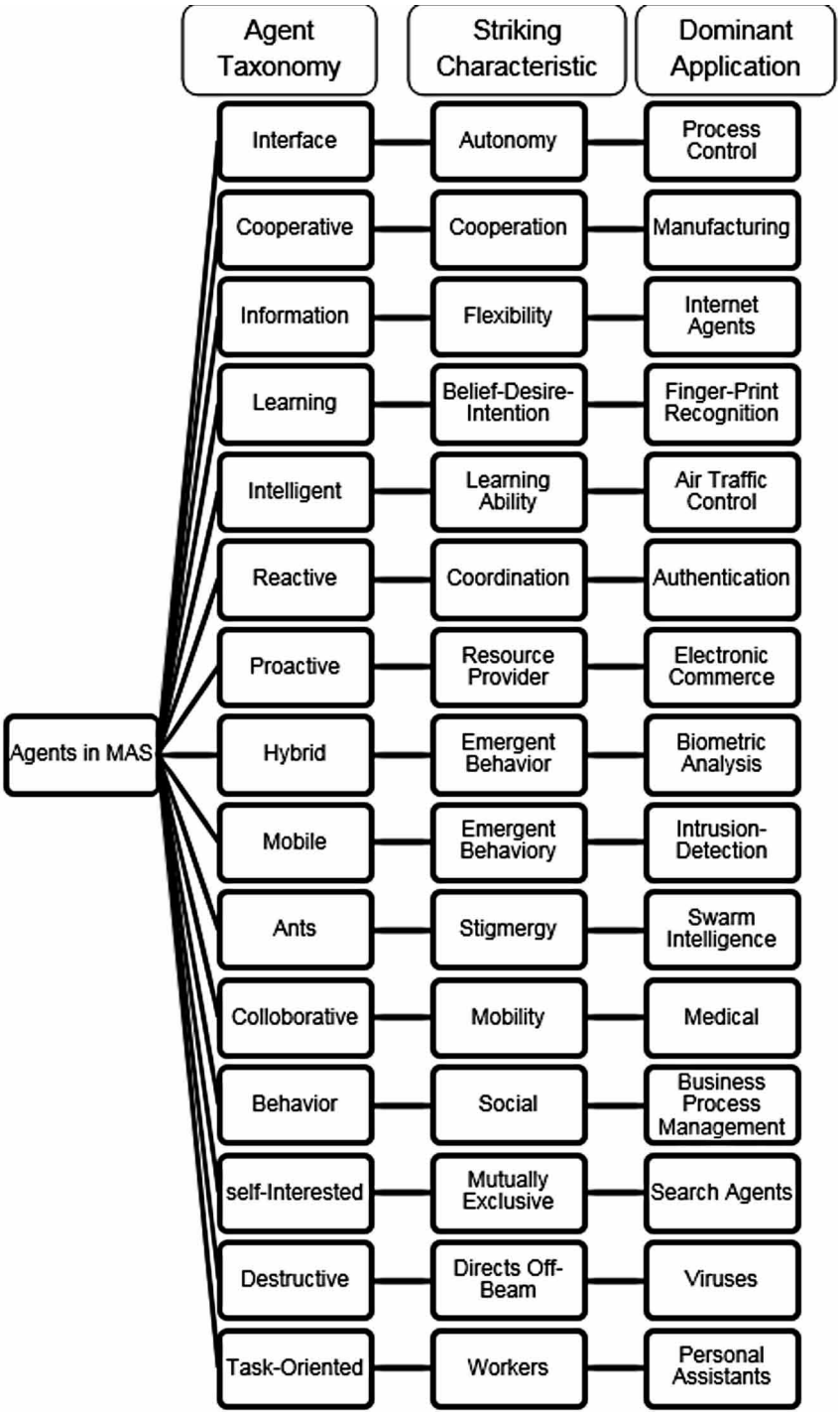
Considering the above significant types of agents, it is evident that over the years the big community of information has been restructured into smaller elements called agents that are now possessed to have many abilities well suited to handle complex applications. In other words, groups of software agents can be treated as a dynamic collection of simple agents with self-describing interfaces who while progressing through their life span continue to collect the information in a "knowledge soup," and same is presented to the user according to the application in which these agents operate. In such an environment, individual agents do not manage the entire system to achieve the target where as they would interact and collaborate actively forming a multi-agent system (MAS) (Sycara,1998). In such a system agents integrate to provide the needed information and such a system in which a community of software agents works in collaboration and coordination with each other to realize a shared objective is called an agent-based system. It has evolved as a result of sociological relationships among agents. In MAS, at least one agent has an ascribed set of goals and other agents in the system may adopt the goal. On the basis of above study, Figure 3 concludes following types of prominent agents and their most striking characteristics.

Agent Communication and Interaction

In order to maximally utilize the potential of agents, agents form a multi-agent system and moreover, agents in a multi-agent system and an entire system in turn must communicate and interact to realize shared targets. Further, in order to establish the communication, various agents in MAS need to interact and hence these need a language to communicate, a commonly agreed upon ontology and also are required to follow certain rules known as Interaction Protocols. Upcoming sections briefly extract the literature to discuss the three prime components listed above (see Figure 4).

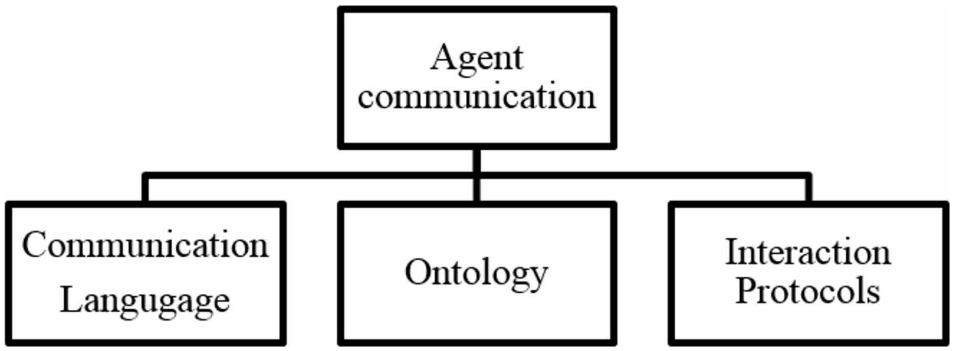
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Taxonomy of Agents with Characteristics and Applications



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Components of Agent Communication



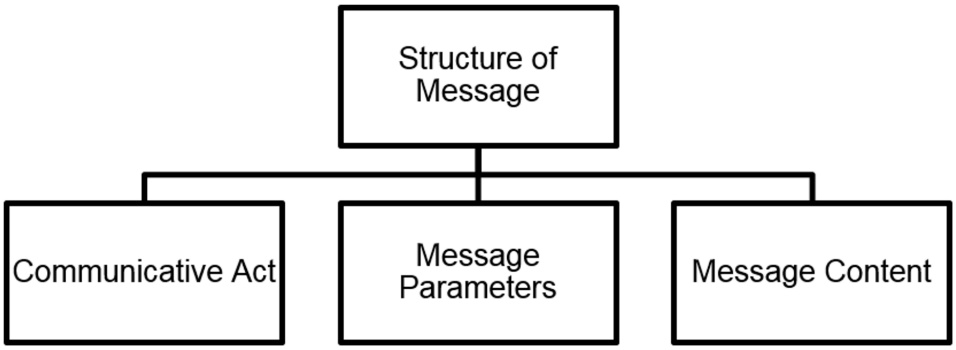
The Communication Languages

The communication languages for agent, more popularly known as Agent Communication Languages (ACL) (Kone, Shimazu & Nakajima, 2000) are utilized to exchange facts and figures among agents in multi-agent environment. ACL provides an interface that allows diverse agents to communicate with semantically true statement that deliver information about their environments. Although literature reflects many developments (Chaib & Dignum, 2002), (Labrou, Finin & Peng, 1999) in the domain of ACL but the most popular languages are FIPA-ACL, KQML, KIF.

A) FIPA-Agent Communication Language

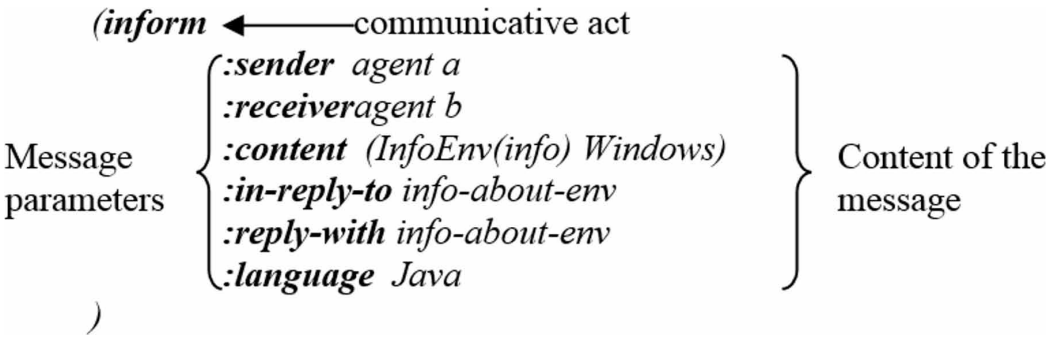
Specification for FIPA-ACL was proposed by Foundation for Intelligent Physical Agents (FIPA). The language enables interoperability between different agents and is adopted as the standard language for agent communication (Juneja, Jagga & Singh, 2015). Speech act theory forms the underlying backbone for FIPA-ACL and alike speech act theory, makes use of communicative acts for exchanging messages. A FIPA-ACL message comprises three components shown in Figure 5 and same is being illustrated through the example shown in Figure 6.

Figure 5. Structure of FIPA-ACL Message



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Example of FIPA-ACL message

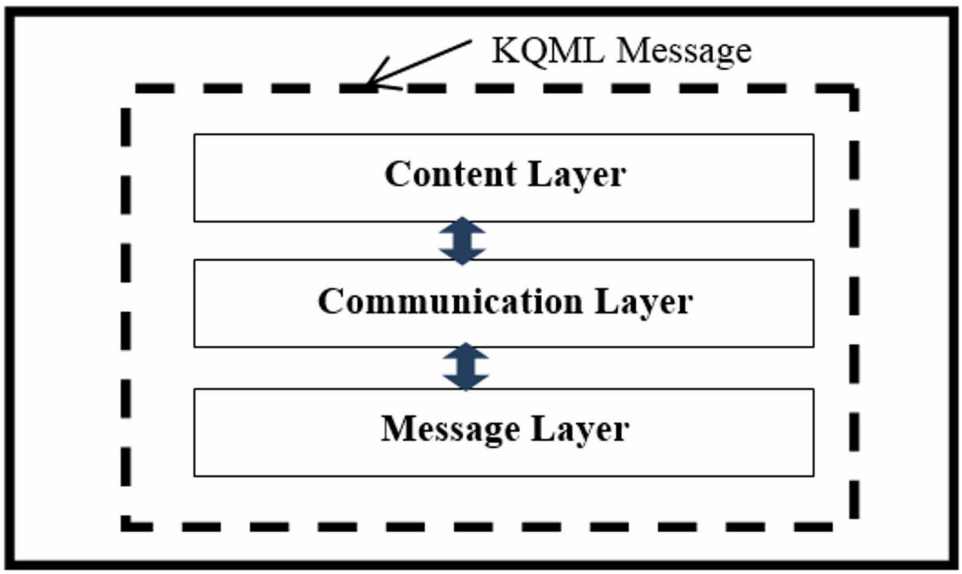


Here, inform is the communicative act,:sender,:receiver,:content, etc., are the message parameters and agent a, agent b etc. represents the content of message. It shall be noted that every FIPA-ACL message begins and ends with a round bracket. The complete list of reserved message parameters along with their description permitted by FIPA-ACL is given in Table 2 and Table 3 lists the communicative acts defined in FIPA-ACL. For more details, readers are deferred to technical document by FIPA[2].

B) Knowledge Query and Manipulation Language

KQML (Finin, Fritzson, McKay &McEntire, 1994; Labrou & Finin, 1997) is the result of further efforts by research community to support agent communication. KQML is not only a language but also a protocol that supports sending and receiving general purpose messages among distributed agents. It supports exchange of messages which are syntax and ontology independent. Primarily a KQML message comprises of three layers i.e. content, communication and message as shown in Figure 7.

Figure 7. Layers of a KQML Message



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The content layer bears the actual content of the message in the program's own representation language. KQML can carry any representation language, including languages expressed as ASCII strings and those expressed using binary notation.

Every KQML implementation ignores the content portion of the message, except to determine where it ends. The communication layer encodes a set of features of the message that describe the lower-level communication parameters, such as the identity of the sender and receiver, and a unique identifier associated with the communication. This layer is the core of KQML and determines the kind of interactions one can have with a KQML-speaking agent. Function of this layer is to identify the speech act or performatives that the sender attaches with the content. This speech act indicates whether the message is an assertion, a query, a command, or any other from a set of known performatives, where a performative stands for primitive message type. Additionally, as the content is hidden from KQML, the message layer also includes optional features that describe the content language, the ontology it assumes, and some type of description of the content, such as a descriptor naming atopic within the ontology. These features make it possible for KQML implementations to analyze and properly deliver messages with hidden contents. Alike FIPA-ACL, the structure of a KQML message is being illustrated with the example given in Figure 8.

Alike FIPA-ACL, a KQML message also begins and ends with a round bracket. It also has some reserved parameters (see Table 4) mentioned within a predefined performative. In this message, the KQML performative is *ask-one*, the content is *(InfoEnv?info)*, the ontology assumed by the query is identified by the token *Operating System,* the receiver of the message is agent 2 and the query is written in a Java. The detailed list of KQML performatives is given in Table 5.

C) Knowledge Interchange Format (KIF)

KIF (Genesereth& Fikes, 1992) is based on first order predicate logic and provides a high-level intermediate language for communication between agents. It is capable of expressing facts and rules. It provides constructs for describing procedures, i.e. programs to be executed by an agent. KIF specifies its syntax and semantics similar to Lisp and is more or less alike FIPA-Semantic language (FIPASL) (McArthur, Davidson, Catterson, Dimeas, Hatziargyriou, Ponci & Funabashi,2007).

Table 2. Description of Message Parameters

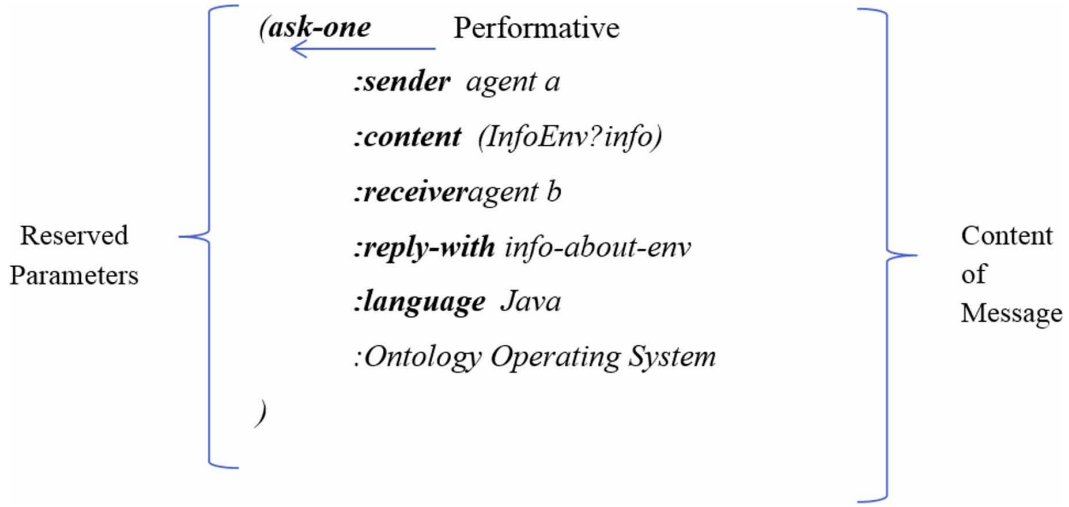
Parameter	Description
:sender	Represents the name of the agent sending the message
:receiver	Represents name(s) of the recipient(s) of the message.
:content	The content of the message is expressed in this section
:in-reply-to	Represents that this message is a reply referring to an earlier action
:language	Language of representing the content of the message.
:reply-with	Reply desired with an expression or a conversation thread identifying the original message.
:ontology	Ontology gives meaning to the symbols in the contents
:reply-by	Represents the deadline by which reply must be received
:protocol	Protocol being used by communicating agents.
:conversation id	Represents the identification for ongoing sequence of communicative acts.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Description of Communicative Acts

Act	Description
accept-proposal	The action of accepting a previously submitted proposal to perform an action.
agree	The action of agreeing to perform some action, possibly in the future.
cancel	The action of cancelling some previously requested action.
cfp	The action of calling for proposals to perform a given action.
confirm	The sender informs the receiver that a given proposition is true, where the receiver is known to be uncertain about the proposition.
disconfirm	The sender informs the receiver that a given proposition is false, where the receiver is known to believe, or believe it likely that, the proposition is true.
failure	The action of telling another agent that an action was attempted but the attempt failed.
inform	The sender informs the receiver that a given proposition is true.
Inform-if (macro act)	A macro action for the agent of the action to inform the recipient whether or not a proposition is true.
Inform-ref (macro act)	A macro action for sender to inform the receiver the object which corresponds to a definite descriptor (e.g. a name).
Not-understood	The sender of the act informs the receiver that it has not understood the latest action performed by receiver.
propose	The action of submitting a proposal to perform a certain action, given certain preconditions.
Query-if	The action of asking another agent whether or not a given proposition is true.
Query-ref	The action of asking another agent for the object referred to by an expression.
Refuse	The action of refusing to perform a given action, and explaining the reason for the refusal.
Reject-proposal	The action of rejecting a proposal to perform some action during a negotiation.
request	The sender requests the receiver to perform some action.
Request-when	The sender wants the receiver to perform some action when some given proposition becomes true.
Request-whenever	The sender wants the receiver to perform some action as soon as some proposition becomes true and thereafter each time the proposition becomes true again.
subscribe	The act of requesting a persistent intention to notify the sender of the value of a reference, and to notify again whenever the object identified by the reference changes.

Figure 8. Structure of a KQML Message



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Reserved Parameters in KQML

Parameter	Description
:sender	the actual sender of the performative
:receiver	the actual receiver of the performative
:from	the origin of the performative in:content when forward is used
:to	the final destination of the performative in:content when forward is used.
:in-reply-to	the expected label in a response to a previous message (same as the:reply-with value of the previous message)
:reply-with	the expected label in a response to the current message
:language	the name of the representation language of the:content
:ontology	the name of the ontology (e.g., set of term definitions) assumed in the:content parameter.
:content	The information about which the performatives expresses attitude

Ontology

The agent community has borrowed ontologies (Gruber, 1995) because they offer a global worldview which allows all agents to ground their beliefs and actions. Ontologies provide agents with a powerful domain of discourse. Before conversing, agents can agree to use a particular ontology, which defines some aspect of the world. Using such a shared ontology, agents can then discuss a topic with complete confidence that other agents have the same understanding of the items discussed. For instance, mutually exchanged 'airplane' ontology allows one agent to talk about a particular airplane e.g. Boeing 747 or Airbus320, and for the other agents to understand that Boeing 747 is a type of airplane object, which has several properties. Ontologies are playing significant role in semantic web since it aims at automation of searching and extraction of useful information spread across the web through multi-agent systems, which exploit ontologies for understanding their domain. Agents are required to have knowledge about its environment, application for which it is working and should also be aware of its neighbor agents. Communicating agents can communicate only if they share the ontology and the information is expressible using shared ontology. Figure 9 given below illustrates this concept through Use case diagram. Different agents may have similar concepts defined differently as per concerned applications. Thus, while communicating agents need to make use of concerned ontology so as to communicate properly. In Figure 9, agent1 and agent2 are using animal domain ontology which provides them knowledge about white mouse whereas agent1 and agent3 are using computer domain ontology where mouse is defined as a physical input device for computer systems.

It is actually the matter of regret that ontologies are usually expressed ambiguously and are generally embedded into agents and moreover, there is no external repository available that describes the ontology specifically i.e. semantics defining ontology are lacking. Research related to such definitions and structure of ontological databases is lagging behind. Such deficiencies further double when the communicating agents are heterogeneous and hybrid and belong to different multi-agent systems. Each MAS is possessed with its local ontology and hence the external agents do not have the idea of ontologies belonging to other MASs. In order to establish the communication link between heterogeneous MAS, ontology shall be linked using some specified standards that supports interface between different ontologies.

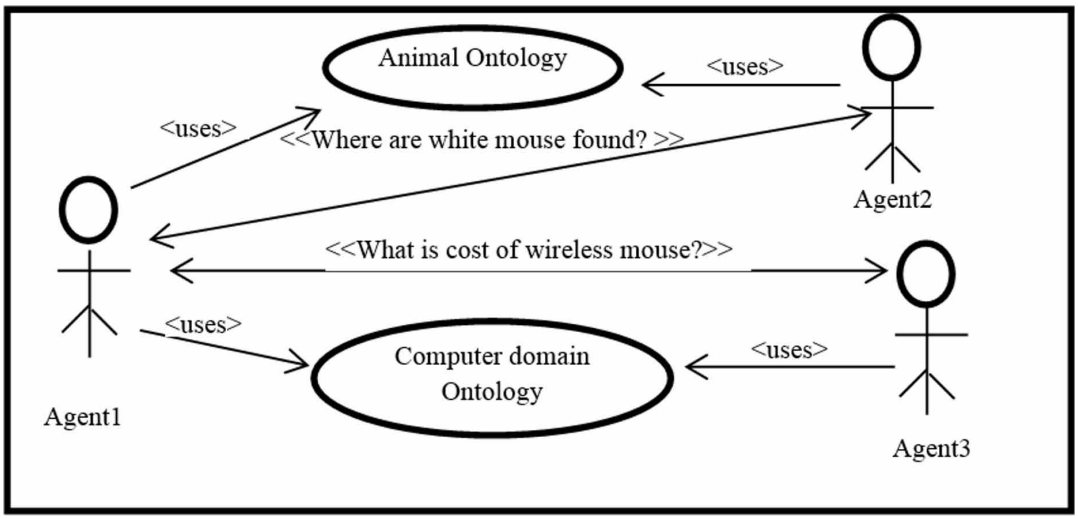
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Performatives in KQML

Sr. No.	Performative	Associated Meaning
37.	ask-if	S wants to know if the:content is in R's VKB
38.	ask-all	S wants all of R's instantiations of the:content that are true of R
39.	ask-one	S wants one of R's instantiations of the:content that is true of R
40.	stream-all	multiple-response version of ask-all
41.	eos	the end-of-stream marker to a multiple-response (stream-all)
42.	tell	the sentence is in S's VKB
43.	untell	the sentence is not in S's VKB
44.	deny	the negation of the sentence is in S's VKB
45.	insert S	asks R to add the:content to its VKB
46.	uninsert	S wants R to reverse the act of a previous insert
47.	delete-one	S wants R to remove one matching sentence from its VKB
48.	delete-all	S wants R to remove all matching sentences from its VKB
49.	undelete	S wants R to reverse the act of a previous delete
50.	achieve	S wants R to do make something true of its physical environment
51.	unachieve	S wants R to reverse the act of a previous achieve
52.	advertise	S wants R to know that S can and will process a message like the one in a:content
53.	unadvertise	S wants R to know that S cancels a previous advertise and will not process anymore messages like the one in the:content
54.	subscribe	S wants updates to R's response to a performative
55.	error	S considers R's earlier message to be malformed
56.	sorry	S understands R's message but cannot provide a more informative response
57.	standby	S wants R to announce its readiness to provide a response to the message in:content
58.	ready	S is ready to respond to a message previously received from R
59.	next	S wants R's next response to a message previously sent by S
60.	rest	S wants R's remaining responses to a messapreviously sent by S
61.	discard	S does not want R's remaining responses to a previous(multi-response) message.
62.	register	S announces to R its presence and symbolic name
63.	unregister	S wants R to reverse the act of a previous register
64.	forward	S wants R to forward the message to the:to agent (R might beThat agent)
65.	broadcast	S wants R to send a message to all agents that R knows of
66.	transport-address	S associates its symbolic name with a new transport address
67.	broker-one	S wants R to find one response to a <performative>(some agent
68.	broker-all	S wants R to find all responses to a <performative>(some agentother than R is going to provide that response)
69.	recommend-one	S wants to learn of an agent who may respond to a <performative>
70.	recommend-all	S wants to learn of all agents who may respond to a <performative>
71.	recruit-one	S wants R to get one suitable agent to respond to a <performative>
72.	recruit-all	S wants R to get all suitable agents to respond to a <performative>

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Use Case Diagram Describing Ontology



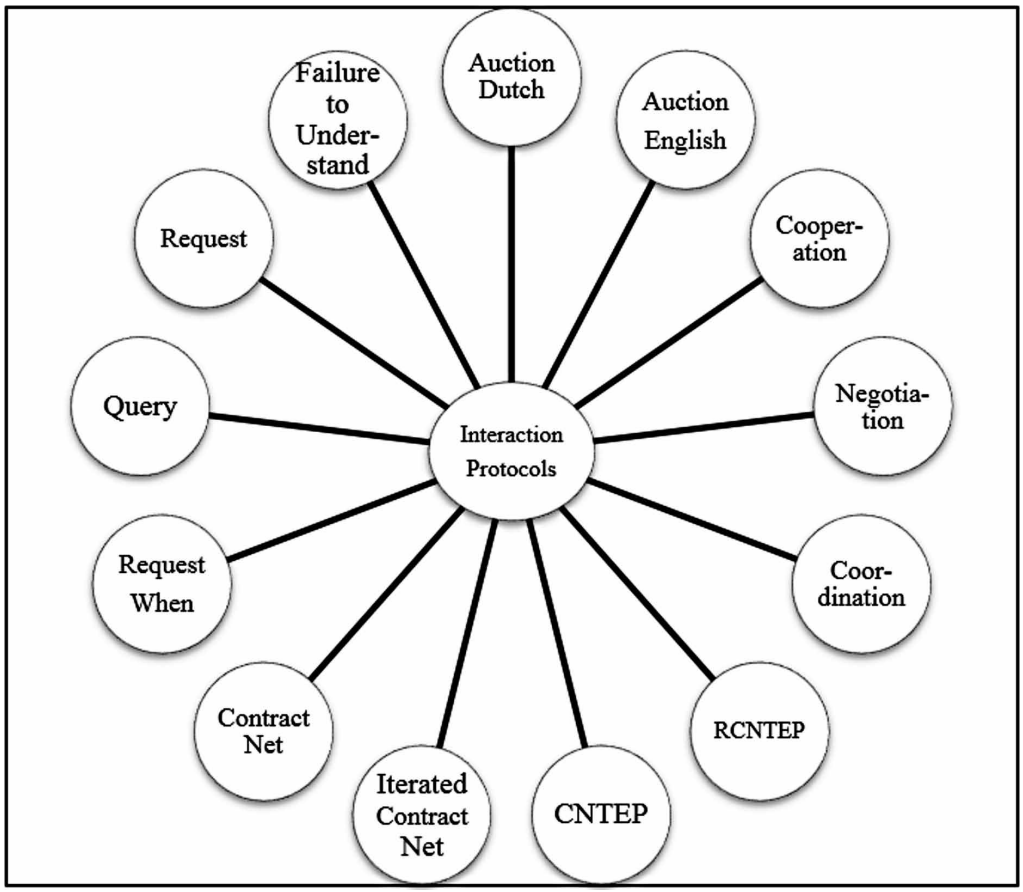
Interaction Protocols

As is already mentioned that an interaction protocol enables agents to exchange messages and establish conversation with other competing agents. Agents communicate with their peers to achieve better goals and enhance the efficiency of the system. During this communication, agents coordinate, cooperate, compete, plan and negotiate on various aspects turning into a social entity. FIPA has put in many efforts to facilitate agent interaction and thus has continually been updating interaction protocols. Few important protocols are shown in Figure 10.

Amongst the various protocols, failure to understand protocol is the simplest protocol and it assumes that all communicating agents are following the same protocol. It offers flexibility in terms of offering "not-understood" communicative act to the agent not able to either understand the message or is unable to fulfil the desired goal. The request protocol is request-response pair protocol as to every request, listening agents are required to respond positively be it that they do not understand the message. An extension of this protocol is request-when protocol which restrains the response only if the set of preconditions are found to be true i.e. if the preconditions are not satisfied, the agents may not respond to the message thus saving the network traffic. The query protocol allows sending query to the receiver and receiver is required to respond to the sender with the results. Above protocols are suitable for homogenous communicating agents but cannot be applied to heterogeneous communicating agent. Contract Net Protocol is a FIPA standardized protocol that is suitable for the latter category of agents operating in distributed MAS. An agent desiring to establish the communication initiates the communication with many agents and selects only few agents for establishing final communication link. The listening agents either complete the task and inform to initiator agent or even if they fail, failure message is communicated. It allows only onetime connection in contrast to iterated contract net protocol which permits iterative call for proposals and thus continues to execute until initiator stop issuing further calls. Both the above protocols do not evaluate the trustworthiness of agents and thus lacked on the issue. Contract Net Trust Establishment Protocol more commonly known as CNTEP overcomes the limitation of CNP by introducing trust in CNP and another extension Reliable Contract Net Trust Establishment Protocol extension (RCNTEP) is

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. Agent Interaction Protocols



also available which considers reliability of agents as an important factor before granting bids to listening agents. The underlying principle behind all above protocols is coordination, cooperation and negotiation and thus are based on the general coordination, cooperation and negotiation protocols. The protocols are executed when no single agent is capable of completing the task alone and need support from peers. In coordination protocol agents coordinate to achieve a common goal while cooperation protocol follows divide and conquer approach and divides a complex problem into sub-problems and solve each one of them cooperatively. Negotiation protocol allows agent to negotiate to reach to a common solution. In such situations agents either reach to a consensus or may raise further conflicts.

Both auction-English and auction-Dutch protocol are two variations of auction protocol and are based on auction procedure i.e. call for bids and select the most suitable bid. However, auction-English protocols begins by selecting the lowest bid while auction-Dutch protocol selects a quite higher bid and then reduces the value of bid gradually to settle to an optimum bid. For further details, readers are defered to (Alberti, Chesani, Daolio, Gavanelli, Lamma, Mello, & Torroni, 2007; Juneja, Jagga & Singh, 2015; Singh, Singh & Mukherjee, 2014).

Multiagent interaction is in existence since the inception of technology wherein the number of agents interact through various communication links. Although various agents have their own globe of authority, However, these may depend on each other in various ways. For instance, two interacting agents do

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

not depend at all on each other while in another case, one agent depends on another but vice-verca is not true. The interacting agents might get mutually dependent for achieving a goal. The next section sheds light on the type of agent interaction, in particular.

MULTIAGENT SYSTEMS: THE PRACTICAL DEVELOPMENTS

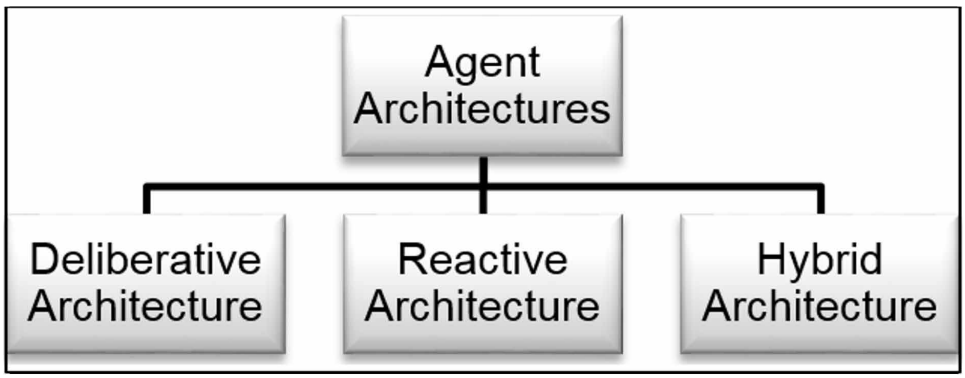
This section aims to highlight the importance of agents in practice and hence various architectures supporting the construction of agents since the inception of agent paradigm are being presented in this section. In general, architecture is known as a precise procedure for constructing autonomous agents. It usually considers decomposition of a problem into sub-problems and thereby the construction of agent is also considered as construction of a few components which are intelligent enough to pursue a goal. The ultimate aim of a practically important architecture is that agents thus developed are able to respond to a query either by input data or by changing its internal states and taking the action. According to the literature, many agent architectures supporting the design and developments of artefacts are available. However, owing to the limited scope and space available, the important architectures (see Figure 11) are being illustrated below.

Deliberative Architectures

Newel and Semon formulated the very first AI paradigm and named it as physical symbol system hypothesis (Newell & Simon, 1976; Newell, 1980). Here physical symbol system represents a physically attainable set of symbols that together form a structure and are capable of running processes as directed. The authors further called such a system as a general intelligent system. On this foundation, Geneserth (Michael & Nils, 1988) introduced deliberative agent architectures in which a deliberative agent takes the decision using logical reasoning, pattern matching and represent symbolic model of the world. In fact, it employs sense-plan-act problem solving paradigm for solving AI problems (see Figure 12).

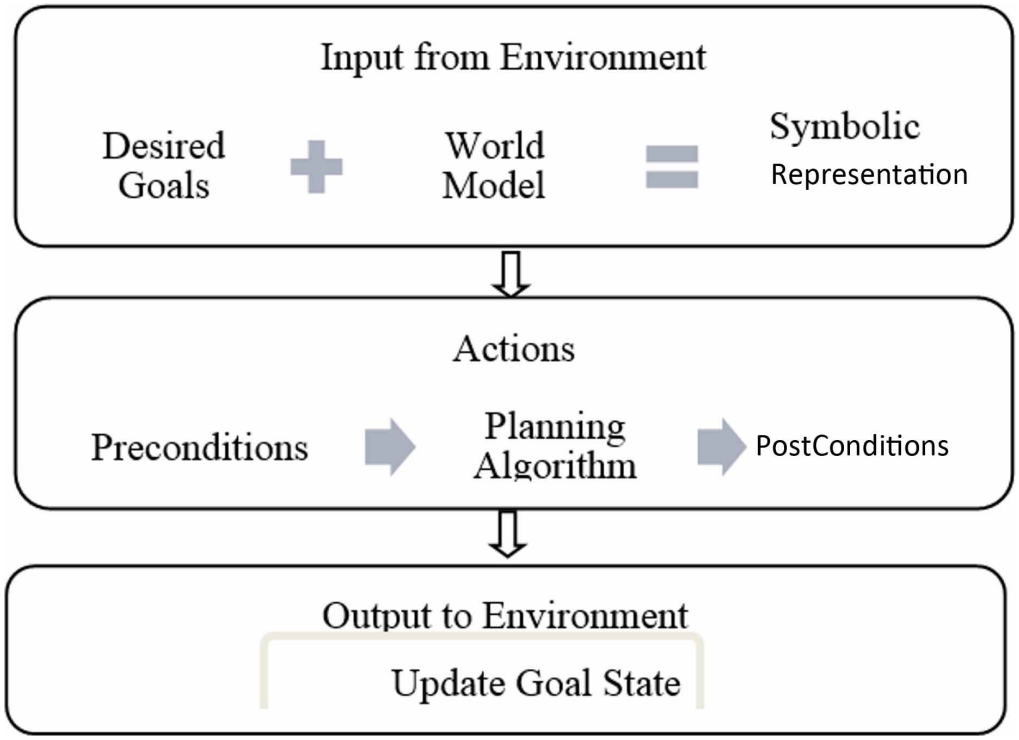
In deliberative architectures, planning agents play the important role and has been in picture since 1970. STRIPS (Fikes & Nilsson, 1971) is one of the earliest systems that has employed planning agents. In such a system, both real world and desired goals are represented symbolically and sub sequentially for

Figure 11. Agent Architectures



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 12. Deliberative Architecture



actions, preconditions pertaining to planning agent execution are imposed to ensure that desired goals are achievable. Planning algorithm involves steps to be executed to achieve the goal and later applies postconditions to guarantee the legitimacy of outcomes. Although, the planning algorithm is simple but the same is inefficient when applied to solve real world complex problems. Hence the demand to have better planning algorithms continues to evolve and lot of time in research has already been devoted, such as developments of systems that support automated planning exists, such as Integrated Planning Execution and Monitoring (IPEM) (Ingerson, Jose & Steel,1988), AutoDrive system, just to list a few.

The developments of such agents are assisted with various design challenges. For instance, representing exact and accurate symbolic description of real world is unrealistic. Moreover, limit on the amount of symbols required to represent physical world is indefinite. In any case, agents designed for an application can neither be granted nor expected to acquire through experience based learning, intelligence equivalent to human developers. Owing to the above challenges, developments in the theories pertaining to learning, knowledge representation, intelligence are still open research gates. Further, research in the direction classified planning as hierarchical planning and non-linear planning (Sacerdoti, 1974; Sacerdoti, 1975). However, (Chapman, 1987) contradicted to above classification also and proved them a failure again in case of real time systems. Results by Chapman had a strong impact on the planning AI and enforced researchers to think about alternative solutions such as reactive agent architectures.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Reactive Architectures

Looking into the unfolded challenges in founding planning architectures, many researchers raised questions on the practicality of agent based computing leading to the origin of reactive architecture (Bonasso, Firby, Gat, Kortenkamp, Miller, & Slack, 1997). In contrast to deliberative architecture, reactive architecture does not focus on symbolic representation and hence symbolic reasoning.

Significant contribution to the development of these architectures is credited to Brooks. Brooks (Brooks,1991a; Brooks,1991b) emphasized that intelligence among agents neither depend on symbolic representation nor on symbolic reasoning. He further mentioned that intelligence among agents is due to the environment in which these are situated and also its emergence is the temporal result of agent interaction with its peers in respective multiagent systems. Reactive architectures represents the world model in very simple form. However, it runs strong bonding between input and respective action. These exhibit behaviour-based paradigms where behaviour persistently matches perceptual input to action output without any reasoning. Such architectures respond as directed. The reactive behaviour is depicted as *output action*: *State→primitive action* where *State* represents the states of the environment, and *primitive action*is the set of basic actions (see Figure 13). Thermostats are the most common example of reactive architecture and actions executed are depicted through equation (1) below:

$$action\left(s\right)=\begin{Bmatrix} Heater\,off\,if\,temperature\,OK \\ Heater\,On\,Otherwise \end{Bmatrix} \tag{1}$$

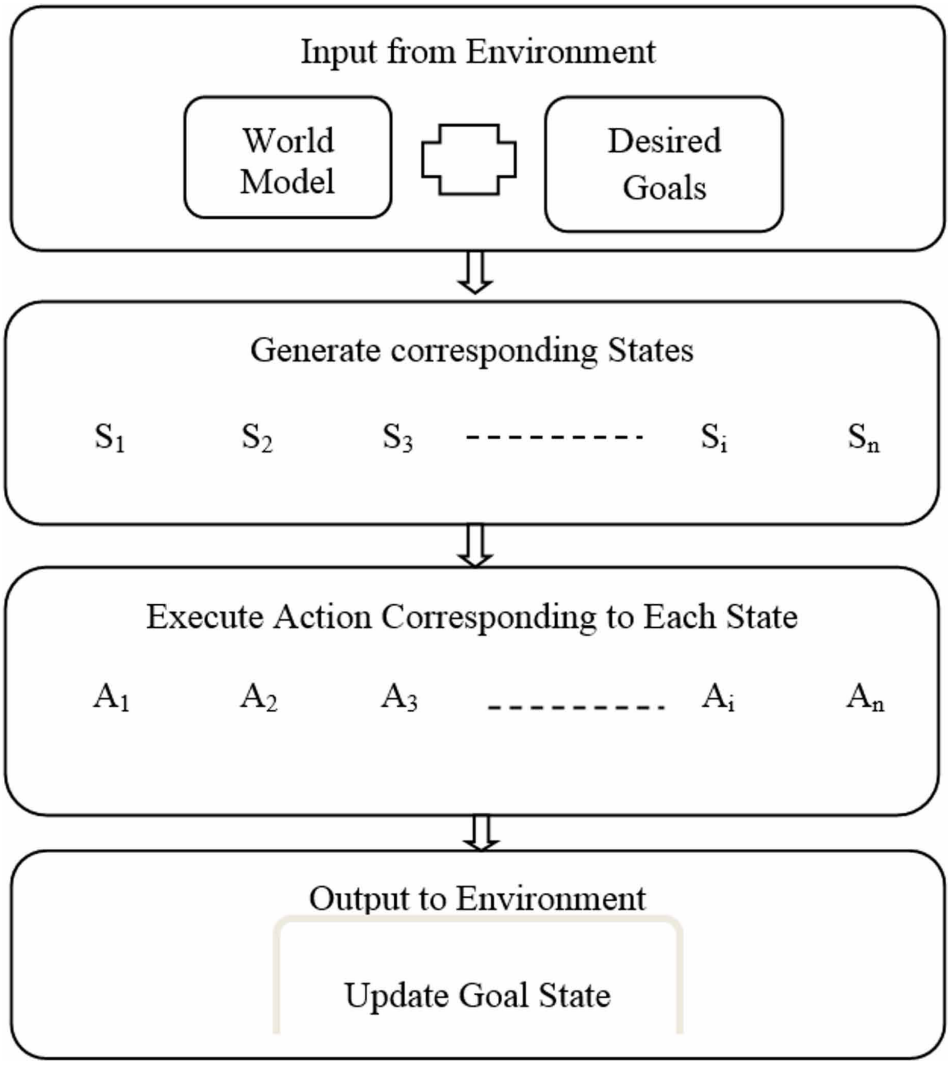
Hybrid Architectures

The research and practical solutions presented by both deliberative and reactive architectures proved that both the architectures are not the best solutions for constructing agents. On-going research efforts further suggested amalgamating both the above architectures and termed the new one as hybrid architecture (Wooldridge & Jennings, 1994). General hybrid architecture is shown in Figure 14. An agent in such architecture comprises of several multiagent systems where each system has its own plans and makes use of deliberative architecture (symbolic reasoning) for taking decisions. Few agents are designed so as to react to an event as directed making reactive architecture a reality. A hybrid architecture usually employs layered approach.

A popular category of agents known as Belief-Desire-Intention (BDI) agents (Haddadi & Sundermeyer, 1996) are also developed under the umbrella of hybrid architectures only. The BDI agents not only uphold a schema of their world but also maintain an explicit standpoint which can be altered using symbolic reasoning. The agents have their set of facts or known inputs known as agent beliefs, desires are the inputs from users or system representations while intentions are the actions that are executed to achieve the stated desires. A BDI agent encapsulates a collection a beliefs or behavior known as goals and the methods for selecting them. Each intention triggers itself when it matches the input in the form of desires either being sensed from the environment or generated due to an event. Depending upon the learning and sensing abilities of agents, new behavioral actions may be added periodically. The agent thread constantly checks the environment to ensure if the current desire is applicable. If it finds the desire to be a false application, it stops its intention by calling the respective methods. The agent on its own is accountable for initiating and dismissing itself.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

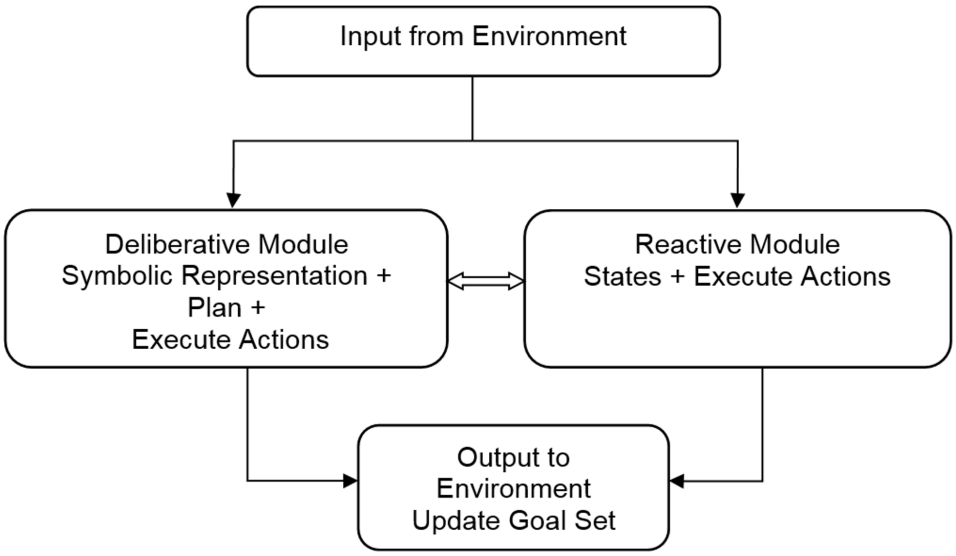
Figure 13. Reactive Architecture



The BDI-based architecture of intelligent agent primarily comprises of five components namely an Agent, an Environment, a Belief Component, Desire Component and Intention Component. The flow diagram as well as class diagram of design of intelligent agent architecture (Sharma & Juneja, 2006) is shown in Figure 15 and Figure 16. Jam-packed agent-based software requires message communication mechanism, information gathering and dissemination services, event notification and lot of signaling information. These are usually provided by agent-friendly platform such as JADE, IBM-Aglet or Aglet Software Development Kit (ASDK), Voyager, Anchor and Zeus and so on. According to the available literature (Macal & North, 2014; Nguyen, Dang, Hluchy, Balogh, Laclavik & Budinska, 2002; Singh, Juneja & Sharma, 2011) more than 100 products are available. However, at fundamental level, each framework provisions three modules for agent development namely creation, communication and discovery. Creation provides the capability to quickly build and execute agents within a supported atmosphere and communication employs speech acts to mitigate inter-agent communication. Discovery lets agents

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

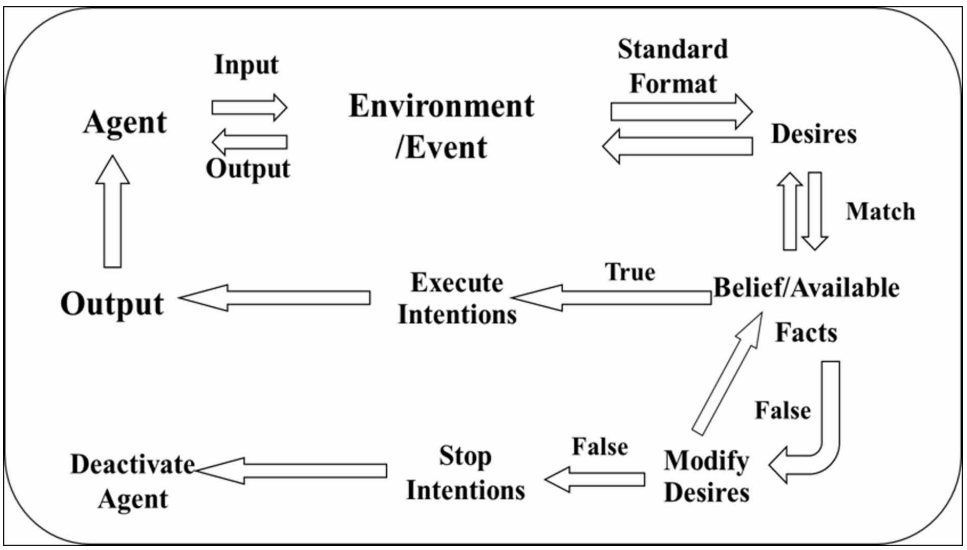
Figure 14. Hybrid Architecture



to explore undiscovered agents using a service-based discovery mechanism. Each framework though offering the above mentioned common services, also offers a unique set of additional features such as standards compliance, mobility, interoperability, knowledge-based ontologies, graphical interfaces etc. Table 6 delineates a comparison of few popular agent development frameworks.

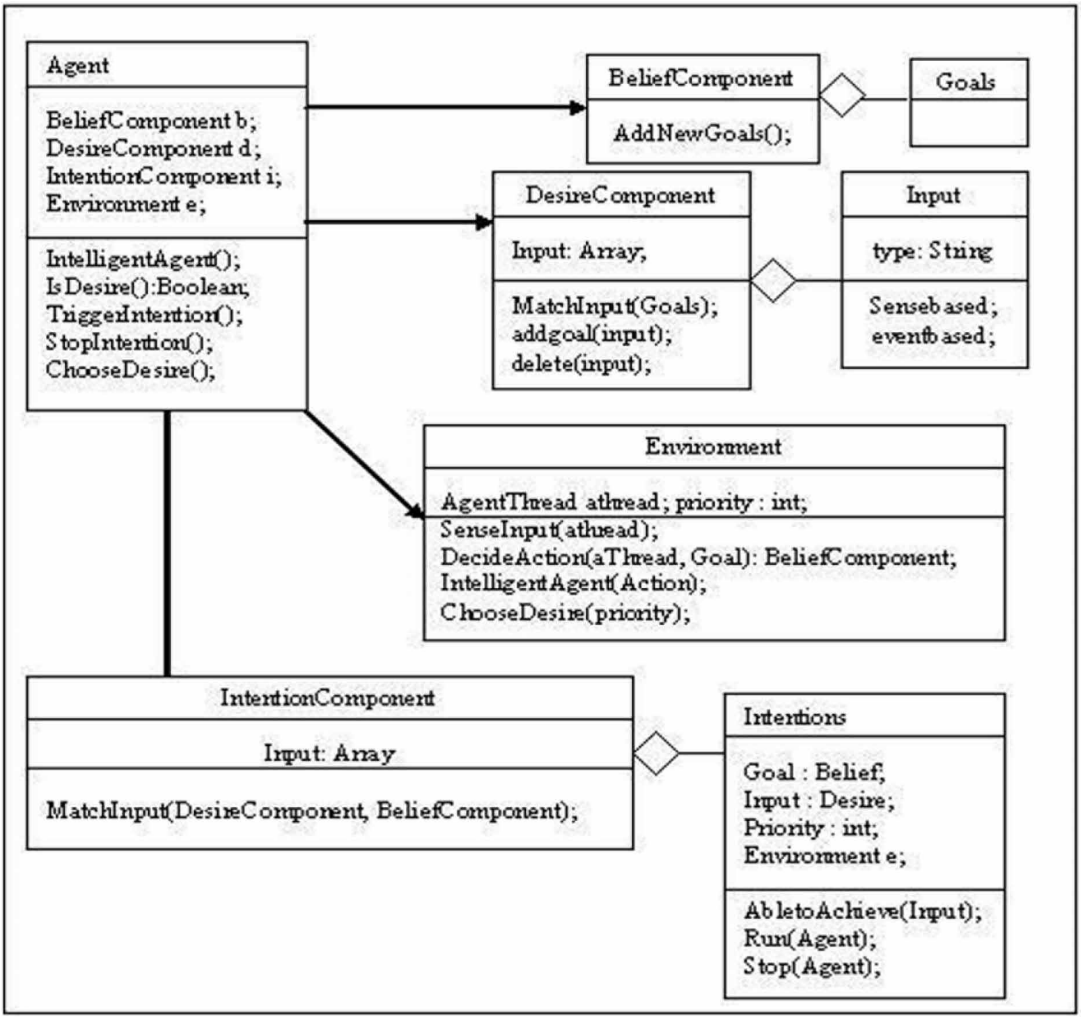
From Table 6 it is clear that most of agent development tools are open source and freely available thereby attracting developers and making intelligent agents appealing solutions for many web based applications.

Figure 15. Information Flow in BDI Architecture



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 16. The Class Diagram of BDI Based Agent (Sharma & Juneja, 2006)



AGENT-BASED SYSTEMS: ARE THESE REALLY UNIVERSAL?

It is evident from the study presented above that agent based systems are proving to be promising solutions to the development of latest applications such as process control, telecommunication, business processing, air-traffic control, medical domain and many more. The basic attributes of agents being autonomous, communal entities make them responsive enough to reflect to alterations in the environment. Owing to the strong behavioral aspects of agents, developers are able to design dependable, strong, extensible and accessible systems. The major advantages of the agent paradigm is that it can overcome delay in communication, reduce user load, execute at different times and independently, adapt with dynamism, can operate in varied environments, and usually have strong and fault-tolerant behavior. Since the design approach is modular, the agent architecture also reveals software reusability. Secure communication mechanisms exist to alleviate inter-agent, agent-to-platform and vice-versa. Once designed, agents can interchange their roles in order to fulfil the user's demand. Agents are atomic and

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. Agent Development Frameworks

Toolkit	Platform	Version	Developer	Operating System	Nature	Resource Link
JADE (Java Agent Development Framework)	Java	JADE 4.4	Telkom Italia Lab	JDK 1.1.x on Win32	Open Source	http://jade.tilab.com/
IBM-Aglet or Aglet Software Development Kit (ASDK)	Java	Aglet 2.0.2	IBM Tokyo research	JDK 1.1 x on win 32	Open Source	http://aglets.sourceforge.net/
Voyager	Java	Voyager 8.0	Object space	Embedded / Wireless OS	Commercial	http://www.recursionsw.com
Anchor	Java	Anchor v2.1.1 Patch	Lawrence Berkeley National Lab	UNIX, Linux, Windows	BSD Style License	http://escholarship.org/uc/item/2594j56c
Zeus	Java	Zeus Agent Toolkit 2.0 patch 2	British Telecommunication labs	Linux, All 32-bit windows	Open source	https://sourceforge.net/projects/zeusagent/

isolated i.e. they don't get distracted easily during processing of any task. Agents can be programmed to avoid flaws of reasoning that usually humans have. The agent concept provides a scalable, modular, flexible and a robust architecture. Also, if the goal is safe, agents are more reliable than object-oriented or component-based solutions. Agents are useful in reducing manual load and can serve as knowledge resources. Agents never get tired. Although an agent-based system can have any number of agents but security tools presents performance limitations that command design decisions and may adversely affect the advantages of agent-based systems.

Despite of the many benefits, agents cannot be applied in all situations. A multi-agent system implicitly is possessed with complicated inter-agent communication which in turn makes the behavior of agent unpredictable. Such complex systems can adversely affect real time applications which demand output in real time. However, with the technologies evolving at a fast pace making the information technology networked as well as distributed, agents have emerged as most fascinating paradigm for system and application development. It appears that more trust than required is being shown in the technology as many applications where agents are being deployed could have been developed and executed without agents as well. In fact, developers are applying agents to almost all distributed domains even if the proposed agent-based solution may not be the most optimal solution. The choice of weather an agent based solution is feasible and shall be applied or not, depends on many factors. Although, agent based solutions are offering many pros but it also has some inherent cons. For instance, MASs are usually application specific and thus are not good for all domains having different constraints especially the real-time systems. Further, in most of the applications where agents are deployed, the response depends on its native state and it is impossible that within MAS, all agents will possess the complete information about the entire system. Hence, the outcome of sub-problems may not be accurate and will finally affect the performance of the entire system. The lack of global knowledge and its adverse effect on the entire system is one of the primary issues and has always been attracting research attention. Moreover, in order to get the job done from agents, users delegate the tasks to agents which demand that these agents are trustworthy and their users have confidence in them. For agents to gain trust and confidence of users these must work

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

efficiently and further, the decision pertaining to the limit on autonomy of agents is also challenging. One can say, that an agent based systems although seems to be fascinating and promising but we must know the limitations these offer and hence cannot be trusted to be applied in every distributed system.

CONCLUDING REMARKS

The domain of multi-agent is highly effervescent and escalating at an exponential rate. The basic concepts offer powerful solutions for designing variety of agent-based solutions. Since last three decades, the advances in the domain are significant specifically pertaining to agent interaction and communication in complex distributed systems. The technology is now finding space in commercial applications and complex and mission critical applications as well. Efforts to improve the computing are still going on and many issues such as defining the limit on trust, intelligence and their applicability are yet open. Full potential of agents is yet to be realized.

REFERENCES

Alberti, M., Chesani, F., Daolio, D., Gavanelli, M., Lamma, E., Mello, P., & Torroni, P. (2007). Specification and verification of agent interaction protocols in a logic-based system. *International Journal for Parallel and Distributed Computing*, *8*(1), 1–13.

Barbati, M. (2012). The design of transportation networks: A multi objective model combining equity, efficiency and efficacy. *OASICS-OpenAccess Series in Informatics*, *22*, 75–83.

Barbati, M., Giuseppe, B., & Andrea, G. (2012). Applications of agent-based models for optimization problems: A literature review. *Expert Systems with Applications*, *39*(5), 6020–6028. doi:10.1016/j.eswa.2011.12.015

Barber, K. S., & Martin, C. E. (1999). Agent autonomy: Specification, measurement, and dynamic adjustment. *Proceedings of the autonomy control software workshop at autonomous agents* (pp. 8-15).

Bauer, B., Jörg, P. M., & Odell, J. (2001). Agent UML: Formalism for specifying multiagent software systems. *International Journal of Software Engineering and Knowledge Engineering*, *11*(3), 207–230. doi:10.1142/S0218194001000517

Brenner, W., Zarnekow, R., & Wittig, H. (2012). *Intelligent software agents: foundations and applications*. Springer Science & Business Media.

Brooks, R. A. (1991a). Intelligence without reason. *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence (IJCAI-91)*, Sydney, Australia (pp. 569-595).

Brooks, R. A. (1991b). Intelligence without representation. *Artificial Intelligence*, *47*(1-3), 139–159. doi:10.1016/0004-3702(91)90053-M

Chaib, B., & Dignum, F. (2002). Trends in agent communication language. *Computational Intelligence*, *2*(5), 89–101. doi:10.1111/1467-8640.00184

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapman, D. (1987). Planning for conjunctive goals. *Artificial Intelligence*, *32*(3), 333–377. doi:10.1016/0004-3702(87)90092-0

Chen, B., & Cheng, H. H. (2010). A review of the applications of agent technology in traffic and transportation systems. *IEEE Transactions on Intelligent Transportation Systems*, *11*(2), 485–497. doi:10.1109/TITS.2010.2048313

Cioppa, T. M., Lucas, T. W., & Sanchez, S. M. (2004). Military applications of agent-based simulations. *Proceedings of the 2004 Winter Simulation Conference*. IEEE. 10.1109/WSC.2004.1371314

Dam, H.K., Pitt, J., Governatori, G., Ito, T., & Xu, Y. (2016). Introduction to the Special Issue on Principles and Practices in Multi-Agent Systems. *Scalable Computing: Practice and Experience, 16*(4).

Etzioni, O., & Weld, D. (1994). A softbot-based interface to the internet. *Communications of the ACM*, *37*(7), 72–76. doi:10.1145/176789.176797

Ferber, J. (1999). *Multi-agent systems: an introduction to distributed artificial intelligence* (Vol. 1). Addison-Wesley.

Fiedrich, F., & Burghardt, P. (2007). Agent-based systems for disaster management. *Communications of the ACM*, *50*(3), 41–42. doi:10.1145/1226736.1226763

Fikes, R. E., & Nilsson, N. J. (1971). STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, *2*(3-4), 189–208. doi:10.1016/0004-3702(71)90010-5

Finin, T., Fritzson, R., McKay, D., & McEntire, R. (1994, November). KQML as an agent communication language. *Proceedings of the third international conference on Information and knowledge management* (pp. 456-463). ACM.

Franklin, S., & Graesser, A. (1996). Is it an Agent, or just a Program?: A Taxonomy for Autonomous Agents. In *Intelligent agents III agent theories, architectures, and languages* (pp. 21–35). Springer Berlin Heidelberg.

Genesereth, M. R., & Fikes, R. E. (1992). Knowledge interchange format-version 3.0: reference manual.

Genesereth, M. R., & Ketchpel, S. P. (1994). Software agents. *Communications of the ACM*, *37*(7), 48–53, ff. doi:10.1145/176789.176794

Georgeff, M., Pell, B., Pollack, M., Tambe, M., & Wooldridge, M. (1998). The belief-desire-intention model of agency. In Intelligent Agents V: Agents Theories, Architectures, and Languages (pp. 1-10). Springer Berlin Heidelberg.

Graesser, A. C., Jeon, M., & Dufty, D. (2008). Agent technologies designed to facilitate interactive knowledge construction. *Discourse Processes*, *45*(4-5), 298–322. doi:10.1080/01638530802145395

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies*, *43*(5), 907–928. doi:10.1006/ijhc.1995.1081

Haddadi, A., &Sundermeyer, K. (1996). Belief-desire-intention agent architectures. In *Foundations of distributed artificial intelligence* (pp. 169-185).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Heger, A. S., & Koen, B. V. (1991). KNOWBOT: An adaptive data base interface. *Nuclear Science and Engineering*, *107*(2), 142–157.

Heppenstall, A. J., Crooks, A. T., See, L. M., & Batty, M. (Eds.). (2011). *Agent-based models of geographical systems*. Springer Science & Business Media.

Hewitt, C., Bishop, P., & Steiger, R. (1973, August). A universal modular actor formalism for artificial intelligence. *Proceedings of the 3rd international joint conference on Artificial intelligence* (pp. 235-245). Morgan Kaufmann Publishers Inc.

van der Hoek, W., & Wooldridge, M. (2008). Multi-agent systems. In Handbook of Knowledge Representation (pp. 887-928).

Hussein, K.M.E.D. & Pe, F. (2006). U.S. Patent No. 7,007,235. Washington, DC: U.S. Patent and Trademark Office.

Ingerson, A., Jose, A., & Steel, S. (1988). Integrating Planning, Execution and Monitoring. Proceedings of AAAI '88.

Jennings, N. R. (2000). On agent-based software engineering. *Artificial Intelligence*, *117*(2), 277–296. doi:10.1016/S0004-3702(99)00107-1

Jennings, N. R., Sycara, K., & Wooldridge, M. (1998). A roadmap of agent research and development. *Autonomous Agents and Multi-Agent Systems*, *1*(1), 7–38. doi:10.1023/A:1010090405266

Juneja, D., Jagga, A., & Singh, A. (2015). A review of FIPA standardized agent communication language and interaction protocols. *Journal of Network Communications and Emerging Technologies*, *5*(2), 179–191.

Kone, M. T., Shimazu, A., & Nakajima, T. (2000). The state of the art in agent communication languages. *Knowledge and Information Systems*, *2*(3), 259–284. doi:10.1007/PL00013712

Labrou, Y., & Finin, T. (1997). *A Proposal for a new KQML Specification* (Technical Report TR-CS-97-03). University of Maryland Baltimore County.

Labrou, Y., Finin, T., & Peng, Y. (1999). Agent communication languages: The current landscape. *IEEE Intelligent Systems*, *2*(2), 45–52. doi:10.1109/5254.757631

Laghari, S., & Niazi, M. A. (2016). Modeling the Internet of Things, Self-Organizing and Other Complex Adaptive Communication Networks: A Cognitive Agent-Based Computing Approach. *PLoS ONE*, *11*(1), e0146760. doi:10.1371/journal.pone.0146760 PMID:26812235

Lange, D. B., & Oshima, M. (1999). Seven good reasons for mobile agents. *Communications of the ACM*, *42*(3), 88–89. doi:10.1145/295685.298136

Macal, C., & North, M. (2014, December). Introductory tutorial: Agent-based modeling and simulation. *Proceedings of the 2014 Winter Simulation Conference* (pp. 6-20). IEEE Press. 10.1109/WSC.2014.7019874

Maes, P. (1994). Agents that reduce work and information overload. *Communications of the ACM*, *37*(7), 30–40. doi:10.1145/176789.176792

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Maes, P. (1995). Artificial life meets entertainment: Lifelike autonomous agents. *Communications of the ACM*, *38*(11), 108–114. doi:10.1145/219717.219808

Magedanz, T., Rothermel, K., & Krause, S. (1996, March). Intelligent agents: An emerging technology for next generation telecommunications. *Proceedings of the IEEE Fifteenth Annual Joint Conference of the IEEE Computer Societies on Networking the Next Generation INFOCOM'96* (Vol. *2*, pp. 464–472).

Manvi, S. S., & Venkataram, P. (2004). Applications of agent technology in communications: A review. *Computer Communications*, *27*(15), 1493–1508. doi:10.1016/j.comcom.2004.05.011

Mařík, V., & Lažanský, J. (2007). Industrial applications of agent technologies. *Control Engineering Practice*, *15*(11), 1364–1380. doi:10.1016/j.conengprac.2006.10.001

McArthur, S. D., Davidson, E. M., Catterson, V. M., Dimeas, A. L., Hatziargyriou, N. D., Ponci, F., & Funabashi, T. (2007). Multi-agent systems for power engineering applications—Part II: Technologies, standards, and tools for building multi-agent systems. *IEEE Transactions on* Power Systems, *22*(4), 1753–1759.

Metzger, M., & Polakow, G. (2011). A survey on applications of agent technology in industrial process control. *IEEE Transactions on* Industrial Informatics, *7*(4), 570–581.

Michael, R. G., & Nils, J. N. (1988). *Logical Foundations of Artificial Intelligence. Morgan Kaufmann Publishers.*

Milojicic, D. (2000). Guest editors introduction: Agent systems and applications. *IEEE Concurrency*, *8*(2), 22–23. doi:10.1109/MCC.2000.846190

Nangle, B., Cunningham, P., & Evans, R. (1998). PALS: Personal Assistants which Learn for intelligent call Screening.

Newell, A. (1980). Physical symbol systems*. *Cognitive Science*, *4*(2), 135–183. doi:10.120715516709cog0402_2

Newell, A., & Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the ACM*, *19*(3), 113–126. doi:10.1145/360018.360022

Nguyen, G., Dang, T. T., Hluchy, L., Balogh, Z., Laclavik, M., & Budinska, I. (2002). *Agent platform evaluation and comparison. Rapport technique*. Bratislava, Slovakia: Institute of Informatics.

Nwana, H. S. (1996). Software agents: An overview. *The Knowledge Engineering Review*, *11*(03), 205–244. doi:10.1017/S026988890000789X

Nwana, H. S., & Wooldridge, M. (1996). Software agent technologies. *BT Technology Journal*, *14*(4).

Panait, L., & Luke, S. (2005). Cooperative multi-agent learning: The state of the art. *Autonomous Agents and Multi-Agent Systems*, *11*(3), 387–434. doi:10.100710458-005-2631-2

Pěchouček, M., & Mařík, V. (2008). Industrial deployment of multi-agent technologies: Review and selected case studies. *Autonomous Agents and Multi-Agent Systems*, *17*(3), 397–431. doi:10.100710458-008-9050-0

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Peter Bonasso, R., James Firby, R., Gat, E., Kortenkamp, D., Miller, D. P., & Slack, M. G. (1997). Experiences with anarchitecture for intelligent, reactive agents. *Journal of Experimental & Theoretical Artificial Intelligence*, *9*(2-3), 237–256. doi:10.1080/095281397147103

Rao, A. S., & Georgeff, M. P. (1995, June). *BDI agents: From theory to practice* (Vol. 95, pp. 312–319). ICMAS.

Russell, S., Norvig, P., & Intelligence, A. (1995). A modern approach to Artificial Intelligence. *Englewood Cliffs:* Prentice-Hall.

Sacerdoti, E. D. (1974). Planning in a hierarchy of abstraction spaces. *Artificial Intelligence*, *5*(2), 115–135. doi:10.1016/0004-3702(74)90026-5

Sacerdoti, E. D. (1975). *The nonlinear nature of plans (No. SRI-TN-101)*. CA: Stanford Research Inst Menlo Park.

Shakshuki, E., Ghenniwa, H., & Kamel, M. (2003). Agent-based system architecture for dynamic and open environments. *International Journal of Information Technology & Decision Making*, *2*(01), 105–133. doi:10.1142/S0219622003000537

Sharma, A. K., & Juneja, D. (2006). *BDI architecture of intelligent agents. Book on Technology Management* (pp. 300–309). Punjab, India: GJIMT.

Shen, W., Hao, Q., Yoon, H. J., & Norrie, D. H. (2006). Applications of agent-based systems in intelligent manufacturing: An updated review. *Advanced Engineering Informatics*, *20*(4), 415–431. doi:10.1016/j.aei.2006.05.004

Sim, K. M. (2012). Agent-based cloud computing. *Services Computing*. *IEEE Transactions on Service Computing*, *5*(4), 564–577. doi:10.1109/TSC.2011.52

Singh, A., Juneja, D., & Sharma, A. K. (2011). Agent development toolkits. arXiv preprint arXiv:1111.5930

Singh, R., Singh, A., & Mukherjee, S. (2014). A Critical Investigation of Agent Interaction Protocols in Multiagent Systems. *International journal of Advancements in Technology*, 5(2), 72-81.

Singh, Y., Gulati, K., & Niranjan, S. (2012). Dimensions and issues of mobile agent technology. arXiv preprint arXiv:1210.4644

Smith, D. C., Cypher, A., & Spohrer, J. (1994). KidSim: Programming agents without a programming language. *Communications of the ACM*, *37*(7), 54–67. doi:10.1145/176789.176795

Suzanne Barber, K., Goel, A., & Martin, C. E. (2000). Dynamic adaptive autonomy in multi-agent systems. *Journal of Experimental & Theoretical Artificial Intelligence*, *12*(2), 129–147. doi:10.1080/095281300409793

Sycara, K., Pannu, A., Williamson, M., Zeng, D., & Decker, K. (1996). Distributed intelligent agents. *IEEE Intelligent Systems*, (6), 36–46.

Sycara, K. P. (1998). Multiagent systems. *AI Magazine*, *19*(2), 79.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Van Diggelen, J., Beun, R. J., Dignum, F., Van Eijk, R. M., & Meyer, J. J. (2006, May). ANEMONE: an effective minimal ontology negotiation environment. *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems* (pp. 899-906). ACM. 10.1145/1160633.1160794

Vigna, G. (2004). Mobile agents: Ten reasons for failure. *Proceedings of the 2004 IEEE International Conference on Mobile Data Management* (pp. 298-299). IEEE.

Weiss, G. (1999). *Multiagent systems: a modern approach to distributed artificial intelligence*. MIT press.

Wooldridge, M., & Jennings, N. R. (1994). Agent theories, architectures, and languages: a survey. In *Intelligent agents* (pp. 1–39). Springer Berlin Heidelberg.

Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: Theory and practice. *The Knowledge Engineering Review*, *10*(2), 115–152. doi:10.1017/S0269888900008122

Wooldridge, M. J. (1998). *Agent technology: foundations, applications, and markets*. Springer Science & Business Media.

ENDNOTES

1 http://agents.media.mit.edu/

2 http://www.fipa.org/repository/aclspecs.html

This research was previously published in the International Journal of Ambient Computing and Intelligence (IJACI), 8(1); edited by Nilanjan Dey; pages 23-49; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 7
A Hybrid Intelligent Risk Identification Model for Configuration Management in Aerospace Systems

Jose Nava
Pinnacle Aerospace, Inc., Mexico

Alejandro Osorio
Pinnacle Aerospace, Inc., Mexico

ABSTRACT

This chapter proposes a multi-dimensional patterns recognition model for Configuration Management's Risk Identification in Aerospace Safety Critical Systems. This work has been designed for Aerospace software systems where companies require full compliance with the Aerospace Standard DO-178b. The solution focuses on Risk Identification for the Configuration Management Process Area. An Anomaly Detection Solution has been designed through the modeling of statistics and artificial intelligence algorithms, following CRISP-DM model standard for data mining solutions. A dimensional architecture was designed to model the problem through three dependent and interconnected dimensions. The first dimension, Behavioral Biometrics, which this model has extended to Human Behavioral Patterns. The second dimension is Infrastructure, which represents all physical specialized equipment, environments, networking, and its configurations. The third dimension is space-time, which in this model represents a time dimension against all geographical information project related (code, files, among others).

INTRODUCTION

Software products are deployed across a wide range of global industries, but when it comes to Aerospace safety critical systems; security and criticality are strictly rigorous. In these kinds of systems, a software failure or malfunction could result in loss of life and significant property damage.

DOI: 10.4018/978-1-7998-0951-7.ch007

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Aviation systems development has been evolving for almost a hundred years now. One of the more notorious development evolving areas is the instruments and displays that were electro-mechanical in the past and now are composed by electronic parts, which make them more reliable, secure and lighter. The biggest priority of this evolution has always been to avoid aircraft accidents and providing security in the air. The rapid increase in the use of software in airborne systems and equipment used on aircraft and engines in the early 1980s resulted in a need for industry-accepted guidance for satisfying airworthiness requirements. DO-178, "Software Considerations in Airborne Systems and Equipment Certification", was written to satisfy this need.

Developing reliable software products becomes more and more challenging for Safety Critical Industries due to specific regulations, short production cycles or high complexity design.

One of the more challenging aspects when working with Aerospace Software is maintaining a good Configuration Management control. A well planned strategy for this area can result in the fault tolerant system we have designed since the beginning.

Configuration Management has always been a complex, dynamic and constantly changing area, which not only interacts with the project team members and the IT department of the service provider, but also with those entities from the client side. This configuration makes it more complex because each company has its own Configuration Management practices, tools, and processes. Other important risk factor is that no matter the technical level of the involved teams, each employee works differently when it comes to practices like code design, SVN activities like commits, check in/out, updates, merges, syncs, cleanups, or other activities like estimation, activities tracking culture, among others.

The Federal Aviation Administration and the Joint Aviation Authorities have defined a list of Failures Condition Categories within DO-178b, as well as a list of Software Level Definitions based upon the contribution of software to potential failure conditions.

The Software level implies that the level of effort required showing compliance with certification requirements varies with the failure condition category (Radio Technical Commission for Aeronautics, 1992). These software level definitions are described below:

- **Level A (Catastrophic):** Software whose anomalous behavior would cause or contribute to a failure of system function resulting in a catastrophic failure condition for the aircraft, and multiple deaths. An example of this anomalous behavior can be the Turbine Engine Monitoring Software, which malfunction can cause a fatal accident.
- **Level B (Hazardous/Severe-Major):** Software whose anomalous behavior would cause or contribute to a failure of system function resulting in a hazardous/severe-major failure condition for the aircraft, and some deaths or severe injuries. An example of this anomalous behavior can be the Landing Gear Control Software, which malfunction can cause a big accident.
- **Level C (Major):** Software whose anomalous behavior would cause or contribute to a failure of system function resulting in a major failure condition for the aircraft, and severe injuries. An example of this anomalous behavior can be the FMS/NAV/COM System, which malfunction can cause damages and accidents.
- **Level D (Minor):** Software whose anomalous behavior would cause or contribute to a failure of system function resulting in a minor failure condition for the aircraft, and minor injuries. An example of this anomalous behavior can be the A/C monitoring system software, which malfunction can cause a failure condition.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- **Level E (No Effect):** Software whose anomalous behavior would cause or contribute to a failure of system function with no effect on aircraft operational capability or pilot workload. An example of this anomalous behavior can be the Entertainment audio system software, which malfunction can cause just a minimal failure condition.

We will describe the software level definitions that we are covering in this work, and the reasons to cover these levels.

The general objective of designing this model is to avoid Configuration Management issues during Aerospace software development and verification projects through artificial intelligence and statistical techniques like anomaly detection and pattern recognition in exploratory data analysis.

The specific objectives of this model are:

- Based on the most threatening issues in Configuration Management, to identify user behaviors through the project life cycle.
- To have a real-time intelligent pattern recognition algorithms tool based on these behaviors to detect anomalies.
- To have a real-time risks notification and diagnosis tool.
- Based on the previous intelligent pattern recognition algorithms, to have valuable statistical information, which cannot be detected or accessed through traditional mathematical methods.
- To use this statistical information to improve real-time decision making, and have informational tools to build performance improvement plans in real-time without having an impact in the project.

BACKGROUND

Over time, Aerospace companies have always put special attention at the moment of creating the Configuration Management strategy for each project at the initial phases. In this sense, special controls, monitoring strategies, resources and processes are put in place to overcome the big risks that the development and testing teams are phasing in a daily basis. These controls are not enough because there are several factors that can trigger a failure in a system; some examples are listed below:

- A failure in a build
- Bad version control
- Poor requirements traceability, bad hooks strategy
- Design complexity
- Disaster recovery issues
- Change control confusions
- Confusing environments' configuration
- Bad release management or deployment management
- High estimations' deviations
- Poor activities tracking
- High overtime rates

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The Software Configuration Management Process, in cooperation with other processes during the software life cycle process assists in satisfying general objectives to:

1. Provide a defined and controlled configuration of the software throughout the software life cycle.
2. Provide the ability to consistently replicate the executable Object Code for Software Manufacture or to regenerate it in case of a need for investigation or modification.
3. Provide control of process inputs and outputs during the software life cycle that ensure consistency and repeatability of process activities.
4. Provide a known point of review, assessing status, and change control by control of configuration items and the establishment of baselines.
5. Provide controls that ensure problems receive attention and changes recorded, approved, and implemented.
6. Provide evidence of approval of the software by control of the outputs of the software life cycle processes.
7. Aid the assessment of the software product compliance with requirements.
8. Ensure that secure physical archiving, recovery and control are maintained for the configuration items.

The daily execution of each of the hundreds of configuration management processes and good practices each engineer has to perform for these complex systems can lead to a minimal to moderate deviation in the execution's exactness. For this reason, we can say that the avionics process could be perfect, and the way a team of several engineers execute within a project could not be exactly the same in detail. In aerospace systems, a minimal deviation is a risk and a possible failure. For creating this model, we did an extensive research in Artificial Intelligence techniques to find a way to use patterns recognition to derive with behavioral biometrics mathematical solutions which can perform as a live system focused in each micro-technical movement of the engineers within a project, to detect these deviations and avoid issues within the configuration management process.

Below we provide detailed background information in Safety Critical Systems, Configuration Management in DO-178, Artificial Intelligence, Data Mining and Anomaly Detection techniques.

AEROSPACE SAFETY CRITICAL SYSTEMS

Safety Critical Systems have been evolving exponentially during the last 50 years. Technology advances are helping to have better processes and better controls over these kinds of Systems.

The ultimate objective of any system is the performance of any function. This function is frequently called the mission. The term often used to describe the overall capability of a system to accomplish its mission is system effectiveness. For consumer products, system effectiveness is related to customer satisfaction, which is related to the overall concept of quality. Quality of products and services is defined and evaluated by the customers and users of the product. System effectiveness is defined as the probability that the system can successfully meet an operational demand within a given time when operating under specified conditions. (Salvendy, 2001).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Safety-critical systems are introduced into a system to prevent, or stop the development of an accident. Safety-critical functions carried out by these systems may range from simple physical barriers to very complex systems including a lot of software. The safety-critical functions can be of technical nature, but the interaction with administrative controls and human actions may need to be included to fully understand how the performance is affected. Several international standards have been introduced related to design and evaluation of safety-critical functions. Among these standards are the IEC61508 standard, and its application specific versions/related standards, such as IEC 61511 for the process industry, IEC 62061 for machinery (ISO 13849), ISO 26262 for automotives, EN 50126/EN 50129/ IEC 50128 for railway (IEC versions are IEC 62278, IEC 62425, and IEC 62269), and IEC 61513 for nuclear.

Some of the more important concepts we should understand to have an accurate idea of what Safety Critical Systems mean for us are: Safety is freedom from accidents or losses. Safety is not reliability.

Reliability is one of the major attributes determining system effectiveness. It is generally defined as the probability that a system will perform its intended function satisfactorily, for its intended life, under specified operative conditions (Salvendy, 2001).

Safety is not security, is protection or defense against attacks, interference, or espionage. Accident is a loss of some kind, such as injury, death, or equipment damage. Risk is a combination of the likelihood of an accident and its severity:

$$risk = p(a) * s(a)$$

Hazard is a set of conditions and/or events that leads to an accident. A failure is the nonperformance or inability of a system or component to perform its intended function for a specified time under specified environmental conditions. Failures are events. e.g.. a component failure. An error is a systematic fault, which is a design error.

Errors are states or conditions, e.g. a software bug. A fault is either a failure or an error.

Safety must be considered in the context of the system, not the component or the software. It is less expensive and far more effective to build in safety early than try to tack it on later. The Hazard Analysis ties together hazards, faults, and safety measures.

Some of the most common types of Hazards are described below:

- **Actions**
  - inappropriate system actions taken
  - appropriate system actions not taken
- **Timing**
  - too soon
  - too late
- **Sequence**
  - skipping actions
  - actions out of order
- **Amount**
  - too much
  - too little

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

System safety is a planned, disciplined and systematic approach to preventing or reducing accidents throughout the life cycle of a system.

As an example of a Safety Critical System we can mention a common instrument system which is a configurable color display that provides backup information for attitude, altitude, airspeed, heading, and navigation all plains have installed in the cockpit. This instrumentation can have the capability of interfacing with several types of sensors and protocols that can be installed in different airplane models.

CONFIGURATION MANAGEMENT IN DO-178

Having a good Configuration Management control is critical in any Industry; a company's quality level can be measured by its configuration management process and control. This is a very complex area because it involves different departments as Development, Testing, IT, and Quality. Maintaining a perfect log database of a picture of every second of the life of the project in terms of configurations it is always very complex and challenging.

The objectives for Software Configuration Management are independent of software level. However, two categories of software life cycle data may exist based on the Software Configuration Management controls applied to the data.

The Software Configuration Management process includes the activities of configuration identification, change control, baseline establishment, and archiving of the software product, including the related software life cycle data. The following guidelines define the objectives for each Software Configuration Management process activity. The Software Configuration Management process does not stop when the software product is accepted by the certification authority, but continues throughout the service life of the airborne system or equipment.

The objective of the Configuration Identification activity is to label unambiguously each configuration item (and its successive versions) so that a basis is established for the control and reference of configuration items. Guidance includes:

1. Configuration identification should be established for the software life cycle data.
2. Configuration identification should be established for each configuration item, for each separately controlled component of a configuration item, and for combinations of configuration items that comprise a software product.
3. Configuration items should be configuration-identified prior to the implementation of change control and traceability data recording.
4. A configuration item should be configuration-identified before being used by other software life cycle processes, referenced by other software life cycle data, or used for software manufacture or software loading.

Baselines and Traceability

It is important to state that at the beginning of a project, we need to establish baselines when planning. The objective of baseline establishment is to define a basis for further software life cycle process activity and allow reference to. Control of configurations and traceability in a project is complex. Guidance includes:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. Baselines should be established for configuration items used for certification credit. (Intermediate baselines may be established to aid in controlling software life cycle process activities).
2. A software product baseline should be established for software product and define in the software configuration index.
3. Baselines should be established in controlled software libraries (physical, electronic, or other) to ensure integrity. Once a baseline is established, it should be protected from change.
4. Change control activities should be followed to develop a derivative baseline from an established baseline.
5. A baseline should be traceable to the baseline from which was derived, if certifications credit is sought for software life cycle process activities or data associated with the development of the previous baseline.
6. A configuration item should be traceable to the configuration item from which it was derived, if certification credit is sought for software life cycle process activities or data associated with the development of the previous configuration item.
7. A baseline of configuration item should be traceable either to the output it identifies or to the process with which it is associated.

Problem Reporting, Tracking and Corrective Action

The objective of problem reporting, tracking and corrective action is to record process non-compliance with software plans and standards, to record deficiencies of outputs of software life cycle processes, to record anomalous behavior of software products, and to ensure resolution of these problems.

Guidance includes:

1. A problem report should be prepared that describes the process non-compliance with plans, output deficiency, or software anomalous behavior, and the corrective action taken.
2. Problem detection should provide for configuration identification of affected configuration items or definition of affected process activities, status reporting of problem reports, and approval and closure of problem reports.
3. Problem reports that require corrective action of the software product or outputs of software life cycle process should invoke the change control activity.

Change Control

The objective of the change control activity is to provide control for recording, evaluation, resolution and approval of changes throughout the software life cycle.

Guidance includes:

1. Change control should preserve the integrity of the configuration items and baselines by providing protection against their change.
2. Change control should ensure that any change to a configuration item requires a change to its configuration identification.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

3. Changes to baselines and to configuration items under change control should be recorded, approved, and tracked. Problem reporting is related to change control, since resolution of a reported problem may result in changes to configuration items or baselines.
4. Software changes should be traced to their origin and the software life cycle processes repeated from the point at which the change affects their inputs. For example, an error discovered at hardware/software integration, that is shown to result from an incorrect design, should result in design correction and repetition of the associated integral process activities.
5. Throughout the change activity, software life cycle data affected by the change should be updated and records should be maintained for the change control activity.

ARTIFICIAL INTELLIGENCE

The term Artificial Intelligence was conceived in 1956 used by researches in universities within different scientific areas. Artificial Intelligence has acquired several definitions along the years. One of the most known is "A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes" (Schalkoff R., 1990). We defined Artificial Intelligence as the science that tries to emulate evolution, and within evolution, natural selection and cognitive modeling.

Artificial Intelligence has several branches, the most known are shown below:

- **Logical AI:** What a program knows about the world in general the facts of the specific situation in which it must act, and its goals are all represented by sentences of some mathematical logical language. The program decides what to do by inferring that certain actions are appropriate for achieving its goals (McCarthy, 2007).
- **Search:** AI programs often examine large numbers of possibilities, e.g. moves in a chess game or inferences by a theorem proving program. Discoveries are continually made about how to do this more efficiently in various domains.
- **Pattern Recognition:** When a program makes observations of some kind, it is often programmed to compare what it sees with a pattern. For example, a vision program may try to match a pattern of eyes and a nose in a scene in order to recognize a face. More complex patterns, e.g. in a natural language text, in a chess position, or in the history of some event are also studied. These more complex patterns require quite different methods than do the simple patterns that have been studied the most.
- **Representation:** Facts about the world have to be represented in some way. Usually languages of mathematical logic are used.
- **Inference:** From some facts, others can be inferred. Mathematical logical deduction is adequate for some purposes, but new methods of non-monotonic inference have been added to logic since the 1970s. The simplest kind of non-monotonic reasoning is default reasoning in which a conclusion is to be inferred by default, but the conclusion can be withdrawn if there is evidence to the contrary. For example, when we hear of a bird, we can infer that it can fly, but this conclusion can be reversed when we hear that it is a penguin. It is the possibility that a conclusion may have to be withdrawn that constitutes the non-monotonic character of the reasoning. Ordinary logical reasoning is monotonic in that the set of conclusions that can be drawn from a set of premises is a monotonic increasing function of the premises. Circumscription is another form of non-monotonic reasoning.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- **Common Sense Knowledge and Reasoning:** This is the area in which AI is farthest from human-level, in spite of the fact that it has been an active research area since the 1950s. While there has been considerable progress, e.g. in developing systems of non-monotonic reasoning and theories of action, yet more new ideas are needed.
- **Learning from Experience:** Programs do that. The approaches to AI based on connectionism and neural nets specialize in that. There is also learning of laws expressed in logic. The field of Machine Learning is concerned with the question of how to construct computer programs that automatically improve with experience (Mitchell, 1997). Programs can only learn what facts or behaviors their formalisms can represent, and unfortunately learning systems are almost all based on very limited abilities to represent information.
- **Planning:** Planning programs start with general facts about the world (especially facts about the effects of actions), facts about the particular situation and a statement of a goal. From these, they generate a strategy for achieving the goal. In the most common cases, the strategy is just a sequence of actions.
- **Epistemology:** This is a study of the kinds of knowledge that are required for solving problems in the world.
- **Ontology:** Ontology is the study of the kinds of things that exist. In AI, the programs and sentences deal with various kinds of objects, and we study what these kinds are and what their basic properties are. Emphasis on ontology begins in the 1990s.
- **Heuristics:** A heuristic is a way of trying to discover something or an idea embedded in a program. The term is used variously in AI. Heuristic functions are used in some approaches to search to measure how far a node in a search tree seems to be from a goal. Heuristic predicates that compare two nodes in a search tree to see if one is better than the other, i.e. constitutes an advance toward the goal, may be more useful.
- **Genetic Programming:** Genetic programming is a technique for getting programs to solve a task by mating random Lisp programs and selecting fittest in millions of generations.

We will explain which Artificial Intelligence techniques we have chosen to design this model and help solving this problem, from the most mentioned nowadays like Anomaly Detection, to using a clustering-based technique called K-means for Pattern Recognition and descriptive statistics techniques.

Data Mining

Data mining is a complex process that involves the interaction of the application of human knowledge and skills, and technology. This must be supported by clearly defined processes and procedures.

The main purpose of data mining is the extraction of hidden and useful knowledge from large volumes of raw data. However data mining is not simply about technology, it brings together different disciplines like software engineering, computer science, business intelligence, human-computer interaction and analysis techniques. Phases of these disciplines must be combined for data mining project outcomes. CRISP-DM methodology defines its processes hierarchically at four levels of abstraction, from general to specific, allowing a project to be structured modularly, being more maintainable, scalable and the most important, to reduce complexity. At the first level we have the phases, the second level consists of generic tasks, the third level is the specialized task level, and the fourth level is the process instance.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

CRISP-DM describes the life cycle of a data mining project consisting of six phases: business understanding, data understanding, data preparation, modeling, evaluation and deployment. The sequence of the phases is not rigid, moving back and forth between phases is always required and allowing a data mining project to have a flexible development during its life cycle.

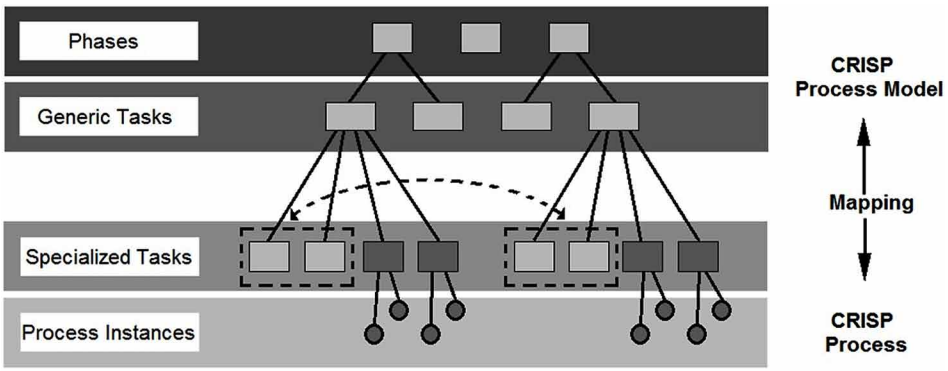
The most important and core in CRISP-DM methodology is that defines its processes hierarchically, consisting of a set of tasks at four levels of abstraction, from general to specific, explained below:

- Using levels of abstraction, at the top level we have the phases, the data mining process is organized into a number of phases, where each phase consists of several second-level generic tasks.
- At the second level we have the generic tasks, which must be general enough to cover all possible data mining situations and must be as complete and stable as possible. Complete means covering all possible applications and the whole process of data mining. Stable means this model and process must be valid for any kind of problems, and for old and new modeling techniques.
- The third level is the specialized task level, which describe how actions in the generic tasks should be carried out in specific situations. Ideally, the phases and tasks must be described as discrete steps and must be performed in a specific order. In practice, many of the tasks can be performed in a different order and it will often be necessary to repeatedly backtrack to previous tasks and repeat certain actions.
- The fourth level is the process instance, a record of the actions, decisions and results of an actual data mining engagement. A process instance is organized according to the tasks defined at the higher levels, but represents what actually happened in a particular engagement, rather than what happens in general.

According to CRISP-DM, the life cycle of a data mining project contains the phases of the project and the respective tasks and relationships between these tasks. At this description level, it is not possible to identify all relationships (see figure 1).

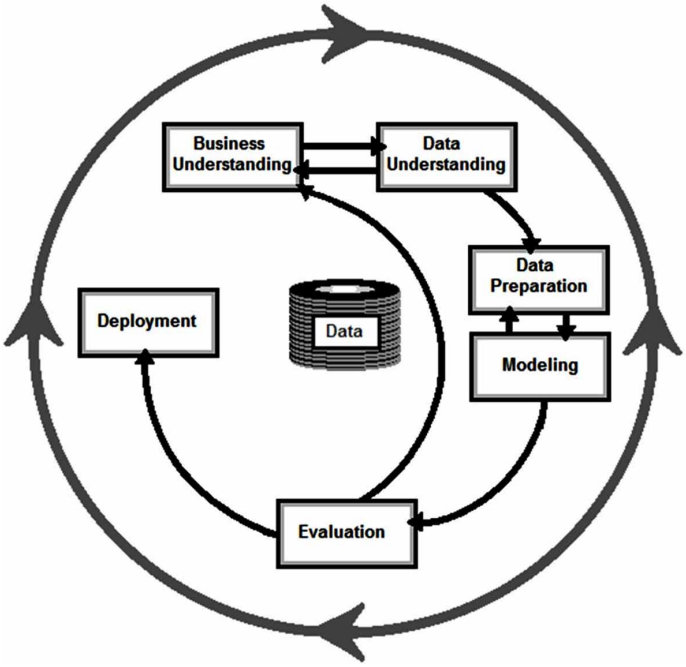
Essentially, relationships could exist between any data mining tasks depending on the goals, the background and interest of the user and most importantly on the data.

Figure 1. Four level breakdown of the CRISP-DM methodology



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Phases of the CRISP-DM reference model



A data mining project involves more than modeling. The modeling phase is just one phase in the implementation process of a data mining project. Steps of critical importance proceed and follow model building and have a significant effect on the success of the project (Chorianopoulos & Tsiptsis, 2009).

The CRISP-DM process model contributes to data mining as a process, which is reflected in its origins (McCue, 2007).

The life cycle of a data mining project consists of six phases. Figure 2 shows the phases of a data mining process. The sequence of the phases is not rigid. Moving back and forth between different phases is always required. It depends on the outcome of each phase which phase or which particular task of a phase, has to be performed next. The arrows indicate the most important and frequent dependencies between phases.

The outer circle in Figure 2 symbolizes the cyclical nature of data mining itself. Data mining is not over once a solution is deployed. The lessons learned during the process and from the deployed solution can trigger new, often more focused business questions. Subsequent data mining processes will benefit from the experiences of previous ones.

The Data Mining Phases

- **Business Understanding:** This initial phase focuses on understanding the project objectives and requirements from a business perspective, then transforming this knowledge into a data mining problem definition and a preliminary plan designed to achieve the objectives. This phase also includes assessing the current situation, establishing data mining goals, and developing a project plan (Delen & Olson, 2008).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- **Data Understanding:** The data understanding phase starts with an initial data collection and proceeds with activities in order to get familiar with the data, to identify data quality problems, to discover first insights into the data or to detect interesting subsets to form hypotheses for hidden information.
- **Data Preparation:** The data preparation phase covers all activities to construct the final dataset (data that will be fed into the modeling tool(s)) from the initial raw data. Data preparation tasks are likely to be performed multiple times and not in any prescribed order. Tasks include table, record and attribute selection as well as transformation and cleaning of data for modeling tools.
- **Modeling:** In this phase, various modeling techniques are selected and applied and their parameters are calibrated to optimal values. Typically, there are several techniques for the same data mining problem type. Some techniques have specific requirements on the form of data. Therefore, stepping back to the data preparation phase is often necessary.
- **Evaluation:** Model results should be evaluated in the context of business objectives established in the first phase (Delen & Olson, 2008). At this stage in the project you have built a model (or models) that appears to have high quality from a data analysis perspective. Before proceeding to final deployment of the model, it is important to more thoroughly evaluate the model and review the steps executed to construct the model to be certain it properly achieves the business objectives. A key objective is to determine if there is some important business issue that has not been sufficiently considered. At the end of this phase, a decision on the use of the data mining results should be reached.
- **Deployment:** Creation of the model is generally not the end of the project. Even if the purpose of the model is to increase knowledge of the data, the knowledge gained will need to be organized and presented in a way that the customer can use it. It often involves applying "live" models within an organization's decision making processes, for example in real-time personalization of Web pages or repeated scoring of marketing databases. However, depending on the requirements, the deployment phase can be as simple as generating a report or as complex as implementing a repeatable data mining process across the enterprise. In many cases it is the customer, not the data analyst, who carries out the deployment steps. However, even if the analyst will not carry out the deployment effort it is important for the customer to understand up front what actions need to be carried out in order to actually make use of the created models.

Anomaly Detection

Anomaly Detection is an artificial intelligence area which focuses in detecting patterns in data that are not complying with the expected behavior. These anomalies are also known as outliers in the data. These kinds of solutions are widely used in critical industry applications like intrusion detection in security, fraud detection, stock market, insurance, fault detection in critical electronic systems, among others.

Anomaly detection is about finding what we don't know to look for. We are searching for anomalies, but we don't know what their characteristics or properties will be present in the information. When we know the data in detail, we can use machine learning techniques, like classification, we can just write specific rules to find the anomalies. Classification is a form of supervised learning where in the data; we have examples of each kind of thing we are looking for. We can use a learning algorithm for these examples to build a model that can use features of new data to classify them into categories that represent each kind of data of interest. When we have examples of normal and some number of abnormal situations, classifiers can help

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

us to mark new situations as normal or abnormal. Even when we know about some kinds of anomalies characteristics, it is good to keep an eye to monitor new kinds that we don't know about. That is where anomaly detection is applied, so we use the unsupervised-learning approach of anomaly detection when you don't know exactly what you are looking for. Outliers (anomalies) are those events that don't match what you already know. Together, anomaly detection and classification make for a useful pair when it comes to finding a solution to real-world problems. We will use the anomaly-detection model to spot outliers.

So the first step in anomaly detection would be using our model to discover what is normal. This also requires human insight both to structure the model mathematically as well as to interpret what aspect of a pattern is of interest and whether or not it represents a reasonable view of a normal situation.

We then must experiment to determine at what sensitivity you want your model to flag data as anomalous. In one of the algorithms used for this problem, this is possible by calibrating a variable. If it is set too sensitive, random noise will get flagged, and with huge amounts of data, it will be essentially impossible to find anything useful beyond all the noise. Even if you've adjusted the sensitivity to a coarser resolution such that your model is automatically flagging actual outliers, you still have a choice to make about the level of detection that is useful to you.

For an anomaly detection solution, we need to identify the cost of false-positives and false-negatives, in order to know how we will be calibrating our solution. To exemplify a case, we can say that if a false-positive is too expensive, we will try to calibrate the algorithms to get a low quantity of anomalies.

For this problem, we will need to define a strategy to calculate these costs, and design the architecture based on this. There could be a need of a very high performance infrastructure in case the cost of false-positives is low for the solution, this means that the algorithms will end up with a results of a high volume of anomalies.

SOLUTION ARCHITECTURE

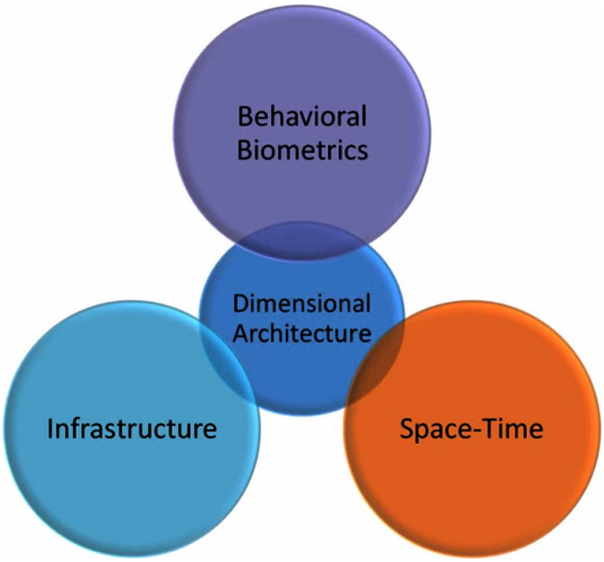
An Anomaly Detection Solution (within the solutions spectrum of Data Mining) has been designed through the modeling of statistics and artificial intelligence techniques, following CRISP-DM model standard for data mining solutions.

We describe how we have implemented CRISP-DM model and its activities below, for each phase of the methodology.

- **Business Understanding:** For this initial phase, we have analyzed and understood the project objectives, which are described in the introduction of the chapter.
- **Data Understanding:** For data understanding and preparation we used the same approach, looking at the initial data collection and analyzing the data. Also we have built a dimensional architecture.
- **Data Preparation:** During this phase, the modeling of a properties vector took place, looking for properties adding value to help to have rich data in the vector.
- **Modeling:** During the Modeling Phase, the selection of algorithms and creation of the model design were done. Iterations for the algorithms selected were run and it was necessary to iterate the design of the model. In this phase, various modeling techniques are selected and applied and their parameters are calibrated to optimal values. Typically, there are several techniques for the same data mining problem type. Some techniques have specific requirements on the form of data. Therefore, stepping back to the data preparation phase is often necessary.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Dimensional Architecture



- **Evaluation:** In this Phase, we were stressing the algorithms with real data, to iterate over results, and to see how real the data was behaving, and how efficient the algorithms were.
- **Deployment:** As we mentioned before, the creation of the model is generally not the end of the project, so we modeled this system for the results of data to be organized in a way that the information is easily understandable and valuable.

A dimensional architecture was designed to model the problem through three dependent and interconnected dimensions. The first dimension, Behavioral Biometrics, is a scientific area defined by the Human-Computer-Interaction (HCI), which this model has extended to Human Behavioral Patterns. The second dimension is Infrastructure, which represents all physical specialized equipment, environments, networking. The third dimension is space-time, which in this model represents a time dimension against all geographical information project related (code, file, systems and environments and its physical and logical locations).

This dimensional architecture will allow us to identify, separate and relate the different problem components, to understand and abstract each entity separately and be able to design and develop the problem in an organized way (see figure 3).

The dimensional entities are explained below:

- Behavioral Biometrics will represent Human Behavioral Patterns, where we will identify the individuals activities as team members of each project:
  a. Configurations
  b. Development Activities
  c. Testing Activities
  d. Technical Activities
  e. Others...

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Infrastructure is representing the physical entity of the problem:
  a. Equipment
  b. Environments
  c. Network Architecture
  d. Others...
- Space-Time is representing the dimension with which we will be able to contrast the other dimensions:
  a. Geographical information
  b. Time

It is very important to emphasize that the problem would not have a possible solution if one the dimensions was not existing, we exemplify this below with two possible scenarios:

1. Having only Infrastructure and Space-Time dimensions; in this scenario we will not be able to identify who is performing activities, neither to know what this individual is doing (behavior).
2. Having only the Behavioral Biometrics and Infrastructure dimensions; in this scenario we will not be able to contrast these dimensions with time, so we will not have the "when", and the information will not be complete.

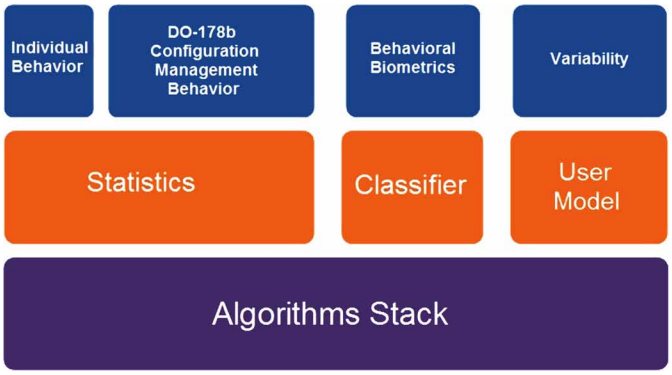
Now we can see that each dimension is critical for the problem, and that all possible data we can enter to each dimension will be important for the model.

For the Model Design and Architecture, a General Solution Framework was designed with two architectural stacks. The first stack is the Algorithms Stack, which will contain the Business Logic of the model containing a Statistics module, a Classifier module and a User Profiling module (see figure 4).

The second stack is the Utilities Stack, which will perform as an ETL standard process, containing a Preprocessing module in charge of the readers and parsers of the information and a Configurations module (see figure 5).

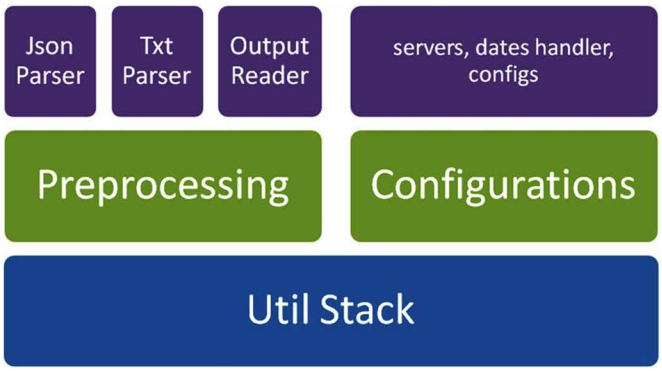
To explain the algorithms we have used in our architectural design, we will develop the methods in detail. In the first step of the model, a filter has been designed implementing a partitioning-based clustering algorithm, ***k-means*****,** which will attempt to find k non-overlapping clusters to recognize patterns

Figure 4. Algorithms Stack, The Business Logic



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Utilities Stack



and identify types of behaviors of each individual. Each cluster is represented by a mean point in the data, called centroid. This algorithm will iterate and group the data based on behavioral patterns until its convergence, this means that the algorithm has grouped the information.

In order to mathematically model the information to be processed, a properties vector has been designed containing Configuration Management behavioral data for each active team member of the project.

Attributes for the Properties Vector:

A1. Commit Size.
A2. Failed Changes Frequency.
A3. Unauthorized configurations Frequency.
A4. Incidents caused by releases.
A5. Number of Bugs.
A6. Code integration issues.
A7. Unit testing´s code coverage.
A8. Release document correctness.
A9. Overtime factor.
A10. Standards Compliance.
A11. Code design accuracy.
A12. Requirements traceability process execution.
A13. Tracking accuracy.
A14. Trunk destabilization.
A15. SVN File-Sharing problems.
A16. Big average file size.
A17. Unexecuted code.

A properties vector with this design would be formed as follows:

$v = \{3, 4, 2, 5, 0, 2, 5, 3, 1, 0, 1, 1, 2, 3, 3, 2, 0\}$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Data matrix example

	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	A13	A14	A15	A16	A17
Data 1	4	2	1	...	...	...	…	…	…	…	…	…	…	…	…	…	…
Data 2	...	…	...	…	…	…	…	…	…	…	...	…	…	…	…	…	…

Then data matrix of this problem will be represented by Table 1:

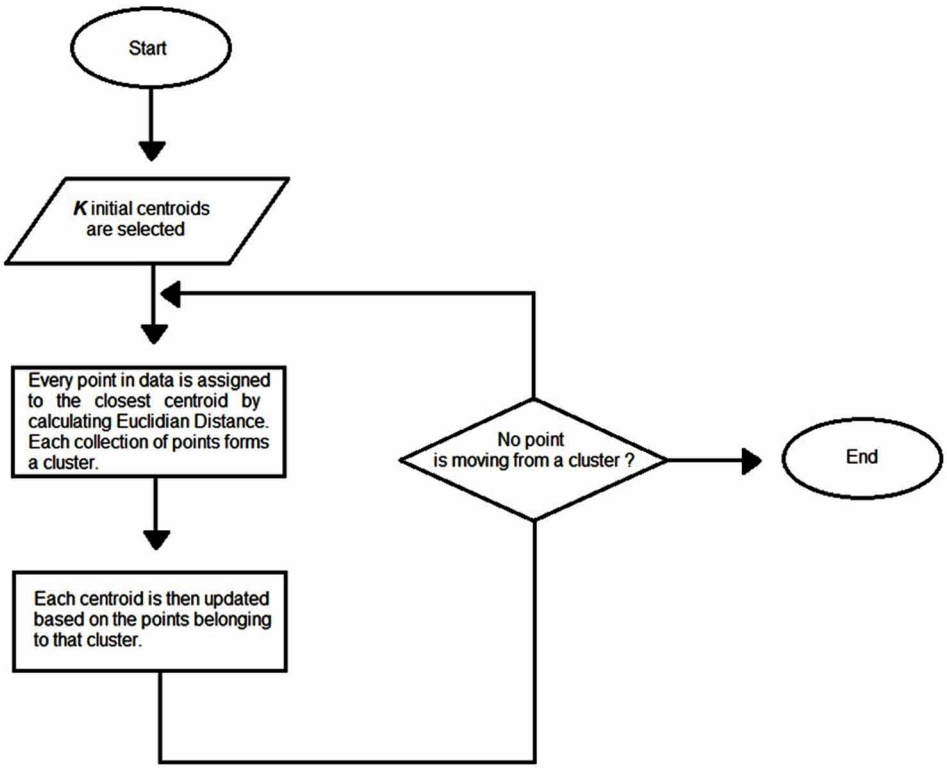
Based on DO-178b Configuration Management guidelines and best practices, specific data from the properties vector are labeled (Table 1).

The information is analyzed separately for Aerospace Software Development teams, and Aerospace Verification or Testing teams; this is because the Configuration Management activities within both groups can vary significantly and cause noise in the system. The objective of this separation is to avoid false-positives in the system results.

The clustering process of the algorithm is as follows (see figure 6):

Our K-means implementation for this problem is detailed in the procedure shown below:

Figure 6. K-means Algorithm Process



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. The Data Matrix

To explain the methodology we have designed with a visual example for representation purposes, we have chosen three of the seventeen properties of the vector model, shown in the following list (this way, a three dimensional figure can visually represent the individuals in our system):

1. Commit Size.
2. Failed Changes.
3. Number of unauthorized configurations.

The data Matrix of ten individuals or team members will be represented by the following table (see table 2):

2. Centroids Selection for Iteration 0:

For the centroids quantity selection, we will determine this number based on what we are looking for at this moment, which is patterns or types of behavior to determine anomalous data for each tester. We defined 4 centroids as the first set of centroids (the first four testers). We will denote the coordinates of the centroids by:

$C_1(4,2,1)$
$C_2(3,0,3)$
$C_3(2,1,2)$
$C_4(4,1,2)$

3. Objects-Centroids Distance for Iteration 0:

For this step of the algorithm, Euclidean Distance formula is used:

$$dij = \sqrt{\sum_{k=1}^{n} (x_{ki} - x_{kj})^2}$$

Table 2. Data matrix

	Tester 1	**Tester 2**	**Tester 3**	**Tester 4**	**Tester 5**	**Tester 6**	**Tester 7**	**Tester 8**	**Tester 9**	**Tester 10**
A1	4	3	2	4	4	3	2	4	2	4
A2	2	0	1	1	2	0	1	1	5	2
A3	1	3	2	2	1	3	2	2	0	1

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In this case, we have *x,y,z* coordinates to be used in the formula to obtain the Distance Matrix for iteration 0, calculating the distance between each centroid to each object (testers data):

$$D^0 \begin{bmatrix} 0 & 3 & 2.45 & 1.41 & 0 \\ 3 & 0 & 1.73 & 1.73 & 3 \\ 2.45 & 1.73 & 0 & 2 & 2.45 \\ 1.41 & 1.73 & 2 & 0 & 1.41 \end{bmatrix} \begin{matrix} C1(4,2,1) group1 \\ C2(3,0,3) group2 \\ C3(2,1,1) group3 \\ C4(4,1,2) group4 \end{matrix}$$

$$\begin{matrix} t1 & t2 & t3 & t4 & t5 \end{matrix}$$
$$\begin{bmatrix} 4 & 3 & 2 & 4 & 4 \\ 2 & 0 & 1 & 1 & 2 \\ 1 & 3 & 2 & 2 & 1 \end{bmatrix} \begin{matrix} x \\ y \\ z \end{matrix}$$

4. Objects Clustering for Iteration 0.

Now we have to assign each object to a group based on the minimum distance. The elements of the Group matrix below has a value of 1 if and only if the object has been assigned to that group or cluster.

$$G^0 \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{matrix} group1 \\ group2 \\ group3 \\ group4 \end{matrix}$$
$$\begin{matrix} t1 & t2 & t3 & t4 & t5 \end{matrix}$$

5. Iteration 1, Determine centroids.

Now we know the members of each group or cluster, we calculate the new centroid of each cluster based on the new memberships, where the centroid of each group is the average coordinate among the members. In groups with one member, the centroid is the only object.

$$C1 = \frac{4+4}{2}, \frac{2+2}{2}, \frac{1+1}{2} = \frac{8}{2}, \frac{4}{2}, \frac{2}{2} = (4,2,1)$$

$$C2 = (3,0,3)$$
$$C3 = (2,1,2)$$
$$C4 = (4,1,2)$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

6. Objects-Centroids Distance for Iteration 1:

We will repeat step 3 with the new centroids coordinates we just obtained.

7. Objects Clustering for Iteration 1.

We will repeat step 4 with the new centroids coordinates we just obtained.

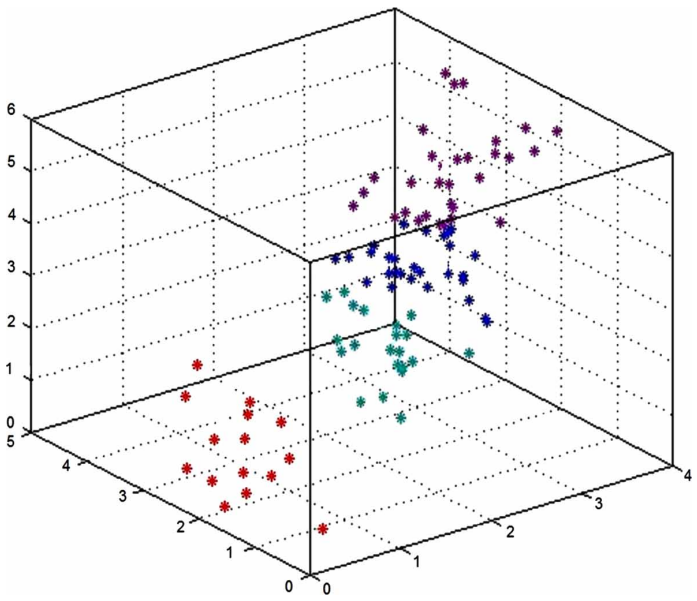
8. When the group *Gn*=*Gn* - 1, the algorithm has converged, and we have a result, no objects are moving, and we have our clusters defined.

$$G^0 \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \begin{matrix} group1 \\ group2 \\ group3 \\ group4 \end{matrix}$$
$$t1 \quad t2 \quad t3 \quad t4 \quad t5$$

The following figure (figure 7) illustrates visually a converged k-means for this problem, but with 91 objects we have added, in order to have an understandable representation of our 4 clusters.

The following step is the cluster selection performing an exploratory analysis, in order to discriminate the rest of the data.

Figure 7. K-means Results Represented in a 3 Dimensional Plot



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

An analysis to try to identify a threshold and categorize a good behavior and a bad behavior with all seventeen attributes of data together will be performed to the different components of the properties vector, this is, k-means clusters results will be analyzed to find which cluster is maximizing or minimizing behavioral values, this is, which cluster represents a good or a bad behavior. This part of the design can be categorized within the Non-Supervised Learning techniques. The cluster with more anomalous characteristics will be selected and will be entering to the next statistical filter which will utilize the exploratory data represented in the figure below (see figure 8).

As we can see in the above image, exploratory data will come from different entities or individuals; testers and developers working with DO-178 standard in its b revision.

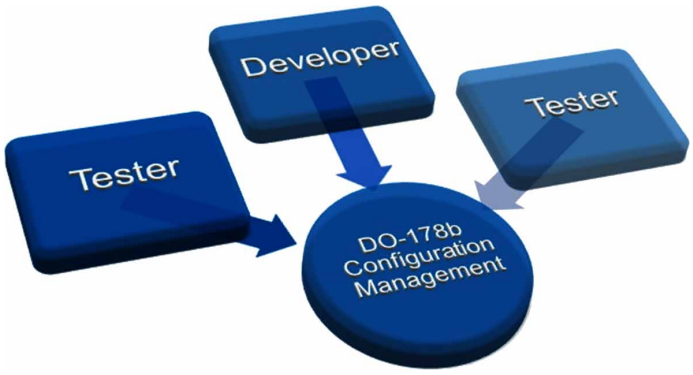
For the next step of this solution, the anomaly detection, an exploratory analysis is performed by using an anomaly detection technique of the descriptive statistics area. In this area, a box plot or boxplot is a convenient way of graphically depicting groups of numerical data through their quartiles. Box plots display variation in samples of a statistical population without making any assumptions of the underlying statistical distribution. The spacings between the different parts of the box indicate the degree of dispersion (spread) and skewness in the data, and show outliers. In addition to the points themselves, they allow one to visually estimate various L-estimators, notably the interquartile range, range, midrange, and trimean. Boxplots can be drawn either horizontally or vertically (McGill, et al., 1978). John W. Tukey, an American mathematician, introduced this type of visual data display in 1969.

A graphical representation of a boxplot and a probability density function, emphasizing the quartiles, is shown in the figure below (see figure 9).

The outliers' detection will be performed by an exploratory data analysis through Tukey Criteria, a statistical method constructing a lower and a higher thresholds to flag data as anomalous, where the lower and upper Tukey limits (lower and higher thresholds) correspond to a distance of 4.5 σ (standard deviations) from the sample mean. Any observations falling beyond these limits are considered as anomalous.

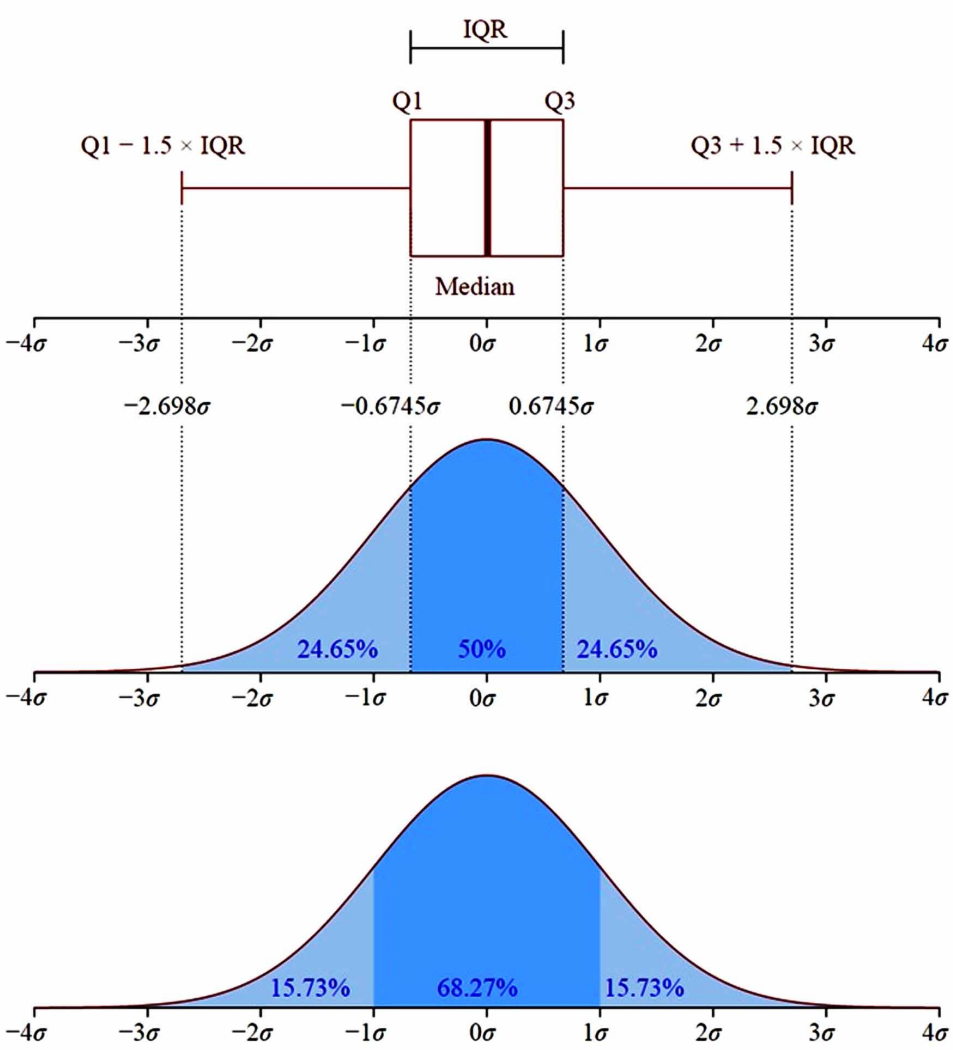
To explain how the probability distribution is formed by the problem data, we have modeled the information with vectors identified by two dimensions: *i* and *j*, where *i* represents the individual ID (i.e. 1714), and *j* represents the day ID (i.e. 01.25.2015).

Figure 8. Exploratory Data



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Boxplot and a probability density function of a Normal N(0,1σ2) Population

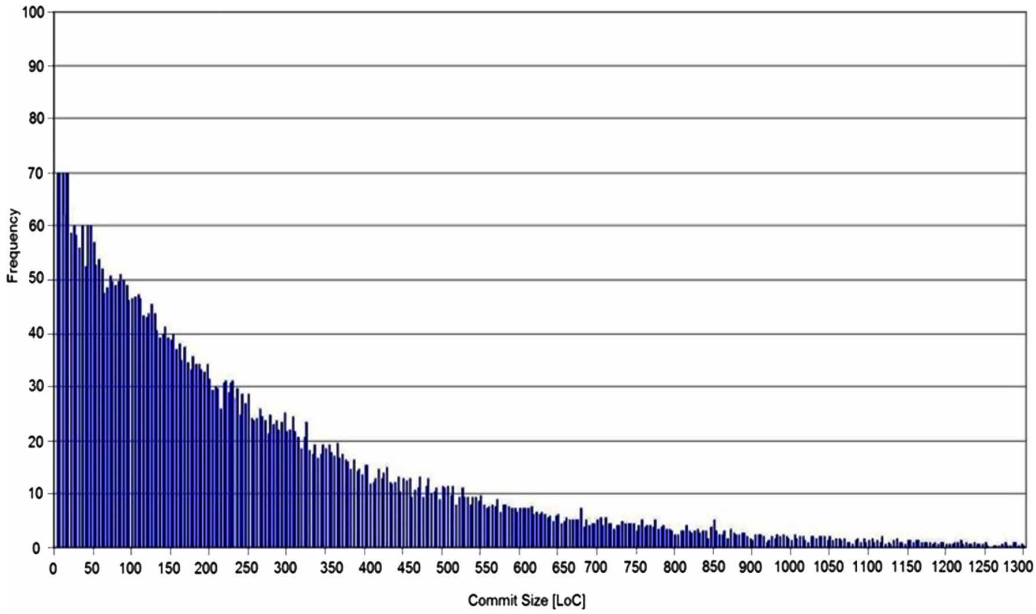


Vectors of different users and dates with the form $V_{ij}\{x,x1,x2\ldots\}$, would be represented as follows:

$$V_{1714,05.19.2015} = \{3,4,2,5,0,2,1,4,2,5,4,2,0,2,1,4,2\}$$
$$V_{2302,05.19.2015} = \{1,4,2,5,4,2,0,4,1,1,0,2,3,5,4,2,0\}$$
$$V_{1995,05.20.2015} = \{0,2,1,4,2,5,4,2,2,5,1,4,1,4,2,5,4\}$$
$$V_{0988,05.20.2015} = \{3,3,2,0,3,2,5,4,2,2,5,1,4,3,2,0,3\}$$
$$V_{1554,05.21.2015} = \{1,4,2,2,5,0,2,1,4,2,5,4,2,4,2,2,5\}$$
$$V_{1645,05.21.2015} = \{3,0,2,4,5,2,2,5,0,2,1,4,2,2,4,5,2\}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. Frequency Distribution of Attribute 1 in our Data Matrix (2 weeks)



This data will be arranged in an ascending order from the smallest to the largest observation. The ordered data is broken into four quarters, the boundary of each quarter defined by Q_1, Q_2, and Q_3, called the 1st quartile, 2nd quartile, and 3rd quartile respectively. The difference $|Q_3 - Q_1|$ is called the inter-quartile range. The Tukey upper and lower thresholds for anomalies respectively are: $ltl=Q_1 - 3|Q_3 - Q_1|$ and $Q_3+ 3|Q_3 - Q_1|$· Observations falling beyond these limits are called serious anomalies and any observation, X_i, $i= 1,2,\ldots,n$ such that $Q_3+ 1.5|Q_3 - Q_1| \leq X_i \leq Q_3+ 3.0|Q_3 - Q_1|$ will be called a possible anomaly (Wang Ch. et al. 2011).

To illustrate one of the 17 distributions types we have worked with, the following distribution is representing the Commit Size Attribute Frequency from our data matrix in a DO-178 Project, in a period of time of 2 weeks (see figure 10).

With this frequency distribution we can see that we are working with an Exponential Distribution (see figure 11). The Tukey criteria formula for this the type of distribution ideal for this project (exponential probability distribution) would be defined as follows:

$$Q3 + 1.5IQR = \ln(4)/\lambda + 1.5\ln(3)/\lambda$$

The results for the Tukey Criteria Analysis will be thrown for each Vector´s Attribute Frequency Distribution which we will refer to as "Rules". After a Tukey Analysis has been performed, we can observe the results in a bar chart represented in the image below to identify anomalies per rule (see figure 12):

Tukey method allows the user flexibility in setting the threshold limits (see table 3). This way we can calibrate the strictness level of the algorithms. This will depend on the problem, the model efficiency, and the needs of the user. The upper and lower limits for Tukey can be set to $Q_3+ k|Q_3 - Q_1|$ and $Q_1 - k|Q_3 - Q_1|$.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Exponential Probability Distribution Function

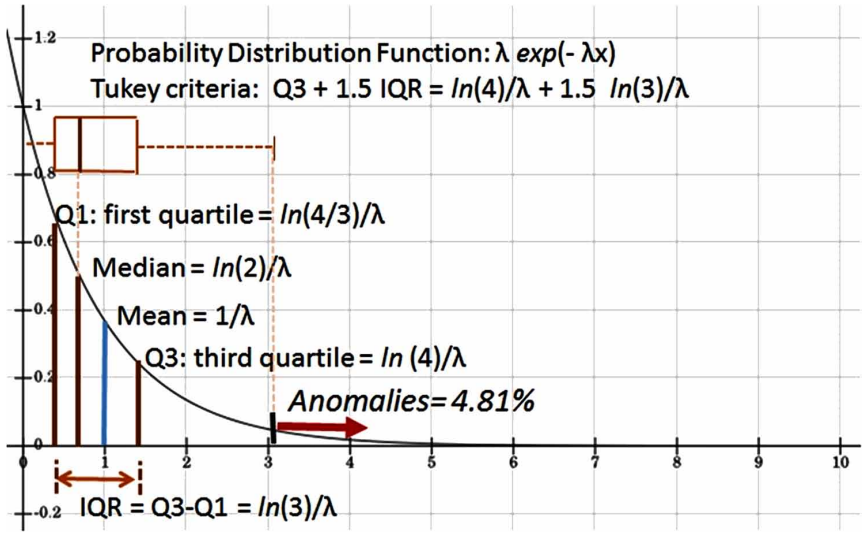


Figure 12. Anomalous Users per Attribute Rule

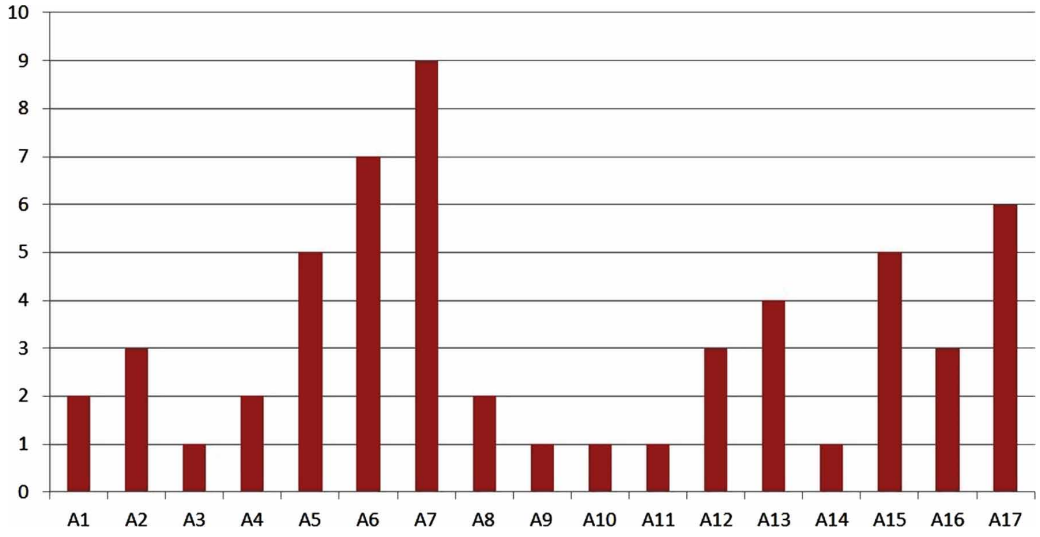


Table 3. Tukey Threshold Calibration

Tukey Criteria	Threshold	Anomalous Population
$Tk_1 = 1.5$	X = 3.034213 / λ	4.81125%
$Tk_2 = 0.834044$	X = 2.302585 / λ	10.0%
$Tk_3 = 0.464974$	X = 1.897120 / λ	15.0%
$Tk_4 = 0.203114$	X = 1.609438 / λ	20.0%

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

It is important to mention that at the beginning of each project, there is a phenomenon we call calibration period for the algorithms that take place during the first two weeks (initiation period). During this period, the system's uncertainty is high because there is not enough data to have a good level of efficiency in the algorithms results. This will be decreasing in parallel with increasing historical data after this period (see figure 13).

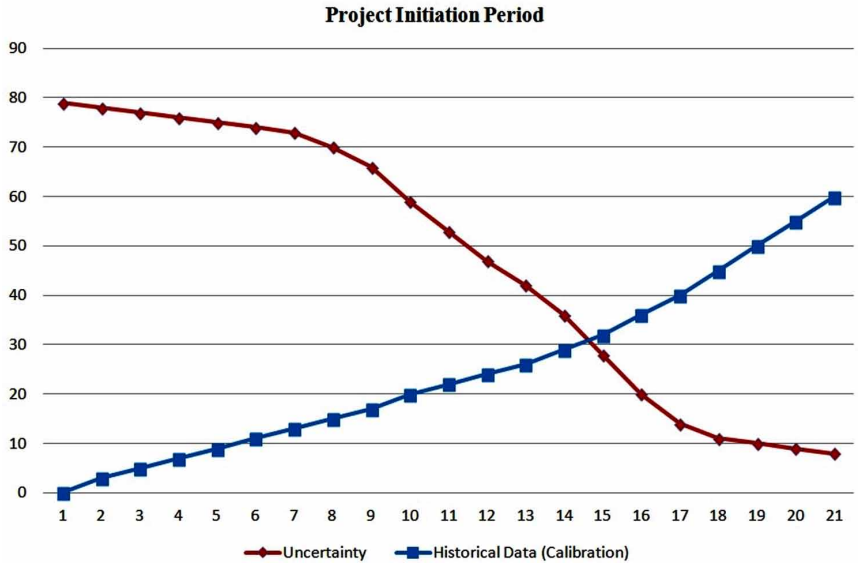
A Remediation Process for noise reduction purposes has been designed; this process is a quarantine that takes place after each algorithm's cycle has delivered results of anomalous data. This Remediation process takes the outliers individuals and start a pre-defined period of time quarantine of deep detailed supervision, where the system compares the behavioral biometrics in deeper detail, and the results for each individual are compared against a longer period of time. This process allows the system to look for false-positives or noise in the system, this way the system is able to eliminate possible noise.

The pre-defined period of time for the remediation process will depend on the combination of two constants of the system, the false-positives cost and the true-negatives cost for the needs of the project.

The results of this Remediation Process can represent two scenarios: The first one is when the outlier(s) was detected as noise in the system, then this individual is extracted of the remediation process. The second scenario is when the remediation process result is a true-positive, then the individual is a real outlier, and specific actions are to be taken.

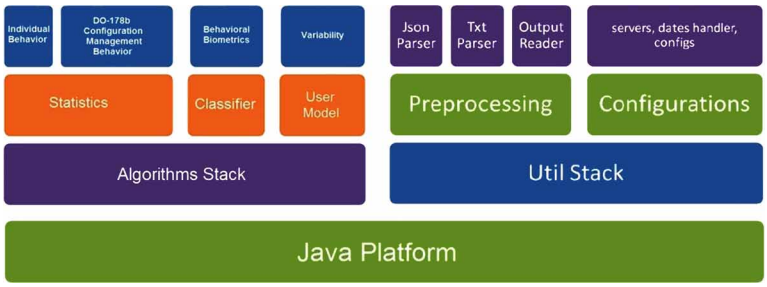
A real-time notifications system has been designed, which will notify when an individual enters or exits Remediation, this will trigger an alert to the manager, so a deep analysis of the case can be performed with the information provided by the system. The outliers' notification will report a graphical behavioral diagnosis to demonstrate the results (see figure 14).

Figure 13. Calibration Period Representation



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 14. Solution Framework



FUTURE RESEARCH DIRECTIONS

The design of solutions for risk identification within the spectrum of complex systems will be always involving different areas of the science and different approaches.

After working with this project, we were identifying possible future work for scope augmentation on this problem. We can add some other techniques like Principal Component Analysis if we increase the number of factors and variables to identify risks company wise, instead of identifying risks at a project level. We also identify the needs of a prediction model in the future, to be able to exploit a big quantity of projects historical information with forecasting techniques like Time Series, Linear Regression, Support Vector Machines or Neural Networks. A prediction model would let a company to be prepared in advanced, and create strategies to improve its projects efficiency.

CONCLUSION

Configuration Management provides all the important control elements to the process: configuration items (CIs), environments' configurations, project libraries, version control, records and other configurations to identify parts and software that should be perfectly working in compliance with the approved versions, but this does not have an impact in the real-time detection of process and human errors, or requirements' specifications during the execution.

Configuration Management risks and issues have not been decreasing over time or with technology improvements, on the contrary, its complexity is growing with time. This model represents a significant help in risks identification for software development and testing projects in the aerospace industry. The behavioral diagnosis when reporting outliers are a valuable visual tool not only to avoid impacts in projects, but also to help to create real-time performance improvement plans for engineers, and to optimize and continuously improve the Configuration Management strategy, with valuable statistical data obtained by this model.

REFERENCES

Chorianopoulos, A., & Tsiptsis, K. (2009). *Data Mining Techniques in CRM: Inside Customer Segmentation*. Wiley.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Delen, D., & Olson, D. (2008). *Advanced Data Mining Techniques*. Springer. doi:10.1002/9780470172339

McCarthy, J. (2007). *What Is Artificial Intelligence?* Stanford, CA: Stanford University.

McCue, C. (2007). *Data Mining and Predictive Analysis: Intelligence Gathering and Crime Analysis*. Butterworth-Heinemann.

McGill, R., Tukey, J. W., & Larsen, W. A. (1978). Variations of Box Plots. *The American Statistician*.

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.

Radio Technical Commission for Aeronautics. (1992). *Software Considerations in Airborne Systems and Equipment Certification*. Washington, DC: RTCA, Inc.

Salvendy, G. (2001). *Handbook of Industrial Engineering: Technology and Operations Management*. New York, NY: John Wiley & Sons, Inc.

Schalkoff, R. J. (1990). *Artificial Intelligence: An Engineering Approach*. McGraw-Hill College.

Wang, Ch., Viswanathan, K., Choudur, L., Talwar, V., Satterfield, W., & Schwan, K. (2011). Statistical Techniques for Online Anomaly Detection in Data Centers. *12th IFIP/IEEE International Symposium on Integrated Network Management (IM 2011)*. 10.1109/INM.2011.5990537

KEY TERMS AND DEFINITIONS

Anomaly Detection: A Data Mining technique in charge of the identification of events, items and observations not expected as a pattern.

Biometrics: In Computer Science, this term refers to the metrics related to human characteristics.

Clustering: An Unsupervised Learning technique of Data Mining used for identification and grouping of objects with similar characteristics.

Configuration Management: An area of the Software Industry related to configuration control activities.

Data Mining: An area of Artificial Intelligence related to the Knowledge Discovery through exploiting big quantities of data.

Pattern Recognition: A science that focuses in finding regularities in data.

Safety Critical Systems: Systems where a failure can result in big damages like loss of life, environment or property damage.

This research was previously published in the Handbook of Research on Military, Aeronautical, and Maritime Logistics and Operations edited by Alberto Ochoa-Zezzatti, Jöns Sánchez, Miguel Gastón Cedillo-Campos, and Margain de Lourdes; pages 319-345; copyright year 2016 by Business Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 8
Smart Grid and Demand Side Management:
Application of Metaheuristic and Artificial Intelligence Algorithms

Alper Ozpinar
Istanbul Commerce University, Turkey

Eralp Ozil
Zeta Information Technologies, Turkey

ABSTRACT

Energy becoming more and more crucial and critical in the civilized populations and locates itself as one of the major requirements of living standards. Obtaining the energy from fossil fuels still is one of the common sources of energy production; however, there is a common understanding of increasing the potential use of renewables, carbon capture and storage, energy efficiency and intelligence and smart applications for collecting, distributing and transmission of the energy between the supply and demand locations. Those applications and generating the new policies, roadmaps in order to make an energy revolution and increase the usage of low-carbon energy technologies targeting the decrease of energy related emissions. In this chapter, the authors explains the common issues about smart grid and demand side management and possible use artificial intelligence and metaheuristic algorithms for smart grid and demand side management optimization and scheduling.

INTRODUCTION

From the beginning of life in our universe, the mankind has needed and depended on energy in one form or another, from food to the operation of the most sophisticated machine we know today, for its existence, survival and progress. As the key commodity in the world, most physical laws, axioms and rules that are valid today either directly or indirectly related to energy. Especially with the start of industrial revolution it has become clear that the man needs to control the energy and the energy controls

DOI: 10.4018/978-1-7998-0951-7.ch008

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

the man. This dual relationship finds its true meanings from the laws of thermodynamics. The first law of thermodynamics is actually nothing but the accounting process of the energy, which eventually leads to the somewhat philosophic statement of the concept of conservation of energy which means that the energy can be transformed from one form to another, but cannot be created or destroyed. In other words, as we engineers say, the total energy of an isolated system is constant. The second law identifies which processes naturally occur in the environment and which do not. Because these two laws and many others related ones tell us that it is impossible to create new energy in our universe, we have over all these years taken the only plausible route and have chosen to transform otherwise non useful forms of energy, however abundant they may be, such as fossil fuels and brought it in one useful form or another to our homes, to industry and put them into use wherever and whenever we need it. Transforming processes include exploration, extraction, enrichment, cleaning, transporting, conducting, transmitting, distributing of energy among others.

Among the forms of energy known to all of us, electricity has become the leading form of energy starting from the second half of the twentieth century. With the recent improvements on the global education, increased awareness on the environment, and the goal of all the countries on increasing the welfare and prosperity of their people, the demand for electricity is increasing at a faster rate than any other form of energy. It is a known fact that new technological advancements and related end products most likely require electricity. In addition, environmental concerns has inevitably increased the demand for electricity consuming equipment at end use levels. Especially ever increasing demand for use electricity both for short and long range transportation in the sector which is in dire need for pollution control, will skyrocket the demand for electrical power. Finally, the diurnal variation for electricity demand in the industry, buildings and the commercial sector necessitates larger production facilities with high peak loads. This unstoppable increase in demand for electricity puts the existing power grids in a tremendous stress in all the countries.

Power grid is defined as interconnected network of power stations, substations, transmission/distribution lines, transformers that produce electricity from the power plants and deliver to the consumers. Power stations are operated by fossil fuels such as natural gas, coal, oil, nuclear fuel or by renewable resources. For the majority of last century electricity was mostly generated in large centralized power plants (Chan et al., 2012). And no care was paid whether or not they were placed and operated near heavily populated area. Some of the major characters of this type of a power grid is that the power is produced centrally. Secondly, in these classical power grids, electricity flows only in one direction, from the production point to the consumption site. Finally, the longer the transmission and distribution network, lower is the efficiency, reliability, and security. we may also add to these drawbacks, the issue of sustainability.

However when the first signs of climate change and increasing pollution appeared during the second half of the last century, it became apparent that there was a need for implementing new policies on the choice and quality of the fuel types, production processes, and setting new stringent pollution targets. In other words, the whole system of electricity production, transmission, distribution and consumption on the supply side and the energy consuming systems on the demand side needed new thinking, research, design, production and operation methods to be redefined and realized.

The reduction of greenhouse gases and emissions by using carbon free and clean production technologies like renewables is fast becoming one of the major alternatives to energy harnessing. But the renewables have issues of their own. First of all, if not most of renewables such as solar, wind, and hydraulic energy to certain extend all are intermittently available and their intensities are low. Just like fossil fuels

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

their availability widely varies on the earth. Some locations has a great wind power potential, some has solar, some has a small hydro potential as well as the biomass and other possible renewables. But none of them will be available at certain capacity at all times everywhere. We should never the less note that the number of renewable power stations are increasing at a promising pace all over the world. However, still theoretical and practical research on capturing and converting methods for most of the renewable systems, short and long term storage options as well as durability, maintenance and related fields.

We should emphasize that insisting on the continued use of fossil fuels for producing electricity as it was rightfully put in front of the humanity first by United Nations Framework Convention on Climate Change (UNFCC) in 1992, later by KYOTO Protocol and Copenhagen Convention, will increasingly produce wide range of pollutants and irreversibly harm the atmosphere and the ecosystem globally. In addition, more centralized energy production based on fossil fuels will result adverse effects on human health as proven statistically by many countries in Europe.

Nuclear power seems to be an alternative for fossil fuel based energy production for many years and it is the second-largest source of low-carbon electricity generation worldwide, after hydropower. According to IEA nearly all new nuclear construction in recent years has taken place in price-regulated markets or in markets where government-owned entities build, own and operate the plants (IEA, 2015b). However long term nuclear waste storage problems, huge risk potential of catastrophic accidents such as 1986 nuclear meltdown in Chernobyl and 2011 earthquake and the resulting tsunami's destruction of various Japanese nuclear plants, decommissioning problems and security issues needs to policy makers reconsider their nuclear investments.

The idea of decentralized power production and smart grids has evolved from the fact that, switching completely to new renewables energies, that is replacing all the fossil fuel using power plants, is not going to be possible in a very short time or it may be never be possible. Smart grid approach basically considers both sides of the operation, one side being the supply side where more clean electricity production will be achieved and the demand side will truly have decentralized production or even micro-generation at homes, buildings, large housing complexes as well as renewable energy based generation and cogeneration plants. All these will be planned and regulated by the newer set of regulations and rules. Inevitably, decentralized generation means transmission and distribution has to evolve into a new form and options like onside storage, short term switching and the most crucial idea of "two way operating" which makes each and every nodal point both a consumer and a producer in the grid. So the overall progress has to be optimized and scheduled as the constraints of the grid needs to become more intelligent.

The main objectives of the chapter is to provide a common understanding of the use of new and improving algorithms for optimization and scheduling for smart grids and demand side management applications in order to meet the future requirements of energy markets and usage. There are already a number of classical optimization and planning applications and algorithms available without the use of artificial intelligence and metaheuristics. However due to the success and widely adaptability and the increase the computational power of the computers these methods are becoming more popular applicable.

BACKGROUND

This part of the chapter starts with a short review of the related topics in order to provide the understanding of the topics and relationship between them. There are some great recommendations in the references and further readings part who wants to get involved in to these algorithms more detail. The authors do

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

not provide any language specific codes whereas algorithms and general parameter selection recommendations has been made. There are abundance of commercial and open source software packages in the market for the general use of artificial intelligence and metaheuristic methods.

Future Projections and Scenarios

IEA International Energy Agency, which is aiming to promote energy security to physical disruptions in oil supply, looking for ensure reliable, affordable and clean energy for the world publishes several reports on energy policies and perspectives and future projections each year. Their reports are key sources that provides the user with a common understanding to the needs of the energy issues. According to IEA reports, since the first Conference of the Parties (COP) held in Berlin Germany, in 1995, greenhouse-gas (GHG) emissions have risen by more than % 25 and the atmospheric concentration of these gases has increased steadily to 435 parts per million carbon-dioxide equivalent (ppm CO2-eq) in 2012. According to the International Panel on Climate Change (IPCC), humanity and all the countries have to take fully committed and urgent action. Otherwise climate change will have severe and irreversible impacts across the world. The international commitment to keep the increase in long-term average temperatures to below two degrees Celsius (2 °C), relative to pre-industrial levels, will require substantial and sustained reductions in global emissions

In order to obtain this level and reach the goals energy sector needs to be transformed since two-thirds of all anthropogenic greenhouse-gas emissions produced by energy sector. There are some good news in recent years about the application of these policies and precautions. Global energy production oriented CO2 emissions stayed flat in 2014 even there is an increase of around 3% in the global economy. (IEA, 2015a; IEA, 2006; IEA, 2015b) It is assumed to be the energy sector can achieve a peak in GHG emissions by the year 2020, with almost the same level of economic growth and development.

There are three main scenarios widely accepted in order to control the global warming and climate change. Kyoto Protocol and Copenhagen Climate Change Conference studies and outcomes becoming outdated and replaced with new policies and scenarios. Those scenarios have a common goal but they differ according to their assumptions about the evolution and application of government policies. These scenarios are the “Intended Nationally Determined Contributions (INDC) Scenario”, the “Bridge Scenario” and the “450 Scenario”.

In the Intended Nationally Determined Contribution (INDC), countries and unions like Switzerland, European Union, Norway, Mexico, United States, Gabon, Russia, Liechtenstein and Andorra which collectively represent more than one third of the global energy-related CO2 emissions submit their contributions in order to obtain the goals. All of the contributors have their own targets mostly based on reducing their GHG emissions by at least x % compared with 1990 levels by the year “y”. The numbers for x and y changes from country to country.

The “Bridge Scenario” is an improvement to the INDC and creating a bridge between INDC and 2°C target by focusing on research and development activities for the key energy technologies and making these technological steps and improvements to build up the bridge. Some examples of these technological improvements will be based on increasing energy efficiency, limiting least-efficient coal power, raising renewable investments, reforming fossil fuel subsidies and reducing upstream methane.

450 Scenario on the other hand have a different perspective but same target of global increase in temperature to 2°C can be obtained by limiting concentration of GHG in the atmosphere to around 450 parts per million of CO2.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Smart grid, microgeneration and demand side management are key factors in all of the above policies and targets.

Smart Grid, Microgeneration, and Onside Storage

Power grid was defined above, as an interconnected network of power stations, substations, transmission/distribution lines, transformers that produce electricity from the power plants and deliver to the consumers. The classical grids have a history over one hundred years starting from the invention of electricity. However making a grid smarter requires new technologies in order to collaborate the market, producers, operators and various consumers and end point devices. According the definition of NIST, "Smart Grid" refers to a modernization of the current electricity delivery system so that the system able to monitor, protect and automatically optimizes the way of operation between the interconnected elements in the system. Energy transferred from centralized or decentralized distributed energy generator through the high-voltage transmission network and the distribution system, to industrial users, factories, commercial buildings and building automation systems, as well as energy storage installations and also to end-use consumers and their thermostats, electric vehicles, appliances and other household devices.

The Smart Grid mainly differs from one way flow operation of classical girds with a two-way flow of electricity and information to create an automated, widely distributed energy delivery network based on real time information and optimization. (Von Dollen, 2009)

There are major benefits that can be obtained from smart grids. First of all by the use of information technology, cloud computing, high speed communication between the devices in the smart grid system in order to automate and control the grid, the system behaves like an autonomous structure so power reliability and power quality can be obtained within the grid. This autonomous grid has a self-monitoring property to detect unsafe, insecure and risky situations which may result in malfunction or unreliable working conditions. Encryption, secure communication, cyber security included in to the all components of the smart grid when required and available. These standards and applications makes grid more safe and secure comparing with the classical grid. One major benefit of the smart grid which is mostly related with our chapter is the energy efficiency benefits. Smart grid can adapt itself to operate in different conditions providing higher efficiency, reducing the peak demand, minimizing the energy losses and interact with the end user. As mentioned on the previous paragraphs, there is a need for smarter technologies in order to reach the goals of different scenarios. Smart grid helps to reduce the GHG and other pollutants by reducing or swapping the energy production from inefficient energy sources, least-efficient coal power, supports the use of renewable energies even in the household level and moreover to replacement in the transportation from fossil fuel operated vehicles to electric vehicles. And finally smart grid has direct financial benefits for all contributors and actors. Operational costs reduced, avoided or balanced with different policies and operational models between customers and the market. The customers have a direct feedback option via display devices and indirect feedback within the bills. This will enhances the demand control and provide energy savings from reducing the peak demands. Controlling the demand will be explained more in detail in demand side management. (Gellings, 2009)

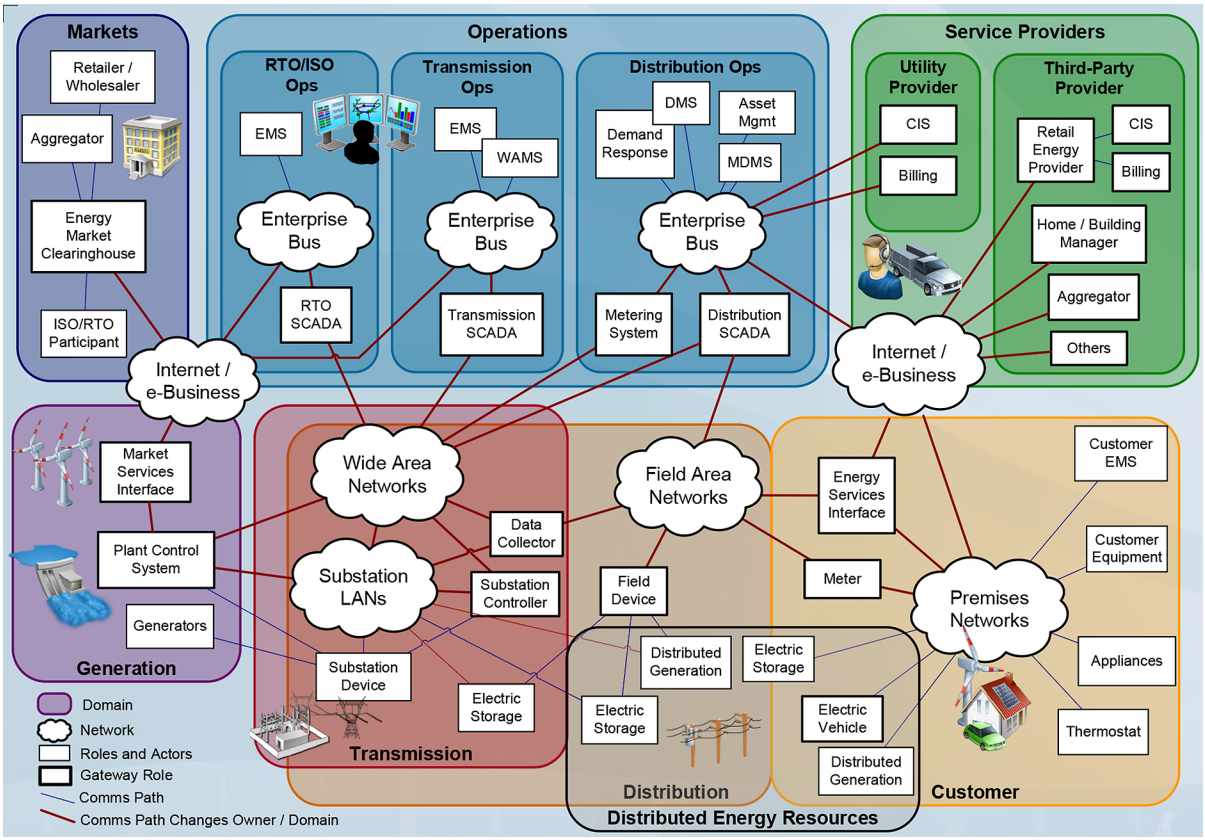
Smart grid also improves operational efficiency. In order to obtain and maximize these benefits grid has to be more smarter with the implementation and use of communications, electronics, and other related technologies like automatic sensors, data loggers, PLC's and control cards and units such as phase measurement units (PMUs) which measures and tracks voltages and currents at transmission lines and distribution stations. All these devices, units stations makes the grid smarter for reliable detection,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

isolation, and recovery of faults and greatly simplifies distributed state estimations in power systems for better monitoring for grid system operating stability, planning, scheduling, optimization, and recovery. Since there are lots of computational requirements for the monitoring and calculation of frequency, amplitudes, and transients of power signals different mathematical models and algorithms implemented in to smart grids. (Chan et al., 2012)

Two way operation system gives the potential to the customers being energy generators to their neighbors and the grid. Especially use of small wind turbines, smart buildings with PV's, solar water heating, air source heap pumps, micro CHP allows on-site energy efficiency production and storage. In order to support and encourage microgeneration and on-site storage systems governments offers various tax discounts, financial plans and credits with a guarantee for buying the excess energy produced via smart grid. Considering the future use of electricity driven transportation vehicles, and networks, microgeneration is becoming more important by removing most if not all the emission from vehicles and thus minimizing the carbon footprint. Figure 1 shows the NIST conceptual reference diagram for the smart grid and microgeneration. (Locke & Gallagher, 2012).

Figure 1. NIST conceptual reference diagram for the smart grid information networks



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Demand Side Management and Planning

As mentioned on the previous paragraphs smart grid includes the generation, transmission, distribution and finally consumption in two directions. Although the improvements and the innovations in the field of generation, transmission and distribution provides better results and efficiency in order to reach the goals in the scenarios effective controlling and management of the demand is also essential. A good demand side management policy, when implemented, helps to plan and optimize the generation capacity, plant utilization and efficiency by adapting the flexible demand hours on the time domain. Controlling the peak demand, reducing the duration of the peak hour operations, directly eliminates the inefficient, high carbon cost power stations from total generation capacity. They may still be used in the grid as backup and emergency generators. Controlling the peak and fluctuations also improves the operation and the efficiencies of the transmission and distribution networks. The smart grid network expected to be operate even loss of a single circuit due to a fault, malfunction or overload due to peak requirements the rest of the circuits in the grid takes over the operating load of the faulty line. So as a good operating practice requirement during peak-load conditions, grid circuits in the interconnected transmission network are generally loaded below 50% capacity. (Strbac, 2008).

There are various demand side management programs based on the energy efficiency on the demand side by replacing the devices with better, efficient and green technologies which is not directly related with the smart grid and optimization. A dynamic demand planning as a part of demand side management, monitors the smart grid's power factor which is defined by the ratio of the real power flowing to the load to the apparent power in the grid, and aims to increase the diversity factor of the grid by smart opening and closing or shifting the operating cycles which can be controlled and programmed automatically. A management on the demand response is the most widely and easily adaptable part of demand side management programs. The aim of the program is to stabilize the demand in to some limits regarding from the time and season and avoid peaks, fluctuations in the grid. This type of program can be achieved mostly by changing the demand characteristics of the consumers. In order to convince and control consumers' consumption habits educational means, social responsibility and awareness schemes can be employed. However, use of dynamic pricing based on the time of the day and the season is the most effective ways of achieving the goals. For example critical-peak pricing (CPP) in which prices during the top 40–150 hours of the year rise to previously specified levels designed to recover the full capacity and energy cost of power plants that run primarily during those hours. During all other 210 hours of the year, prices are lower than existing rates by an amount sufficient to leave the bill unchanged for a customer whose load shape mirrors that of the rate class. (Sioshansi, 2011) Practical use of CPP can be operating the washing machine, dishwasher and electrical heaters for taking a shower in the other than the peak hours mostly late in the evening or early in the morning where the general demand by the industry is low.

Methods for Decision Taking, Optimization, and Scheduling

Smart grid and demand side management applications and problems can be solved by various mathematical methods and algorithms. The overall solution progress includes both deterministic and probabilistic approaches. Energy production based on fossil fuel can be modelled and controlled by deterministic models, however in case of renewables and customer behavior that generates the demand and as well as the possible microgeneration within the grid cannot be modelled by using simple deterministic models. It is possible to calculate and simulate the model by using assumptions and sophisticated deterministic

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

models operated with calculated and defined parameters and rules. Even these models can calculate the demand and supply, the deterministic part of the model mostly static not dynamic. Making them dynamic requires lots of computational power, based on iterations and long calculations. Another point on the deterministic models they does not include randomness. Since the smart grid has the tools for information and gathering statistical historical data and real time operating data includes a bias and some uncertainty due to human and nature behaviors hybrid models includes both deterministic and probabilistic parts can be used.

Artificial intelligence can be considered as the art of perceiving the situation and taking decisions by the machines in which the rules are not clearly defined and there are some uncertainties within the situation so it is not an automation performed by a machine. Artificial intelligence has been misinterpreted with the keyword machine automation. An air conditioning system that operates according to the indoors and outdoors temperature levels looks like an intelligent system for most of the users. However, a programmable logic control based on the simple rules operated by the temperature information coming from the sensors gives the indoor, outdoor and comfort temperature values based on the readings and assumptions. Operating the air conditioner in cooling mode based on a simple proportional feedback controller solves all the problem. There is no decision nor an intelligence required in this progress after it has been programmed. Programming this controller involves a decision and intelligence progress by humans, but operating is totally different process. An improved version of automation can be considered as machine learning and expert systems. Where the lack of information and necessary acts can be taken by both rules and statistical learning progress. Table 1 shows a conceptual definition of automation and intelligence. As explained a plc automation is a pure rule based system without and decision. A bank ATM can be considered as a rule based automation machine, if the customer have a positive bank account then depending on the daily limits and amounts cash withdrawal can be automated. During this progress the ATM machine does not make any decision, all the progress clearly defined by the bank standard operating procedures. However when the customer asks for a loan or a credit the system transforms in to an intelligent machine acting like an expert system operating if-then rules and knowledge databases, moreover a much more intelligent system learning from previous customer experiences and evaluation of financial indicators also with standard operating rules of the bank, the ATM machine makes a decision of giving a credit or not. Moreover to this decision, the machine follows the payback period and progress, compares with the previous credits decisions segmenting them false positive, false negative, true positive and true negative and learns to make better decisions for the next customer.

Table 1. Automation versus intelligence by means of information, tools and operation data

PLC and Machine Control	Expert Systems	Metaheuristics	Machine Learning	Self-Aware AI
Automation with rules and SOP	if-then rules Knowledge database Decision tree	Optimization search Imperfect info Incomplete info Limited computation Resource allocation Stochastic process	Statistics Knowledge Optimization Learning Pattern Recognition Decision trees Artificial Intelligence	Decision without rules and SOP
100% automation 0% AI	0% < AI, Automation <100%			0% automation 100% AI

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Demand side management and smart grid operations are similar decision making processes. There are some rules based on operational procedures but cannot be classified as pure automation. There is no simple weather model accurately forecasting the weather for the short, mid and long term. The operators and the market need both supervised and unsupervised learning and which will improve the system outcomes for the policy and scenario targets. For example a load forecasting in the smart grid that uses weather forecasting by artificial intelligence and price optimization by metaheuristic model can be applicable. Main advantages of these methodologies are that, they are mostly dynamic, self-learning and self-improving algorithms so they can provide sustainable solutions for the future based on the current and historical data. So if the growth rate of the county will change in the future due to economic conditions and global competition, algorithms can improve the working parameters, set the new critical peak hour pricing values. Also increase and decrease in the pricing is based on common optimization rules.

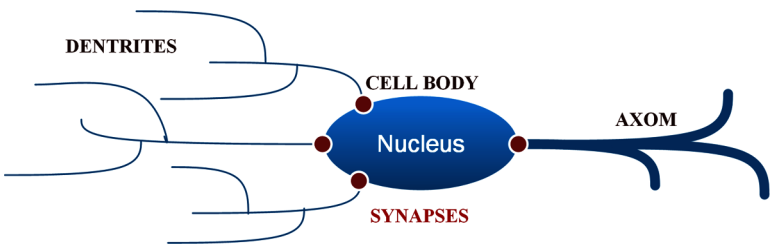
ARTIFICIAL INTELLIGENCE AND METAHEURISTIC ALGORTIHMS

This part of the chapter organized as a brief explanation of the algorithm, general usage and a pseudo code adapted to the proper example for smart grid and demand side management. As mentioned before there are a great number of resources, publications and books available related with the algorithms, so the authors aim to give a common usage and understanding about the algorithms.

Artificial Neural Networks (ANN)

Artificial Neural Network (ANN) is inspired from human neural information processing structure which formed by biological neurons. In the actual biological system a treelike nerve connectors, namely dendrites, are connected to the cell body. These dendrites act like communication bridges and receive signals from other nearby neurons. Cell bodies have a single long connector namely axon which connects the cell body to other neurons via dendrites or cell bodies, the connection point or the junction called the synapse. Axons can have tree diagram like strands and sub strands giving a possibility to connect a neuron to many other synapses. The transmission of information signal from one neuron cell to another at a synapse performed by a complex chemical process which is out of the scope of this chapter. But, in general, the dendrites serves as receptors for signals from other neurons, whereas the purpose of an axon is transmission of the generated neural activity to cells. (Yegnanarayana, 2009) An illustrational drawing for the biological neural network given is in Figure 2.

Figure 2. The biological neural network

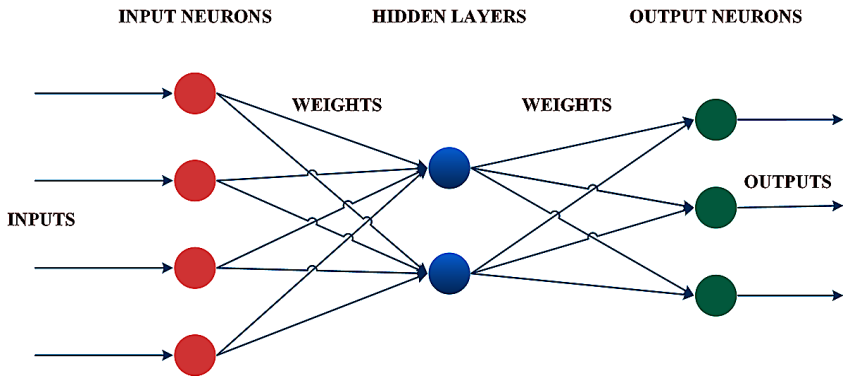


EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Up to now we have explained a perfect biological system which may seem a little irrelevant to our subject. However replacing the cell body with the concept of any node in the smart grid such customers, market, operations, service providers as in Figure 1, it suddenly becomes more logical and turns into a smart grid information network. In the case of artificial neural networks, the system is composed of artificial neurons connected to information inputs balanced with weights for each input, like the axons in biological neurons, artificial ones also connected to other neurons. There neurons can be hidden in order to improve the learning and adaptations progress. Figure 3 shows the general structure of artificial neural networks. Within in each node of the ANN, there is a transfer function used to collect the data from the inputs balanced with connection strengths or weight parameters and then an activation functions like step, linear, log-sigmoid, tan-sigmoid and Gaussian helps us translate the transfer functions input values to output values.

An artificial neural network needs an adaptation, or in other words, a learning process to understand and interpret the information flow according to inputs and outputs. Accordingly, in ANN theory, there are three main types of learning methods. In supervised learning, a historical data from the smart grid about supply and demand, peak hours, and operating zones are provided as an input and the output to the system from previous real data or data generated from deterministic models are tested in the field. So the input data is propagated forward through the ANN network model defined in the beginning. The output data from the system compared with the correct outputs and depending on the bias and the errors if obtained, the next step is to change the weights programmatically and use the new weights to start all over. This learning progress similar to a learning progress in a school when an instructor or a supervisor tells that the outputs or progress is good or some parts has to be updated. In supervised learning a set of input vectors and output vector used and considered to be learning a mapping from input set to output set. In most of the supervised learning, the learning data is prepared and used, after successful learning system works with the learned parameters and weights. The second type of learning is the unsupervised learning and is mostly used as a part of machine learning systems. In this type of the learning, the system looks for the grid information flow and the demand side information only the inputs to the system. Then the ANN network is doing self-learning depending on structure in the input data like the demographics, income level, geographical location, weather station inputs for those days, time based electricity consumptions numbers. Clustering, grouping or segmentation is common ways of this type of learning.

Figure 3. The artificial neural network



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In the final version of learning, reinforcement learning, combines two of the better learnings, so the system uses input to generate output depending on the initial parameters. The feedback from the grid or the customers is simply a scalar value which may be delayed in the time domain depending on the output decision type. The success and the failure of the decision made. So the system is learning the progress online when the grid is actually working. This type of learning can cause some malfunctions and critical errors if applied to decision related with the distribution or transmission lines. However, it can be used as a part of customer experience where there is no previous data available about the system while testing a new dynamic pricing model. The inputs will be the possible segmentation of price and time zones and the output will be the peak reduction levels.

A most convenient way of working with ANN is to use a package program in the beginning while making the test and then code or integrate the package to the system. The main disadvantage of the ANN is the type of the network, number of hidden layers, activation function selection, the number of hidden layer input, outputs and the learning type cannot be estimated in the beginning. There are some recommendations about which type of network and settings to use but in most cases researchers has to do some trial and error possible studies depending on the data and the information. ANN mostly used for modelling, curve fitting and forecasting the data.

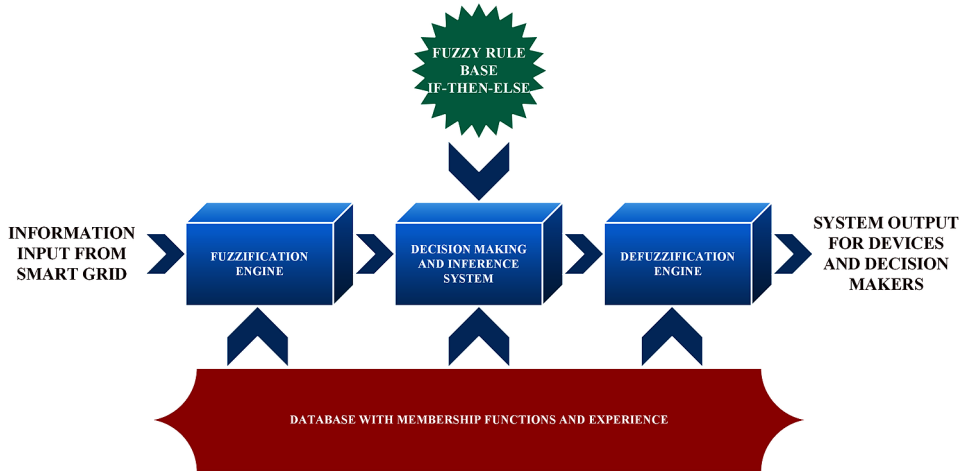
Fuzzy Logic

Fuzzy logic is an mathematical approach to modelling based on the degrees of membership or truth instead of the classical Aristotle logic which was later adapted to computer world by a system of symbolic logic by George Boole where there is only two options: true and false or a member of something or not. Actually, what the fuzzy logic does is to set the membership function to 1 or 0. Fuzzy logic based fuzzy information systems are very suitable for modelling and control of the devices and systems without dealing with complicated equations where possible. Most of the devices within the smart grid and also most of the demand side customer behaviors cannot easily be explained by classical logic. In general, fuzzy information systems create a mapping between the experiences based on linguistic fuzzy rules to real life values. Fuzzy logic systems are robust, can be modified or tweaked easily, tolerant to imprecise data, based on natural language statements but requires those expressed as a part of historical or theoretical experience. The most classical example of fuzzy logic about the decision making for the room temperature can be perfectly adapted into a smart grid system. If the household tries to make a plan for energy saving from air conditioner in the house he has to be precise about when to start the cooling and when to stop. Setting a limit value to start and stop in classical control theory can solve the problem however most people prefers to say the weather is too hot instead of telling as the exact temperature in numbers. If home automation system based on classical logic and a little old system for variable operating conditions, there will be too much on and offs in the operating cycle which may generate fluctuations in the grid. However using a smart controller based on fuzzy logic, the previous experience of the household can be expressed by fuzzy rules in language statements like, if the weather is hot, and humidity is low then run the air conditioner in the first stage and so on. A fuzzy sets and fuzzy membership functions will be used to define the linguistic experiences in to fuzzy engine.

So in normal operation, the inputs to the fuzzy information system from the sensors, inputs from the systems, read values from metering devices and those real scalar numbers are converted as real fuzzy values. This process called, “Fuzzification” as shown in Figure 4. Different fuzzy membership functions are available for this conversion; such as triangular, trapezoidal, gaussian, and generalized bell shape.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. The fuzzy information system for smart grids



There are two ways of expressing a value by means of linguistic variable, first giving a separate set name for each status. For example, considering the temperature case, the fuzzy sets will be hot, normal and cold. But sometimes finding the correct linguistic definitions is not that easy, in that case a degree of negative, neutral and positive or small, medium and big can be used also with adjectives like "very" to make it more precise. In general 3, 5, 7 adjective naming system used for different sets. After the input values are processed by fuzzification, these fuzzy values are subjected to fuzzy decision making logic which is controlled by the fuzzy rule base. Mamdani and Takagi - Sugeno methods are possible approximation and information extraction models of the fuzzy logic. Last process of fuzzy information system is the "Defuzzification" of the information to real scalar numbers where they can be processed by controllers, programs and applications.

Fuzzy logic can be used in the demand side management system for optimizing the demand response and flexible pricing. Experience based rules can be written easily to perform the operations.

A further application of fuzzy logic is to combine it with artificial neural networks. Integrating the linguistic reasoning and rule based system of fuzzy logic with the learning capabilities of neural networks results in hybrid Fuzzy Neural Network (FNN) or Neuro-Fuzzy System (NFS). This type of hybrid application improves the fuzzy logic systems in their field of experience requirement and lack of learning.

Particle Swarm Optimization (PSO)

Similar to the inspiration of the biological neurons, particle swarm optimization is inspired from the behaviors of the various swarms in the nature such as birds, fishes and ants. Particle swarm optimization is considered to be a stochastic population based optimization. This is the reason why it is suitable for smart grid and demand side management applications. In the literature of the PSO it is accepted as there are five major swarm behaviors and actions in the populations (Millonas, 1992). Considering these populations as the grid member the new behaviors can be listed as;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. **Proximity:** Ability to perform space and time computations, as described in the grid section of the chapter with intelligent devices at the generation and consumption locations, sensors, actuators and controllers connected to the information grid. Most of the instrumentation will have enough microprocessor power to perform computations or at least with a connection to information network or high performance computing also with the use of Internet of Things (IoT) devices and the computational power of cloud computing where the computations can be made.
2. **Quality:** Ability to respond to environmental quality factors, so the smart grid node devices responds to the pricing and tariff changes and communicate with the other devices for dynamic response.
3. **Diverse Response:** Ability to produce a plurality of different responses, so each device or each node in the grid can perform also independently which in turn could be very useful in the case of microgeneration and on-site storage.
4. **Stability:** Ability to retain robust behaviors under mild environmental changes, so small fluctuations in the demand response does not affect the overall system stability.
5. **Adaptability:** Ability to change behavior when it is dictated by external factors, so the smart grid continuously adapts itself to the new economical and sociological requirements and working conditions.

In general a PSO algorithm is used mostly for optimization, candidate solution sets or working parameters for the smart grid population modelled according to the general behaviors. These candidate solution sets which are called particles in solution swarm, coexist and evolve simultaneously based on knowledge sharing with neighboring particles. Each particle will be either be acting individually or also in coordination with the other particles. Onsite storing capacity of the households, scheduled working hours for the machines in the industry for batch processes, charge location of each car within the city can be some example uses of this particle solutions. So each particle generates a solution based on the problem statement and initial starting point. In each turn all the particles changes their location, parameters, settings by using a velocity vector to reach its next position in the solution space. Each particle changes its velocity in order to reach a better position than the previous one by using its own best previous experience data and also the best previous experience data from the neighbors within the grid. Particles update their velocities as shown in equations below:

$$v_{t+1}^{i} = w_t v_t^{i} + c_1 r_1 \left(p_t^{i} - x_t^{i} \right) + c_2 r_2 \left(p_t^{g} - x_t^{i} \right)$$

and the new positions are calculated according to

$$x_{t+1}^{i} = x_t^{i} + v_{t+1}^{i}$$

where x_t^i represents the current position of particle i in solution space and subscript t indicates an iteration, p_t^i is the best-found position of particle i up to iteration count t and represents the cognitive contribution to the search velocity v. Each component of v_t^i can be limited to the range $[-v_{max}, v_{max}]$ to control excessive roaming of particles within the smart grid and creating possible risks of malfunction and system stability. p_t^g is the global best-found position among all particles in the smart grid swarm

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

up to iteration count t and forms the smart grid information flow to the velocity vector; r_1 and r_2 are random numbers uniformly distributed in the interval (0,1), while c_1 and c_2 are the cognitive and social scaling parameters depending on the structure of the grid and demand side policies. Finally w_t is the particle inertia, which is reduced dynamically to decrease the search area in a gradual fashion. The variable w_t is updated as;

$$w_t = \left(w_{max} - w_{min}\right) * \frac{t_{max} - t}{t_{max}} + w_{min}$$

where, $w_{\max}$ and $w_{\min}$ denote the maximum and minimum of w_t respectively; $t_{\max}$ is a given number of maximum iterations.

So the smart grid devices repeated the communication and changes their parameters until a time limit or a stopping criteria values reached.

Finally, the general PSO algorithm that can be used for smart grid applications is as follows:

```
For each device in the smart grid
   Initialize device with working parameters
End
While iteration number < maximum number of iterations or an acceptable optimum
reached Do
   For each device
   Calculate fitness value for better
   If the fitness value is better than the best fitness value (p_Best) in history
set current value as the new p_Best
   End
   Choose the device with the best fitness value of all the devices as the g_Best
   For each device
      Calculate device adapting velocity
      Update device working parameters
   End
End
```

Genetic Algorithms (GA)

Genetic algorithm (GA), is the most widely used version of evolutionary algorithms among the meta-heuristic methods. Similar to the inspiration from the nature as well as ANN and PSO, genetic algorithms are used to structure of evolution in order to reach the perfect solution that fits in to the energy environment. The possible working parameters of a particular device in the smart grid for demand side management optimization or scheduling problem taken as a member of the solution domain. As in the nature, every individual member device has set of properties and working parameters which differs from rest of the devices in the grid.

These properties and working parameters corresponds the chromosomes or genotypes similar to the ones in the living organisms (John, 1992).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

As the nature evolves in the time domain with new generation reproduced or composed of their parents or ancestors genotypes by crossover and mutation. Each new next generation member has a possible chance to better fit to its environment. The rules of genetic algorithms can be adapted to optimization problems in the smart grid and demand side management problems where the parameters for the devices or the pricing policies coded as chromosomes and the overall objective function for the grid optimization and scheduling considered to be the fitness of the member to the system. Most of the numerical adaptations of the problems numerical values decoded as binary strings so easy to make different levels and sources of crossovers and mutations (Kumar & Mohan, 2011). The common steps of genetic algorithm is as follows arranged for smart grid applications as follows

1. **Representations:** Proper coding of the smart grid optimization and scheduling in to mathematical chromosomes while the decision variables within a solution (chromosome) are genes. The possible values of variables (genes) are the alleles and the position of an element (gene) within a chromosome is named locus. Most common representation is based on bits. A simple example can be stated as the working status of any generator or electric device can be represented as 1/0 as on and off during any particular time or condition. Binary, discrete and permutation representations are available in the literature.
2. **Population Initialization:** Using the representation coding an initial state of the smart grid devices and their settings created and represented as defined on the first step. As to continue from the previous step, initial population contains some power stations are operational and some are not in a particular initial timing.
3. **Objective Function:** Proper coding of the grid optimization problem in to mathematical chromosomes which will be used to calculate the goodness of the solution.
4. **Selection Strategy:** The best parameter sets or working plans which fits into the objective function selected for the next round of reproduction. Assume there are 32 different working schedule options to start with as a power generation or consumption. Some of the options may generate a peak demand in the system whereas the others do not. So the selection strategy will be based on the objective function and the ability to conform to working parameters and schedule to the system. Selection strategy can be based on proportional conformance (fitness) or rank-based fitness. Since the capacity and parameters within the smart grid devices are different from each other depending on the location of use, industry and priority proportional fitness based selection gives better results. In the rank based selections, without looking at the proportions, best solutions are selected for the reproduction. Thus, considering the demand and supply capacities, it is possible to select huge power stations together with micro generation stations. Roulette Wheel Selection, Stochastic Universal Sampling, Tournament Selection, and Rank-Based Selection are common selection methodologies.
5. **Reproduction:** Using the current solution sets, they will produce the next generation alternatives by using;
    a. **Crossover:** Crossover is used to increase the diversity of the possible solutions, in the selection of energy stations, operating status of the stations replaced with other stations. This can be useful when comparing the different sources of energy and switching from hydroelectric power plants to natural gas power plant in some parts of the smart grid or vice versa. It will also be possible to include small reservoir or pumped hydro stations optimization within the grid.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

b. **Mutations:** New genetic material into the gene at some minor ratio. With a small probability, randomly chosen bits of the offspring genotypes change from '0' to '1' and vice versa for binary coded strings. Moreover to the switching the working status of the power stations, a mutation starts or stops a power station after the crossover applied. This will increase the possibility of the new searches in the domain.

6. **Replacement Strategy:** The best solutions which fit in to the smart grid objective function are selected for the next round of reproduction
7. **Stopping Criteria:** The best solutions which fit in to the smart grid objective function selected are for the next round of reproduction

Like most of the other artificial optimization methods, the following genetic algorithm can be used for complex problems.

```
Choose an initial grid operating parameters by means of chromosomes;
While termination condition not satisfied Do
   Repeat
      If crossover condition satisfied Then
            select parent chromosomes
            choose crossover parameters
            perform crossover
      End if
      If mutation condition satisfied Then
            choose mutation points
            perform mutation
      End if
      evaluate fitness of offspring
   Until sufficient offspring created;
select new grid operating parameters;
End While
```

Ant Colony Algorithm (ACO)

Ant Colony Optimization (ACO) is another swarm intelligence based metaheuristic algorithm that is inspired by the pheromone trail laying and tracking of the remaining of ants while finding their way to food or home. Use of artificial ants in ACO for smart grid applications are stochastic solutions that build possible candidate solutions for the optimization settings and parameters of the smart grid by adapting the artificial pheromone in other words fitness information. ACO is mostly used for discrete optimization problems for scheduling, routing and assignment of the resources in the smart grid. There are various ACO algorithms available like Ant System (AS), Elitist Ant System (EAS), Rank-Based Ant System (ASrank), Min-Max Ant System (MMAS), Ant Colony System (ACS), Approximate Nondeterministic Tree Search (ANTS). The real ants uses stigmergy which is a form of indirect communication by modifying the environment for further usage or communication. In case of the ACO, stigmergy based on the pheromone trails leaved by the ants while passing through a path or way.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

While adapting the ACO to the smart grid optimization ants will be artificial computational agents for the selection of smart devices in the grid operational. Traces or the pheromone levels considered to be the information within the smart grid like the consumption, prices. The memory for the grid is the tabu list of the possible solutions obtained within the system. In case of adapting the travelling salesman problem like selecting the order of the cities during the sales smart grids equipment operational timings aiming dynamic demand mode can be modelled. The ACO algorithm modified for smart grid is as follows:

```
Initialize the base efficiency and fitness (attractiveness), τ, and visibil-
ity, η, for each solution set;
While termination condition not satisfied Do
   For each smart agent Do
      choose probabilistically (based on previous equation) the next state to
move into
      add that move to the tabu list for each smart agent
      repeat until each smart agent completed a solution for dynamic response
   End
   For each smart agent that completed a solution Do
      update energy efficiency and fitness τ for each edge that the smart
agent traversed;
   End
   If (local best solution better than global solution)
      save local best solution as global solution;
   End
End
```

SOLUTIONS AND RECOMMENDATIONS

Linear programming, integer programming, dynamic programming, stochastic programming and many other methods can be used as traditional methods for smart grid applications and there a vast number of different examples available in the literature. Use of artificial intelligence and metaheuristic methods in smart grid and demand side management applications is summarized as in Table 2 according to the main controllable aspects of the issue. For smart grid applications in order to improve the benefits a good load balancing and forecasting required with a proper resource and production planning. Usage matrix summarized potentials of the methods and algorithms as M: Modelling, F: Forecasting, O: Optimization, S: Scheduling.

FUTURE RESEARCH DIRECTIONS

As the global awareness for the global warming and climate change increase, the idea of smart grids and demand side management will gain more significance and related programs will be urgently needed. Both the smart grid and demand side management applications means interaction with a vast number of instruments, devices, sensors, computers, power lines, transmission lines, distribution lines, new

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Methods and algorithms usage matrix

	Smart Grid				Demand Side Management	
	Load Balancing	**Load Forecasting**	**Microgeneration and Storage**	**Resource Allocation Scenarios**	**Dynamic Demand**	**Demand Pricing**
ANN Supervised	M,F	M,F		M,F	M,F	M,F
ANN Unsupervised						M,F
ANN Reinforced						M,F
PSO	O	O,S	O,S	O,S	O	O
GA	O	O	O	O		O
ACO	O,S				O	O,S
FL	M	M	M	M		M

technologies and smart equipment like IoT. It is quite clear that a synchronization and harmony between these technologies is a big challenge for engineers. And making it all possible and fully operational at once is a chore that is easier said than done. However, the concept of smart grid is fast evolving, device by device, customer to customer, node by node for the very near future; thus creating new research venues for the researcher.

Future research opportunities by the authors within the domain of the topic includes the analysis of the other metaheuristic methods and artificial algorithms like memetic algorithms, multi start methods, tabu search, guided local search, iterated local search, reactive search, stochastic search, simulated annealing and hybrid metaheuristics for better and alternative optimization, modelling, forecast and scheduling.

CONCLUSION

Demand for clean energy is becoming more vital and critical in every single day. Smart grids and demand side management tools are the most promising alternatives for successfully achieving energy related scenarios like INDC, bridge and 450 ppm scenario in order to keep the global warming level at or below 2°C. Moreover the research and development activities for new technologies, low emission processes demands efforts be focused on modelling, forecasting, optimization and scheduling of appropriate bodies and devices that can be used in smart grids. In this chapter both artificial intelligence and metaheuristic methods and algorithms are summarized and their possible application areas are identified. As expected artificial intelligence is more suitable for modelling and forecasting, whereas metaheuristic methods are better for searching optimum solutions and schedules.

REFERENCES

Chan, S. C., Tsui, K. M., Wu, H. C., Yunhe, H., Yik-Chung, W., & Wu, F. F. (2012). Load/Price Forecasting and Managing Demand Response for Smart Grids: Methodologies and Challenges. *Signal Processing Magazine, IEEE*, *29*(5), 68–85. doi:10.1109/MSP.2012.2186531

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dincer, I. (2002). The role of exergy in energy policy making. *Energy Policy*, *30*(2), 137–149. doi:10.1016/S0301-4215(01)00079-9

Gellings, C. W. (2009). *The smart grid: enabling energy efficiency and demand response*. The Fairmont Press, Inc.

IEA. (2006). Energy Technology Perspectives. In *Support of the G8 Plan of Action: Scenarios & Strategies to 2050*. Organization for Economic Cooperation & Devel.

IEA. (2015a). *Energy Technology Perspectives 2015: Mobilising Innovation to Accelerate Climate Action*. Organization for Economic Cooperation.

IEA. (2015b). *World Energy Outlook Special Report 2015: Energy and Climate Change*. Organization for Economic Cooperation.

John, H. H. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT Press.

Kumar, V. S., & Mohan, M. R. (2011). A genetic algorithm solution to the optimal short-term hydrothermal scheduling. *International Journal of Electrical Power & Energy Systems*, *33*(4), 827–835. doi:10.1016/j.ijepes.2010.11.008

Locke, G., & Gallagher, P. D. (2012). NIST framework and roadmap for smart grid interoperability standards, release 2.0. National Institute of Standards and Technology, 33.

Millonas, M. M. (1992). *Swarms, phase transitions, and collective intelligence*. Los Alamos National Lab.

Sioshansi, F. P. (2011). *Smart grid: integrating renewable, distributed & efficient energy*. Academic Press.

Strbac, G. (2008). Demand side management: Benefits and challenges. *Energy Policy*, *36*(12), 4419–4426. doi:10.1016/j.enpol.2008.09.030

Von Dollen, D. (2009). Report to NIST on the smart grid interoperability standards roadmap. Electric Power Research Institute for NIST.

Yegnanarayana, B. (2009). *Artificial neural networks*. PHI Learning Pvt. Ltd.

ADDITIONAL READING

Atzeni, I., Ordóñez, L. G., Scutari, G., Palomar, D. P., & Fonollosa, J. R. (2013). Demand-side management via distributed energy generation and storage optimization. Smart Grid. *IEEE Transactions on*, *4*(2), 866–876.

Awad, I. A. (2013). Smart Grid Applications.

Brown, R. E. (2008, July). Impact of smart grid on distribution system design. In Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century, 2008 IEEE (pp. 1-4). IEEE. doi:10.1109/PES.2008.4596843

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Budka, K. C., Deshpande, J. G., & Thottan, M. (2014). Smart Grid Applications. In Communication Networks for Smart Grids (pp. 111-147). Springer London. doi:10.1007/978-1-4471-6302-2_5

Farhangi, H. (2010). The path of the smart grid. *Power and Energy Magazine, IEEE*, *8*(1), 18–28. doi:10.1109/MPE.2009.934876

Gellings, C. W., & Chamberlin, J. H. (1987). Demand-side management: concepts and methods

Güngör, V. C., Sahin, D., Kocak, T., Ergüt, S., Buccella, C., Cecati, C., & Hancke, G. P. (2011). Smart grid technologies: communication technologies and standards. Industrial informatics, IEEE transactions on, 7(4), 529-539

Ipakchi, A., & Albuyeh, F. (2009). Grid of the future. *Power and Energy Magazine, IEEE*, *7*(2), 52–62. doi:10.1109/MPE.2008.931384

Keyhani, A. (2011). *Design of smart power grid renewable energy systems*. John Wiley & Sons.

Mohd, A., Ortjohann, E., Schmelter, A., Hamsic, N., & Morton, D. (2008, June). Challenges in integrating distributed energy storage systems into future smart grid. In Industrial Electronics, 2008. ISIE 2008. IEEE International Symposium on (pp. 1627-1632). IEEE. 10.1109/ISIE.2008.4676896

Palensky, P., & Dietrich, D. (2011). Demand side management: Demand response, intelligent energy systems, and smart loads. Industrial Informatics. *IEEE Transactions on*, *7*(3), 381–388.

Potter, C. W., Archambault, A., & Westrick, K. (2009, March). Building a smarter smart grid through better renewable energy information. In Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES (pp. 1-5). IEEE. doi:10.1109/PSCE.2009.4840110

Ramchurn, S. D., Vytelingum, P., Rogers, A., & Jennings, N. (2011, May). Agent-based control for decentralised demand side management in the smart grid. In The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 1 (pp. 5-12). International Foundation for Autonomous Agents and Multiagent Systems.

Rusitschka, S., Eger, K., & Gerdes, C. (2010, October). Smart grid data cloud: A model for utilizing cloud computing in the smart grid domain. In Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on (pp. 483-488). IEEE. 10.1109/SMARTGRID.2010.5622089

Siano, P. (2014). Demand response and smart grids—A survey. *Renewable & Sustainable Energy Reviews*, *30*, 461–478. doi:10.1016/j.rser.2013.10.022

Strbac, G. (2008). Demand side management: Benefits and challenges. *Energy Policy*, *36*(12), 4419–4426. doi:10.1016/j.enpol.2008.09.030

KEY TERMS AND DEFINITIONS

Cloud Computing: A new way of obtaining the information technology services using as a service pay as you use instead of buying.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Distributed Generation: Contrary of centralized generation where energy generation by district or decentralized stations and possible storage in the locations.

Information Network: A digital communication network used for transfer of useful data between the smart grid devices.

Markets: Economic market structure within the grid formed by retailers, aggregators, wholesalers and energy market clearance house.

Metering: The method for calculating and measurement of the device consumption.

Operators: Transmission, distribution, independent system operators, regional transmission operators acting in the grid.

Service Providers: Utility and third party providers for billing and customer information system in the grid.

Smart Grid Devices: Any intelligent device in the smart grid can able to perform basic computations and communicate within the information network.

This research was previously published in Smart Grid as a Solution for Renewable and Efficient Energy edited by Ayaz Ahmad and Naveed Ul Hassan; pages 49-68; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 9
Society of Agents: A Framework for Multi-Agent Collaborative Problem Solving

Steven Walczak
University of South Florida, USA

ABSTRACT

The development of multiple agent systems faces many challenges, including agent coordination and collaboration on tasks. Minsky's The Society of Mind provides a conceptual view for addressing these multi-agent system problems. A new classification ontology is introduced for comparing multi-agent systems. Next, a new framework called the Society of Agents is developed from Minsky's conceptual foundation. A Society of Agents framework-based problem-solving and a Game Society is developed and applied to the domain of single player logic puzzles and two player games. The Game Society solved 100% of presented Sudoku and Kakuro problems and never lost a tic-tac-toe game. The advantage of the Society of Agents approach is the efficient re-utilization of agents across multiple independent game domain problems and a centralized problem-solving architecture with efficient cross-agent information sharing.

INTRODUCTION

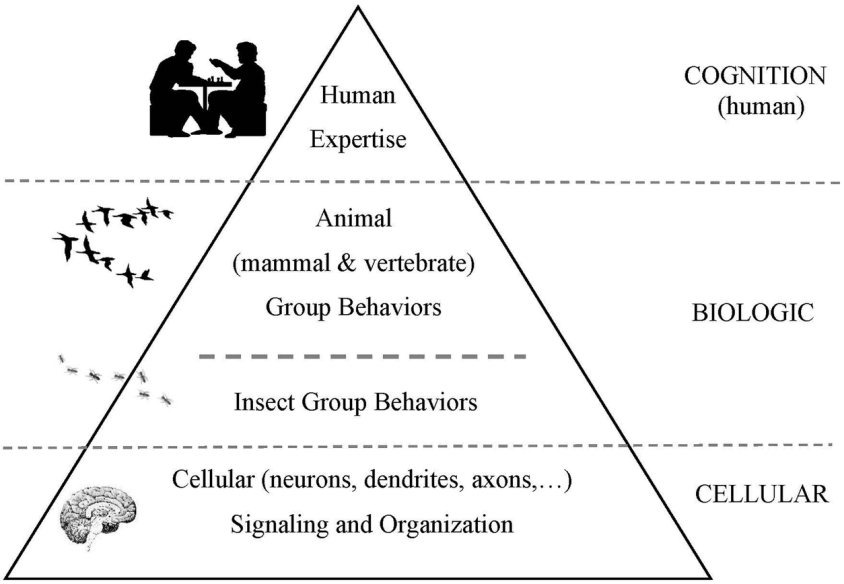
Distributed semi-autonomous intelligent problem solving is the function of multi-agent systems. Ontologisms or formalisms (frameworks) for development of multi-agent systems are recognized as a critical element for further development of multi-agent research (Chen & Chen, 2008). A natural classification of intelligent multi-agent problem solving strategies is implicit in nature. The highest level of self-contained intelligence is reflected in agents that model human intelligent problem solving by embedding evaluation heuristics based on human expertise within the agents, with examples coming from Othello (Samothrakis et al., 2013), a.k.a. Reversi, and checkers (Neto et al., 2014), a.k.a. draughts. The next step down are systems that mimic other biologic systems and include ant colony and beehive optimization agents (Madureira, 2014) among a vast number of other biologically inspired collaboration mechanisms for problem solving. The final classification layer is based on how intelligence develops at the cellular

DOI: 10.4018/978-1-7998-0951-7.ch009

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Intelligence and collaboration tradeoff in multi-agent classification



level and includes artificial neural networks (ANNs) embedded in agents, such as the ANN based agents for playing a specific type of poker (Nicolai & Hilderman, 2009). This agent and multi-agent ontology is shown in Figure 1. The ontology facilitates classification and comparison between agent systems.

As agent systems progress upward in the classification triangle shown in Figure 1, the autonomous intelligence of the system increases, which causes a corresponding decrease in the number of agents required to solve a general problem in the agent's domain. The agent classification triangle in Figure 1 classifies agents based on their relative position within the triangle with increasing knowledge-based problem-solving capability corresponding to the height in the triangle and the width of the triangle indicating the necessary quantity of collaborating agents required for solving problems within their domain. This may be understood in plain language by looking at the number of living agents or corresponding units involved in problem solving for the different communities. A single human can solve a chess endgame problem, whereas animal communities solve survival challenges in groups of several animal agents to several dozen and insect colonies typically manage survival problems using hundreds to thousands of individuals. The human brain contains millions to billions of neurons (West & Gundersen, 1990), which are the building blocks of consciousness and intelligence. Thus, as the foundational model of agents within a multi-agent system moves down the pyramid from cognitive to biologic models and then from biologic to cellular, a greater number of agents are required to collaborate to produce intelligent behaviors.

Although this classification provides a straightforward ontology for evaluating required knowledge and collaboration levels within multi-agent systems, it still lacks a formal framework to define how collaboration between agents functions. The research presented in this paper defines a framework for agent interactions and collaborations in cellular level multi-agent systems. The framework is founded upon Minsky's (1986) *Society of Mind* paradigm, which was never formally instantiated, and is therefore called the Society of Agents (SoA). The idea of utilizing a social perspective to facilitate agent interactions is not new and has been used in developing agent communications for almost two decades (Ferber, 1999;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Huhns & Stephens, 2000), but the framework developed in this article moves the concept of society and social interactions from the biologic level down to the cellular level and provides a uniform way to organize agent societies to accomplish intelligent problem solving.

BACKGROUND

The Society of Mind

In Minsky's (1986) *The Society of Mind*, he proposes that human understanding, intelligence, and consequent problem-solving ability is derived from a large collection of semi-autonomous, non-intelligent, independent, and reactive agents. These propositions were presented as thought exercises and were never instantiated. The way in which Minsky uses the term agent is as an actor that is responsible for a given task. Each actor's task is atomic and may be accomplished by the actor alone or by invoking actions or information from other actors. The use of other actors may be viewed hierarchically, with the initial actor determining a task to be accomplished and then invoking other actors to accomplish subtasks that individually would not accomplish the main task, but together can achieve the desired goal of the initial actor. The actors handle their associated atomic tasks within the system and interact with each other to handle more complex tasks.

Society of Mind (Minsky, 1986) actors are atomic, meaning that each actor is responsible for a single and indivisible component of the overall systems. Some of these atomic actors will be responsible for performing actions, while others detect or maintain state. The individual actors are not meant to be viewed as intelligent, but rather any perceived intelligence of the system is an effect from the various actors interacting to accomplish the needs or tasks of the system. The various actors in Society of Mind are responsible for perception, reasoning, memory, action, and a few other anthropomorphic properties such as emotions and sense of self (Minsky, 1986). Therefore, any multi-agent framework derived from *The Society of Mind* (Minsky, 1986) must require that agents perform atomic behaviors and that these behaviors include perception, memory, and action.

Intelligence occurs through the aggregation of values and actions held by different sets of actors within the society. As such, these actors may be emulated utilizing current software agent technology. Minsky (1986, pg. 18) proposes several important questions that need to be answered to understand the Society of Mind actors: "how do agents work, what are they made of, how do they communicate, where do the first agents come from, how do we make new agents or change old ones, [and most importantly] how can groups of agents do what separate agents cannot do?" The SoA framework presented in this article seeks to provide answers to these questions for cellular agents and multi-agent systems.

A way to begin understanding how multi-agent software agents can answer the final question is to examine biologic-based systems. Although biologic agents are not at the same granularity as the agents in the Society of Mind and the SoA, most animals working in groups or a society can accomplish tasks that are difficult or impossible for individuals to perform.

Biologic Agents

Natural biological systems have inspired a variety of optimization algorithms and agent systems. Biologic-inspired systems (a.k.a. bio-inspired or nature-inspired) may provide ideal solutions to problems that are complex or nonlinear and include either a large number of variables or solutions or both (Siddique &

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Adeli, 2015). The natural biologic systems that have inspired agent-based solutions include: frog leaping, ant colonies, particle swarm methods based on bird flocks and fish schools, bacteria, cuckoo (and other birds), bee hives, fireflies, and plant-oriented algorithms including flower pollination (Kar, 2016).

Biologic agent systems have been used to approach a variety of problems and a sample of this research is presented in Table 1. Since ant colony and particle swarm biologic inspired agent systems are the oldest, they naturally have the largest representation in literature. Biologic agents must adhere to social norms for communication processes between agents, whereas cellular agent communication occurs at a more rudimentary level, exemplified by electrical or chemical signals between neurons.

Cellular Agents

Minsky's (1986) *Society of Minds* attempted to intellectually bridge the gap and make an intellectual connection between traditional knowledge-based artificial intelligence and newer connectionist methods like ANNs that operate at the cellular level and are not considered to be knowledge-based (Reeke, 1991). ANNs and genetic algorithms are well known artificial intelligence cellular level methods.

An example of utilizing a combined biologic and cellular multi-agent system is a bat sonar signal-based ANN multi-agent system for predicting stock time series (Hafezi, Shahrabi, & Hadavandi, 2015). Multi-agent systems with embedded ANNs are well known in game domains, such as: backgammon (Kiourt & Kalles, 2015), checkers (Duarte et al., 2015); poker (Nicolai & Hilderman, 2009), and robot soccer (Jolly et al., 2007; Jolly, Kumar, & Vijayakumar, 2010).

Although genetic algorithms are not commonly used in agents, the approach can be used to fine tune agents that use other learning methods. Kiourt and Kalles (2015) demonstrate a social Darwinism approach to multi-agent system game playing agents such that the agent community eventually determines and selects the optimal game playing agent (i.e., survival of the fittest).

While ANNs and genetic algorithms attempt to model cellular level strategies for learning, the closest current agent systems come to actually modeling cellular behavior is in bacteria-based and human immune system based models, such as the multi-agent green recycling planning system (Mishra, Kumar, & Chan, 2012) and the control of smart grid power systems at the device level (Belkacemi et al., 2011)

Table 1. Examples of biologic system inspired agents

Reference	Bio Types	Problem/Domain
(Liu & Passino, 2006)	foraging swarms	robot sensing & path planning
(Saint Germain et al., 2007)	ant colony	supply chain management
(Grushin & Reggia, 2008)	particle swarm	block construction
(Xiang & Lee, 2008)	ant colony	manufacturing scheduling
(Król & Drożdżowski, 2010)	ant colony & particle swarm	traveling salesman
(Jabbarpour et al., 2014)	ant colony	traffic flow planning
(Liu et al., 2015)	bee colony	task assignment
(Kolomvatsos et al., 2016)	bee colony	e-commerce negotiation
(Byrski et al., 2017)	ant colony	traveling salesman
(Eydi, Panahi, & Kamalabadi, 2017)	ant colony	vehicle path planning

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

both based on the human immune system. There is a need to further develop cellular level multi-agent systems to better capture the unique group dynamics of autonomous cellular units in solving complex problems. Furthermore, development of cellular multi-agent systems would benefit from a framework to enable reproducible results across a wide range of domains.

Existing Agent Architectures and Frameworks

Any agent framework must account for the 3 C's: Collaboration, Communication, and Coordination. Successful heterogeneous agent systems require more than communication and collaboration for successful problem solving (Bellifemine, Poggi, & Rimassa, 2001) and this accounts for the coordination mechanism for assigning tasks to agents and creation of new agents when needed. Collaboration between agents enables a more holistic solution as agents share information and resources to solve problems beyond the ken of single or smaller groups of agents (Arsene, Dumitrache, & Mihu, 2015).

Several architectures or frameworks for software multi-agent systems already exist. The Common Object Request Broker Architecture (CORBA) is an example of an existing architecture framework that focuses primarily on enabling communication between agents (Object Management Group, 1995). CORBA was developed to satisfy a need to enable agents to exchange information without hard coding the information transfers into each agent participating in a multi-agent system.

The Open Agent Architecture (OAA) was developed at SRI International and provided an architecture for distributing or coordinating tasks among a collection of agents (Martin, Cheyer, & Moran, 1999). An extension of the OAA was developed in KAoS (Knowledgeable Agent-oriented System) for performing maintenance problems in industry (Bradshaw et al., 1997). The KAoS system operates in a biological manner by enabling task coordination and sub-goal amongst a community of agents and the cognitive level due to the possible independence of agents in solving complex domain problems.

The SoA framework utilizes ideas from these various architectures which are common to multi-agent systems, communication and task coordination. Each of these architectures specifies how to construct an agent that has multiple non-atomic capabilities. SoA extends these architectures into a cellular level framework which enables each agent to communicate if needed, but distributes the coordination effort to a specific class of agent, thus making it that specific agent's atomic responsibility. The full framework is described in the next section.

SOCIETY OF AGENTS (SoA) FRAMEWORK

The purpose of any framework is to provide a method for developing high quality agents consistently and for instantiations to be evaluated for consistency against the framework. The FIPA organization created a standard for developing heterogeneous agent systems and specifies a programming language for developing agents based on Java (Bellifemine, Poggi, & Rimassa, 2001; Bellifemine, Caire, & Greenwood, 2007). Although other programming languages are used, the common theme is to enable heterogeneous agent development (Russell & Yoon, 2009) and so the SoA framework utilizes Java as its programming language and XML as its data language, which is a common communication protocol for multi-agent systems (Singh, 2007). An instantiation of the SoA framework is known generically as a Society and

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

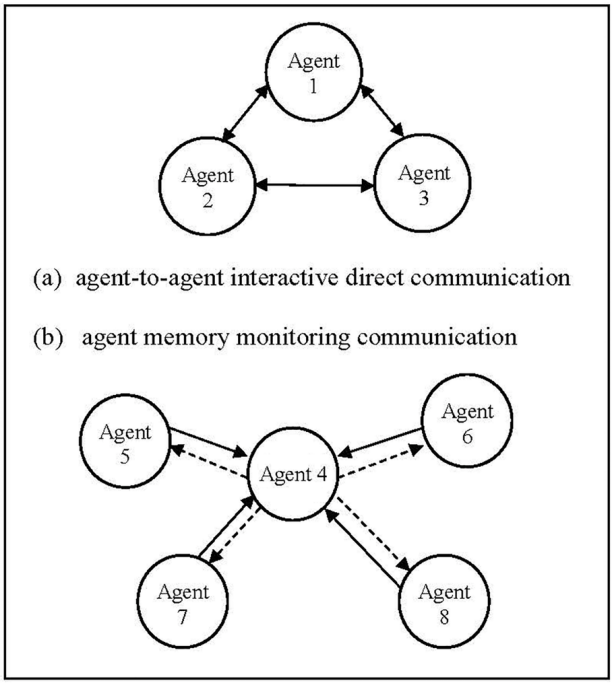
more specifically as either a Domain Society or Society of Domain, where Domain is replaced by the actual domain in which the Society is functioning. Hence for the domain of games, a Society created using the SoA framework would be known as a Games Society or alternately as a Society of Games.

Communication is critical in multi-agent systems (Lian et al., 2006) and the Society of Mind paradigm specifies that while an agent is performing its own atomic task it should not need to communicate with other agents, but when sub-tasking occurs then inter-agent communication is required (Hermans, 2002). The SoA framework enable two types of communication shown in Figure 2: decentralized or direct and centralized or indirect communications. The agent at the center of the indirect communications (agent 4 in Figure 2) may also serve as memory for the other agents and the agents' world. The SoA framework's communications mechanisms are similar to those used in CORBA.

Communication is also a requirement for interfacing with the various instantiated Society users and the SoA framework specifies the creation of input and output agents. The input and output agents are responsible for acquiring information and resources from the society's working environment and for communicating results back to the end users. The SoA framework thus extends and celullarizes the communication protocols of CORBA and KAoS by making communication both a general agent capability as well as the atomic task of specific types of agents.

Coordination affects how agents collaborate through task assignment and delegation. Agent frameworks require a coordination model, similar to OOA and KAoS, but which should also account for shared information spaces (Zambonelli et al., 2014). As with communication, coordination may be either centralized or decentralized (Wang, Zeng, & Cong, 2016). The SoA framework permits both types of coordination. The cellular nature of SoA requires that goals or tasks are allocated to distributed agents that only have knowledge of local tasks and resources (de Weerdt, Zhang, & Klos, 2012). Agent com-

Figure 2. SoA agent communication



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

munities within the larger agent society may organize and coordinate information sharing and tasks to accomplish subgoals of the SoA instantiated Society, thus enabling a decentralized focus. A required agent in the SoA supervises the sharing of information and coordination of agent activities globally, thus providing an overall centralized architecture. As an analogy, in American football, the coach is the centralized controller directing the overall goals and actions of the team (the society of agents), but the offensive line may organize at a lower level to perform blocking assignments for individual agents to promote the success of the whole team.

The SoA framework for multi-agent systems requires that:

- Agents operate at the atomic level, where each agent is responsible only for a single individual non-divisible task, specifically agents must be defined that will:
  - Read input from the domain for the current task (text, graphics, XML, voice, etc.);
  - Reason (problem solve) utilizing various problem-solving strategies, one per agent;
  - If needed, display information or requests for information (alternately this type of agent may also affect actions to occur within the domain such as robot movement);
  - Serve as a local short-term memory for other agents;
- Agents may be, and in fact should be, reused in multiple domain contexts as long as the atomic task of the agent is appropriate within the specified context;
- Agents act independently and thus have a widely distributed locus of control, however;
- Agents must interact with other agents in order to accomplish any task more complex than the atomic tasks for which the agents have been specified; and
- All agents within the society may be active, however it is possible that only a subset of the society may be required to interact in order to solve the current domain problem or task.

These are the fundamental properties of the SoA framework. The social perspective also defines multiple types of agent interaction relationships in multi-agent systems dependent on context (Kolp et al., 2008). The different types of agents defined within the SoA framework and their interactions are specified in Table 2. A UML domain model for the SoA framework is shown in Figure 3. Depending on the problem to be solved, agents with varying tasks/goals may be implemented in each of these classes of SoA agents, so long as they perform the specified purpose and can communicate for collaborative problem-solving activities.

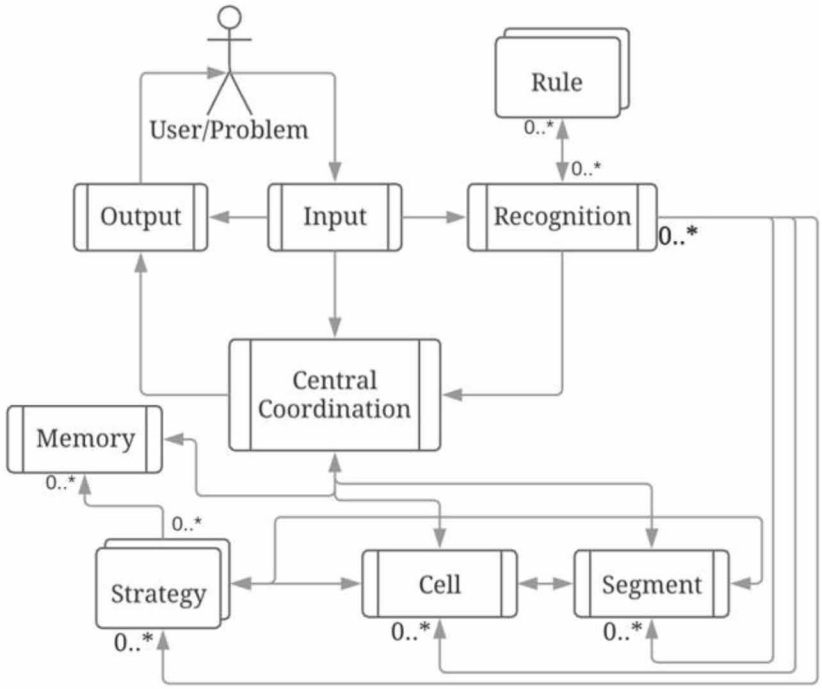
If the division of tasks into the atomic level units specified in Table 2 is followed, then no single agent should be able to solve a standard domain problem solving task individually. Agents must work collaboratively to accomplish tasks. Additionally, one or more Recognition agents operate external to the society and answer the questions where do the first agents come from and, how do we make new agents or change old ones? The Recognition agents have the task of identifying the environment and problem to be solved. These agents are persistent and active within the environment where the instantiated Society performs. The collection of types of problems a SoA Society can work on are pre-specified in the Recognition agents for each domain. The collection of cell and segment agents required to solve the current domain problem are implied through the Rule agents. Once a Recognition agent identifies a known problem type it then creates all of the appropriate agents for the new "Society" as specified by the Rule agents.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. SoA framework agent definitions

Agent	Purpose
Rule agents	Identifies constraints imposed by the environment on other agents and other domain specific information
Recognition agent	Determines the problem to be solved and builds an agent community
Central Coordination agent	Acts as short-term memory to maintain the current status/configuration of the problem environment and coordinates communication between agents
Cell agents	Identifies the contents of a single knowledge item that may be a part of a larger whole
Segment agents	Organizes collections of Cells and permits inter-Cell communication
Input agents	Acquires information from the environment or user/system
Output agents	Responsible for communicating agent results and queries with the user/system
Strategy agents	Implement various reasoning/problem solving strategies
Memory agents	Serve as long term memory

Figure 3. SoA framework UML domain model



Rule agents and Strategy agents serve as long term memory for the SoA societies. Rule agents define limitations or constraints that new society agents must operate under as well as general knowledge about the domain, such as the shape of the world (e.g., the configuration of a game board or puzzle). Strategy agents define generic problem-solving strategies that each individual agent, subgroup of agents, or the full society may employ in solving the problem defined through the Recognition agent and the Input agents, which gather information about this specific problem instance. Strategy agents model both inherent

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

strategies that are present at birth such as fear of following and pain avoidance and also model learned strategies such as fire causes pain. Whenever a new problem type with associated rules or constraints is identified, a new Rule agent is created, and these agents are persistent across SoA implementations. Similarly, if a new problem-solving strategy is identified, then a new Strategy agent is created and again this type of agent is persistent across all instantiations of SoA systems. Therefore, the Rule and Strategy agents serve as a long term memory for the SoA's societies. An additional Memory agent is specified to enable the long-term retention of information across multiple problems that are outside of what constitutes proper knowledge for the Rule and Strategy agents.

Each agent within an instantiated Society understands a subset of information needed for accomplishing domain tasks. Typical agent interaction will be in the form of agents requesting missing information from other agents or suggesting modifications to information currently held by other agents, all while seeking to accomplish their specified goal. This dynamic information processing and resolution is responsible for the intelligent problem-solving capabilities of a Society. Solutions to domain problems may be identified in multiple ways: 1) an agent is specified that has the atomic action of recognizing when a solution has occurred (this may be either a Cell agent or the Central Coordination agent), or 2) solutions may be identified when each active agent has refined its current information so that a unique value or solution is obtained and no further action is required of any agent to obtain their respective goals.

While the details of the SoA framework are specified above, the necessity and implementation of these requirements may be better understood through an example. The next section will examine developing a multi-agent SoA Society for playing games, known as the Games Society.

SOA SOCIETY FOR PUZZLES AND GAMES

The initial domain of application using the SoA framework is the domain of games. Games are chosen because they are well understood and have served as a traditional testing ground for artificial intelligence research (McCarthy, 1990). The types of games to be addressed by this SoA framework implementation is wide ranging and only requires that some type of internal representation of the game is possible for the agents to encapsulate and interact with and the game must be winnable or have a definite conclusion, such as a draw or tie with all game Cell agents occupied, so that the society of agents may determine when a solution is reached. This includes one-person, two-person, and eventually multi-person games that are both bounded and unbounded (indicating the lack of limits for the game board or playing area). The current implementation of the Games Society is capable of playing: Sudoku, Kakuro, tic-tac-toe (TTT), three dimensional tic-tac-toe (3D-TTT), and beginner level Caro, a.k.a. Gomoku (a game similar to TTT, but played on either a 15X15 or 19X19 board and requires 5 similar marks in a row for victory). Future research is needed to extend the Games Society to also play other games such as go, backgammon, and eventually checkers, chess, and shogi.

The current set of games solvable by the Games Society is chosen to incrementally increase the complexity of the games. Sudoku is a single player number puzzle with a consistent bounded board (9X9), while Kakuro is also a single player number puzzle, but introduces additional solution requirements for the agents and uses a bounded yet inconsistent board. Both Sudoku and Kakuro have a unique solution if the puzzle is properly configured. TTT introduces two person games and therefore introduces strategy

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

into a consistent bounded board. 3D-TTT incorporates and additional dimension that the reasoning, input, and output agents must address. Caro increases the overall complexity of the games and the time/distance/size requirements to reach a solution.

Evaluation of Games Society

Evaluation of the Games Society instantiation of the SoA framework is done by determining if the Society is solving a specific game problem at an advanced human player level. For Sudoku and Kakuro this implies the ability to solve arbitrarily difficult puzzles, while for TTT this implies never losing (since TTT will always result in a draw between two knowledgeable players) and winning against a less skilled player. The 3D-TTT game is winnable in four moves by the starting player and as such, if the Games Society is selected to start the game it should always win and play tactically well if going second. The evaluation requirements of Caro/Gomoku is similar to that for 3D-TTT, though the number of possible moves and world states for Caro is significantly larger, the first player is perceived to have an advantage (Tan et al., 2009).

The Game Society will also be compared to existing software agents or multi-agent systems that already operate against the specified games played by the Game Society. This comparison is facilitated by Sudoku becoming the drosophila for agent research because it is an NP-complete problem indicating a reasonable difficulty in solving and the well-defined and constrained nature of Sudoku problems (Kendall, Parkes, & Spoerer, 2008). Existing agents or multi-agent systems typically only try to solve one game or puzzle, which immediately shows an advantage of the SoA framework and its instantiation in the Game Society which currently plays five distinct games or puzzles within this domain.

Implementation of SoA Framework Agent Types

As described in the previous section, several different types of agents are required to form a SoA "society". The full SoA framework derived architecture for the Games Society is shown in Figure 4.

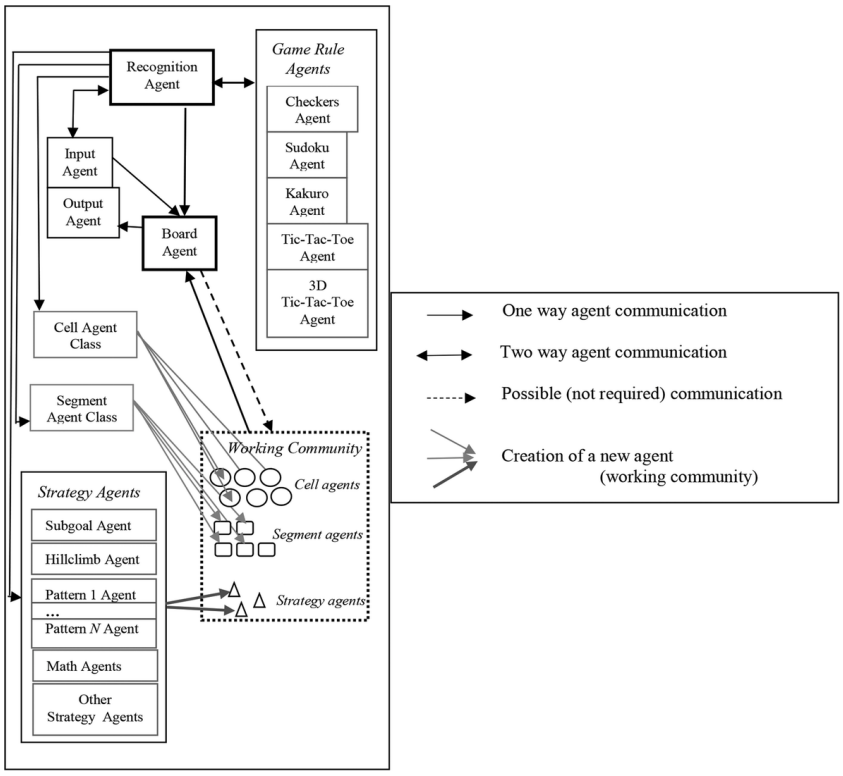
The SoA framework provides a dynamically configurable multi-agent architecture, which is adaptable to a variety of game domain problems. The Rule agents and the Strategy agents appear within their own boxes to indicate that these are persistent across instantiations, similar to a knowledge base for the SoA framework instantiated multi-agent systems.

Each game or puzzle type that the Games Society can play has an associated Rule agent defined. The Rule agents act as a long term memory in the Society, recording pertinent information applicable to each specific game, including the type and number of pieces, size and shape of the game board, starting configuration of the game board, any rules for the play of the game (e.g., for Sudoku and Kakuro this would include a rule that only one of any digit may occur in a Segment agent's collection of Cell agents) and how to determine a winner. Information about the games is kept within the Rule agents in XML format. When future research applies the Games Society to larger scale and longer games such as checkers and chess, the rules for movement of the pieces will also be required as elements of a Rule agent within the Games Society.

The required Recognition agent performs a centralized control to spawn all of the necessary agents needed by the current society. Once these other agents are created, they assume local control to play the game and thus creates a decentralized problem-solving architecture. The Recognition agent identifies the current type of game being played by looking at the game board or puzzle representation being

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Games Society architecture



presented to the Games Society. Once the Recognition agent determines the game being played, it then requests game information from the corresponding Rule agents to determine the quantity of cell and segment agents that are needed. The Recognition agent then creates these agents into a problem specific working community within the larger Games Society, and also creates the Central Coordination agent.

The Central Coordination agent in the Games Society is renamed the Board agent for ease of understanding, since it serves as a centralized short-term (length of this specific game or puzzle solution) memory location for all other agents and in fact represents the current state of the puzzle board or game board. The Board agent identifies which Cell agents correspond to each space on the puzzle or game board and therefore controls the collaboration of Cell and Segment agents. For most two or more player games, the Board agent has a pre-specified starting position and this information is communicated from the Rule agent to the Board agent via the Recognition agent. The current state of the Board agent and hence the puzzle or game is communicated to an Output agent and displayed to the user. The two one player numeric puzzle games currently solved by the Games Society must have the starting position specified either by a user to an Input agent or may be read in from a file by an Input agent. This means that some of the Cell agents have a value pre-determined for them and their task is finished within the larger Games Society problem instance.

Cell agents represent the individual spaces within a puzzle or game board and may therefore each only have a single value representing the current state of that space within the puzzle or on the board. This value may simply be occupied or unoccupied by a specific player or may contain a value for the corresponding location of the Cell agent for the Board agent (e.g., a number from 1 to 9 for Sudoku and Kakuro, or a piece value for games like chess and shogi).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Java code for dynamic agents

```
url= new java.net.URL[] {new java.net.URL("file:"+ClassName+".java")};
java.net.URLClassLoader loader = new java.net.URLClassLoader(url );
Class cls = loader.loadClass(ClassName);
r = (java.lang.Object)cls.getMethod(MethodName,x). invoke(cls.newInstance(),o);
```

As shown in Table 2, the Segment agents form collections of Cell agents that have meaning for the game, as specified by the Rule agents. For example, in Sudoku, Segment agents organize different collections of Cell agents to form rows, columns, and squares for a total of 27 Segment agents. However, the TTT game only requires 8 Segment agents: 3 rows, 3 columns, and 2 diagonals. Note that a Cell agent may simultaneously belong to multiple Segment agent groups so long as this does not violate the Rule agents' constraints for the current environment/game.

Each Cell and Segment agent is created with appropriate Game Rule knowledge/constraints (e.g., Cell agents know the types of values they can hold and if multiple values are permitted). These agents are created dynamically by the Recognition agent as needed using Java code, displayed in Figure 5.

The Input and Output agents are responsible for all communication with users and other systems outside of the Society. The Games Society's Input agents read the puzzle or game board and acquires moves of the opponent in two-player games, while the Output agent displays the current actions taken by the Games Society's agents as represented in the Board agent. Flexibility and adaptability are important features in multi-agent systems (Barber, Goel, & Martin, 2000; Wang & Shen, 2016) and the dynamic creation of these working communities of cells and segments provides this in the SoA framework.

The Recognition agent also loads the Games Society with any Strategy agents that might apply. The Segment and Strategy agents collaborating together create problem solving intelligence for the society. The Segment agents utilize game knowledge (from the Rule agents) to organize the Cell agents and to assist Cell agents in identifying unique or desirable piece values to hold. Strategy agents apply standard problem-solving strategies including:

- Computation (in which the problem is directly solvable utilizing the rules of the game from the corresponding Rule agent);
- Subgoaling or simplification, which breaks the problem into smaller size chunks;
- Pattern recognition for important patterns specified by the Rule agents;
- Hill climbing, which tries to reduce the difference between the current game state held in the Board agent and the desired outcome;
- Analogy to similar problem situations;
- Working backwards, which attempts to move from the goal state backwards to the current state;
- Heuristics, utilizing heuristic problem-solving strategies for specific domains (such as controlling the corners in TTT);

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Mathematics, operation of standard math operators including addition, subtraction, multiplication, division, and exponentiation; and
- Additional agents to assist in discovering numeric patterns and allowing pattern rotation equivalences.

New Strategy agents may be added at any time to a Society to increase the Society's reasoning capabilities. Examples of possible new Strategy agents for future research include: Bayesian reasoning, other probability and statistical operators, and greedy algorithms. While Strategy agents are currently loaded into the working community of Segment and Cell agents working on the current problem, future research is needed to investigate the Board agent realizing that a new unloaded Strategy agent might be needed, perhaps because progress is not being made towards the goal of the Society, and make available other Strategy agents dynamically to the working community.

Previous research has indicated that intelligence in games is usually tightly coupled with the game system (Yu, Shen, & Miao, 2008). The Strategy agent type within the SoA framework provides a mechanism to decouple problem solving intelligence from the multi-agent system and affords a high degree of re-usability across new SoA framework Societies.

RESULTS AND DISCUSSION

The Games Society described in the previous section implements the SoA framework for a multi-agent system in the domain of games. Analysis of the Games Society's performance is accomplished by examining the actual performance of the Games Society across the two logic puzzles and 3 two player games previously mentioned.

Recall that both Segment and Cell agents inherit required domain specific information for a particular game from corresponding Game Rule agents. Simpler games such as Sudoku may be solved using just the Cell and Segment agents by having the Segment agents recognize intersections between the various Segment agents and guide the Cell agents in producing a solution by performing problem space reduction through constraint propagation to help the Cell agents eliminate possible values from consideration. Kakuro works similarly, but does utilize mathematical Strategy agents that are used to identify patterns of numbers that may fit within a specified segment of cells within the Games Society, allowing instantiation of the SoA framework to understand and create non-specified constraints. An example of the use of the Addition Strategy agent is to determine all combinations of N unique (non-repeated) digits with values between 1 and 9 that produce a sum equal to a value X. Similar to Sudoku, the Segment agents then use the values identified by the Addition Strategy agent to identify possible Cell agent values that satisfy the requirements of all the intersecting Segment agents in which they participate.

The TTT and 3-DTTT games could play with just the Segment and Cell agents loaded with Game Rule knowledge, however this play would not be considered intelligent as the moves would be at random except for win blocking and winning moves since these types of moves may be inferred from the Rule agents for TTT and 3D-TTT. These games introduce the requirement for additional Strategy agents to exist in more complex game problem solving tasks. Defensive and offensive heuristics concerning N in a row patterns played on a matrix-like board are commonly known and have been coded as pattern heuristic Strategy agents. The heuristics used in TTT are also re-used in 3D-TTT with one additional

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

heuristic Strategy agent to account for three dimensional patterns. Caro also re-utilizes the various TTT heuristic agents and similar Strategy agent re-use is expected in other pattern-oriented games such as: Othello/reversi and Go.

Clearly, the list of strategies currently implemented for the Games Society given above will require further additions to utilize expert strategies from and for specific games, such as utilizing multiple pieces to defend or attack a Cell or Segment for games like chess and shogi and game specific strategies such as controlling the central area of the Board agent in chess versus controlling the edges of the Board agent in checkers. Segment agents that capture and organize the collection of Cell agents which define different regions of the Board agent, would be required by these new Strategy agents, in each newly defined game.

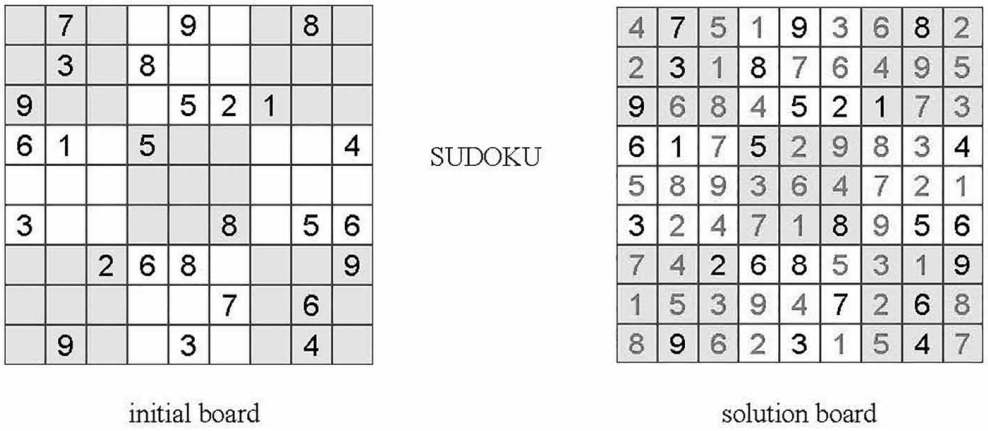
The atomicity requirement for the agents is achieved for all of the specified agents, by making sure that their task and knowledge is specific to a single well defined indivisible task, with the exception of the Recognition agent. Currently the Recognition agent performs two tasks: identification of the domain with transfer of appropriate Rule agent information to other agents and also the creation of the problem solving community. The decision to create the community with the same agent was made for simplicity in programming and could be divided into separate atomic agents in future iterations, where the Recognition agent would solely be responsible for identifying the game or environment and would then pass this information to a Creation agent that would be responsible for the deployment of the working community within the SoA framework instantiation.

Since the Games Society currently solves four different games, with work progressing on a fifth game, each with a unique set of properties, the Games Society clearly enables agents to be re-used across multiple tasks within the domain. Specifically, the Recognition agent, Board agent, Input and Output agents, Cell agents, and Segment agents will be used for every game. This is possible with the use of the Rule agents, whose purpose is to identify environment specific values for individual problems to allow the dynamic reconfiguration of the Board and Segment agents and automatic learning of appropriate Cell content and Output characteristics. Strategy agents are re-used across all games that share specific properties, such as using the math Strategy agents in all games that contain numeric values (i.e., the game pieces) within the Cell agents. Various pattern Strategy agents are used in all of the games currently implemented and help to recognize specific patterns and partially completed patterns. The heuristic strategies are added in response to a performance need within a specific domain game problem, however these may find re-use in other game problems that share similar heuristics (e.g., TTT has a defensive heuristic to control the center of the board and a similar strategy is viable in chess). Thus, the SoA framework as exemplified in the Games Society instantiation enables a successful collaborative multi-agent system.

The final two C requirements of the SoA framework concern locus of coordination control (centralized or decentralized) and agent communication. The current Recognition agent may be viewed as having centralized control authority, however the locus of control is then decentralized and passed to a subset of agents which form the active working community within the society. The different forms of agent communication have been listed and in fact, the atomic nature of the agents requires them to interact and communicate with each other in order to be able to collaboratively solve the various domain tasks. This interaction is especially important between the Board agent and the Segment and Cell agents. The Segment and Cell agents each know very isolated and specific domain task (game) information which when interactively combined enables the resolution of the defined agent system task, as identified by the Recognition agent.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Sample Games Society Sudoku boards



The Games Society currently works on the five games previously mentioned. The Games Society has perfect performance for the domain problem of Sudoku, solving 100 percent of the Sudoku problems given to the Society on puzzles rated from easy (1 star) to difficult (5 star). Sample Sudoku problems were taken from local newspaper puzzles and various open websites providing sample puzzles. A sample beginning board for the Board agent and a Games Society completed board for the Board agent as actually printed by the Output agent (with shading on both boards added for clarity in perceiving the Sudoku puzzle) are shown Figure 6.

There are a large number of existing single agent or multi-agent systems that currently play Sudoku. A small sample of six systems are presented here for comparison. The Game Society Recognition Agent creates a working community within the Society of 81 Cell agents and 27 Segment agents which are supervised by one Board agent and utilizes 3 Strategy agents (single missing value, pattern rotation, and implied future value):

1. Operating at the cognitive agent level, a multi-agent system of 2 agents utilizing backtracking and exhaustive search and other multi-agent systems with 9 agents utilizing sub-tasking and constraint satisfaction built into the agents attempts to solve Sudoku problems with the 2 agent system achieving 100 percent completion and the remaining three systems of 9 agents having successful completion percentages of 0%, 0%, and 44% (Abednego & Nugraheni, 2014). The first system compares favorably against the Game Society with regard to performance. Constraints in the Game Society are specified in Rule agents and enforced by the Board agent through the Segment agents, whereas the constraints are directly programmed into the agents in the 9 agent systems, demonstrating the difficulty of trying to encapsulate all domain knowledge within single agents;
2. Also at the cognitive level is a reported multi-agent system that achieves a 100% solution accuracy in Sudoku and uses 82 agents, representing the 81 cells in a Sudoku puzzle and a central controlling agent (Emery, 2007). This system also uses rollback and exhaustive search and introduces the concept of guessing a cell value when the multi0agent system gets stuck in an unsolved position. Each agent has its own blackboard for communicating with other agents, where the blackboard memory function is handled by the Board agent in the Gam Society. The agents are programmed using a Java-based distributed system agent architecture called Cougaar (COGnitive Agent Architecture), using the same programming language as the Game Society;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

3. The next comparison creates individual agents that attempt to solve Sudoku and are organized into collectives that are called a "society" (Leu, Tang, & Abbass, 2014). Each individual agent has both memory for awareness of the current state of the Sudoku problem and skills for determining cell focus and number selection. Therefore, each individual agent which is performing the solution task is a cognitive agent, but the collection of agents which attempt collaboration to identify more useful skills may be viewed as a pseudo-biological/societal solution. At easier difficulty levels this collective of agents is able to solve all the Sudoku problems, but as difficulty increase the solution accuracy drops dramatically. The Game Society was able to solve all Sudoku problems presented from easy to very hard difficulty levels;
4. Also at the cognitive level, but showing some similarities to the cellular Game Society is a multi-agent system that employs numerous agents each with a different pre-defined solution strategy or combination of strategies (Agerbeck & Hansen, 2008). This method partially utilizes the idea of Strategy agents used in the SoA framework, but is cognitive because strategies may be combined in single agents and the agents search the board to determine if they can be used at any given location. The individual strategies defined for this multi-agent system align closely with the strategies employed by the Game Society;
5. An ant colony-based agent system that uses from 1 to 3 agents and implements various tabu-search algorithms with the agents is able to solve from 68% to 81% of the puzzles (Mullaney, 2006). The 2 and 3 ant agent systems generally outperformed the single agent system as the number of runs increased. This performance is significantly less than the Game Society's and may demonstrate that biologic and cellular level systems simply require larger agent populations to solve NP-complete problems;
6. The final system is another biologic level system that utilizes a bee colony optimization model with 310 bee agents that solve 100% of the puzzles presented (Pacurib, Seno, & Yusiong, 2009), similar to the SoA, which uses a large number of agents, but fewer than 310 for this particular problem. The agents solve the puzzle using exhaustive search by randomly placing values, remembering incorrect solutions, and then replacing new values until the solution is found.

For the Kakuro game domain problems the Games Society currently solves almost all of the problems given to the Society (98 percent) and work is progressing to resolve the agent interactions and use of new Strategy agents to enable 100 percent performance for this domain problem as well. The Kakuro problems used to evaluate the Games Society again came from newspaper and open website games and ranged from easy to very difficult in difficulty level. Sample beginning and solved Board agent representations as actually printed by the Output agent are shown in Figure 7. No other multi-agent systems have been found that play Kakuro.

TTT if played correctly should always result in a draw. The Games Society has never lost a TTT game, with most games resulting in a draw, but occasionally winning if a human player plays at a suboptimal level. The initial testing of the TTT domain problem is accomplished by having a Games Society system play against another Games Society system, thus playing against another agent with similar knowledge. Later testing is performed by having the Games Society play against an expert human player. A sample TTT board produced by the Games Society is shown in Figure 8.

A previous single agent system at the cognitive level, but which incorporates a cellular ANN as its board evaluation function, that plays TTT claims that it can be trained to never lose against an opponent that uses a consistent strategy (Sardinha et al., 2003). This result is expected in any intelligent agent

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Sample Games Society Kakuro boards

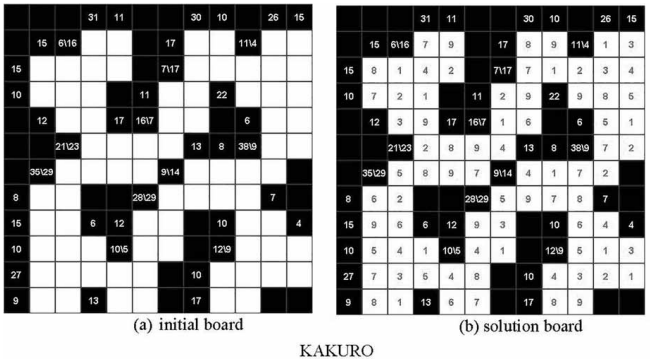
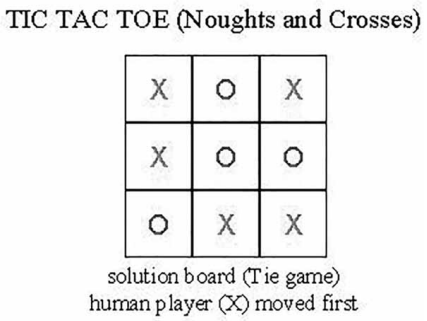


Figure 8. Sample Games Society completed Tic-Tac-Toe board

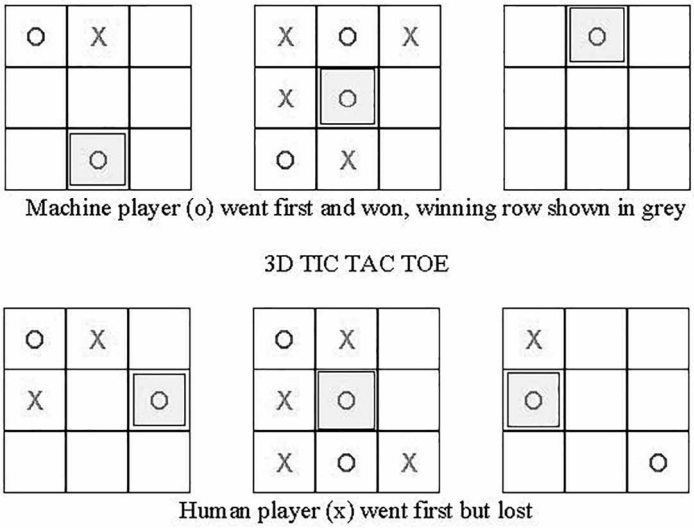


system and the Game Society does not lose regardless of opponent strategies. This agent system uses a training strategy similar to the Game Society by having individual agents play each other to improve their performance. Another cognitive level agent system utilizes 2 player agents, each of which incorporates different reinforcement learning strategies, and one board agent to teach itself to play TTT, but this system continues to lose (as the first player agent) after significant training (Batalov & Oommen, 2001).

The 3D-TTT game is winnable and the Games Society wins approximately 60 percent of its games, which depends on two factors: if it is the starting player (the Games Society wins 100 percent of the games in which it starts) and the level of play of the opponent. Again, initial testing is accomplished by letting the society play against a copy of itself and then later against an experienced human player, who intentionally played at various skill levels. Sample Output agent output for two completed 3D-TTT games are shown in Figure 9 (with shaded squares to highlight the Games Society's winning moves added), where the bottom board is shown on the left and the top board is shown on the right. The output from the Board agent in the first game example indicates that the Games Society is not playing a perfect game, since it should have only taken four moves to win versus the six moves indicated, thus the Games Society for 3D-TTT still has learning to accomplish or requires a new Strategy agent heuristic strategies to improve its performance. Also seen in the second 3D-TTT game in Figure 9 is that the Games Society for TTT and 3D-TTT always represents the opponent as the X player. A new Rule agent is needed for both TTT and 3D-TTT to let the Games Society know that the player who moves first should be represented with an X mark.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Sample Games Society 3-D Tic-Tac-Toe boards



The Caro board is unbounded and initially looks like a large TTT board. Because of the unbounded board, the game may not have a limit, but does have a winning condition (exactly 5 marks of the same player in a row, either horizontal, vertical, or diagonal). The Games Society was trained against a strong intermediate or mid-level player who intentionally played at both a beginner level and an intermediate level. The Games Society operated continuously for a pre-specified number of games prior to playing 5 evaluation games at each level. Due to the large board, no image is provided. The Games Society is able to progress in ability as the agents recognize different moves. The progression of the Games Society in learning to play Caro against a beginning player is shown in Table 3, which demonstrates the SoA capability for agents to adapt and learn from their environment during continuous play.

Because the Games Society is having difficulty winning Caro against a more advanced player, this implies the need for additional Strategy agent strategies that are not currently encoded. The Games Society does not currently implement any Memory agents. Addition of Memory agents, which would be initially created by the Recognition agent and then populated with information from Cell and Segment agents through the Board agent, would benefit games like Caro and also checkers and chess where knowledge

Table 3. Caro/Gomoku performance of the Games Society over multiple training examples

Games Trained	Opponent's Level	Games Society Performance		
		Win	Tie	Loss
0	Low	2	1	2
	Mid	0	0	5
10	Low	4	0	1
	Mid	0	0	5
25	Low	5	0	0
	Mid	0	0	5

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

of previous opponent move patterns could be exploited by the agents in the working community. The purpose of the Memory agent is to recall well known or solved positions, such as opening and closing books currently employed in checkers and chess programs. Future research is needed to investigate the implementation and effect of Memory agents and their utilization by Strategy agents, and other working community type agents.

Although the Games Society has demonstrated that it can learn, the intricacies of stronger play require more advanced heuristic Strategy agents. The response times of this Games Society distributed agent community approach to playing games are very fast, with no perceived (real-time) delay due to computation and interaction amongst the agents for each of the five solved games.

As indicated above and demonstrated in the Games Society instantiation, the SoA framework enables both centralized and decentralized agent coordination and multiple communication mechanisms to achieve collaborative problem solving, thus providing for the 3 necessary C's of multi-agent architectures. Furthermore, the definition and consequent instantiation of agents from the SoA framework limits agents to atomic well-defined tasks, including sensing the environment (perception), reasoning, memory (short term in the current games Society), and action. The SoA framework requires agents to work together to solve more complex (non-atomic) tasks. These cellular features of the SoA framework follow directly from the *Society of Mind* (Minsky, 1986) paradigm that inspired the SoA framework for multi-agent systems.

CONCLUSION AND FUTURE RESEARCH

The research reported in this article makes several advancements for multi-agent systems. First a new ontology for classifying and comparing agent and multi-agent systems is presented. Next, a new framework for development of multi-agent systems at the cellular level called the Society of Agents has been presented. This SoA framework is then utilized to develop a Games Society instantiation of the framework to demonstrate a collaborative multi-agent system for solving a wide variety of one and two person games with very high reliability and performance.

Additional research will progress in two distinct areas. The first will be to increment the quantity and type of games playable by the Games Society and to examine SoA Societies for other problem domains such as vehicle routing. As indicated work is currently progressing on solving the game Caro and future work is needed to examine checkers, chess, shogi, and Go. Additional games that rely on chance, such as backgammon, monopoly, and others, but which have recognized strategies and heuristics will also be developed. These additional games will increase the type and quantity of Strategy agents and Rule agents within the society. The application of the SoA framework to non-game domains will require defining new Rule agents for how to interpret the SoA Society's environment and possibly how the domain is represented in the Central Coordination agent. Tasks that are highly similar to games and have definable Central Coordination representations include network traffic analysis, power grid maintenance, supply chain management, and vehicle routing.

The second area for the research to progress in is the development of a learning capability within the SoA. Biologic-based agent systems have shown that the incorporation of cognitive capability does not improve agent performance over a traditional ant colony model (Sekara et al., 2015), but Stone (2007) states that multi-agent systems need to incorporate learning if they are truly to be qualified as intelligent. Currently the Games Society does not learn from past games and relies on its own inherent knowledge

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

from the Segment agents and various Strategy agents. The issue of short-term memory has been identified as a necessity for the SoA framework. If the SoA framework is to be used for development of "intelligent" multi-agent systems then the Memory agents defined within the SoA framework will need to be instantiated and added to the Games Society as well as new Strategy or Segment agents defined that will be able to utilize this long-term memory to analyze the efficacy of various Strategy agents within specific domain games and perhaps assist in the development of new Strategy agents dynamically. This instantiation of Memory agents into the Games Society would enable well-known strategies in games like chess, where opening and some ending moves in the domain are well known. A longer-term memory will also greatly enhance the analogy Strategy agent's applicability.

Cellular-level agents in multi-agent systems are shown to be adaptive problem solvers able to communicate and work collaboratively and cooperatively with other cellular-level agents to solve a diverse set of problems through reapplication of specific agent types across multiple domain problems. The SoA framework, operating at a cellular level, permits both centralized and decentralized communication amongst its atomic agents to support collaborative and cooperative problem solving. Based on the evidence produced in this research, new cellular-level agent research is needed to explore the capabilities of cellular-level multi-agent systems across a variety of problem types and domains.

ACKNOWLEDGMENT

The author would like to thank his colleague Dr. Reijer Grimbergen who first raised the idea of using the *Society of Mind* paradigm for creating intelligent systems and as a possibility for multi-agent systems. I would also like to thank my former doctoral student Dr. Toan Ong, who helped realize the Game Society instantiation of the proposed SoA framework by doing much of the Java programming and XML encoding.

REFERENCES

Abednego, L., & Nugraheni, C. (2014). A Block World Problem Based Sudoku Solver. *International Journal of Computer, Electrical, Automation. Control and Information Engineering*, *8*(8), 1425–1429.

Agerbeck, C., & Hansen, M. O. (2008). *A Multi-Agent Approach to Solving NP-Complete Problems* (Master's thesis]. Technical University of Denmark, Lyngby, Denmark.

Arsene, O., Dumitrache, I., & Mihu, I. (2015). Expert system for medicine diagnosis using software agents. *Expert Systems with Applications*, *42*(4), 1825–1834. doi:10.1016/j.eswa.2014.10.026

Barber, K. S., Goel, A., & Martin, C. E. (2000). Dynamic adaptive autonomy in multi-agent systems. *Journal of Experimental & Theoretical Artificial Intelligence*, *12*(2), 129–147. doi:10.1080/095281300409793

Batalov, D. V., & Oommen, B. J. (2001). On playing games without knowing the rules. In *Joint 9th IFSA World Congress and 20th NAFIPS International Conference* (Vol. 4, pp. 1862-1868). IEEE. 10.1109/NAFIPS.2001.944350

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Belkacemi, R., Feliachi, A., Choudhry, M. A., & Saymansky, J. E. (2011). Multi-Agent systems hardware development and deployment for smart grid control applications. In *2011 IEEE Power and Energy Society General Meeting*. IEEE. 10.1109/PES.2011.6039822

Bellifemine, F., Poggi, A., & Rimassa, G. (2001). Developing multi-agent systems with a FIPA-compliant agent framework. *Software, Practice & Experience*, *31*(2), 103–128. doi:10.1002/1097-024X(200102)31:2<103::AID-SPE358>3.0.CO;2-O

Bellifemine, F. L., Caire, G., & Greenwood, D. (2007). *Developing multi-agent systems with JADE* (Vol. 7). New York: John Wiley & Sons. doi:10.1002/9780470058411

Bradshaw, J. M., Dutfield, S., Benoit, P., & Woolley, J. D. (1997). KAoS: Toward an industrial-strength open agent architecture. In *Software agents* (pp. 375-418).

Byrski, A., Swiderska, E., Lasisz, J., Kisiel-Dorohinicki, M., Lenaerts, T., Smason, D., ... Nowe, A. (2017). Socio-cognitively inspired ant colony optimization. *Journal of Computational Science*, *21*, 397–406. doi:10.1016/j.jocs.2016.10.010

Chen, R. S., & Chen, D. K. (2008). Apply ontology and agent technology to construct virtual observatory. *Expert Systems with Applications*, *34*(3), 2019–2028. doi:10.1016/j.eswa.2007.02.018

de Weerdt, M. M., Zhang, Y., & Klos, T. (2012). Multiagent task allocation in social networks. *Autonomous Agents and Multi-Agent Systems*, *25*(1), 46–86. doi:10.100710458-011-9168-3

Duarte, V. A. R., Julia, R. M. S., Albertini, M. K., & Neto, H. C. (2015). Mp-draughts: Unsupervised learning multi-agent system based on MLP and adaptive neural networks. In *2015 IEEE 27th International Conference on Tools with Artificial Intelligence (ICTAI)* (pp. 920-927). IEEE.

Emery, M. R. (2007). *Solving Sudoku Puzzles with the COUGAAR Agent Architecture* [Master's thesis]. Montana State University-Bozeman, College of Engineering.

Eydi, A., Panahi, S., & Kamalabadi, I. (2017). User-based vehicle route guidance in urban networks based on intelligent multi agent systems and the ANT-Q algorithm. *International Journal of Transportation Engineering*, *4*(3), 147–161.

Ferber, J. (1999). *Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence*. New York: Addison-Wesley.

Grushin, A., & Reggia, J. A. (2008). Automated design of distributed control rules for the self-assembly of prespecified artificial structures. *Robotics and Autonomous Systems*, *56*(4), 334–359. doi:10.1016/j.robot.2007.08.006

Hafezi, R., Shahrabi, J., & Hadavandi, E. (2015). A bat-neural network multi-agent system (BNNMAS) for stock price prediction: Case study of DAX stock price. *Applied Soft Computing*, *29*, 196–210. doi:10.1016/j.asoc.2014.12.028

Hermans, H. J. (2002). The dialogical self as a society of mind: Introduction. *Theory & Psychology*, *12*(2), 147–160.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Huhns, M. N., & Stephens, L. M. (2000). Multiagent Systems and Societies of Agents. In G. Weiss (Ed.), *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. Cambridge, MA: MIT Press.

Jabbarpour, M. R., Malakooti, H., Noor, R. M., Anuar, N. B., & Khamis, N. (2014). Ant colony optimisation for vehicle traffic systems: Applications and challenges. *International Journal of Bio-inspired Computation*, *6*(1), 32–56. doi:10.1504/IJBIC.2014.059970

Jolly, K. G., Kumar, R. S., & Vijayakumar, R. (2010). Intelligent task planning and action selection of a mobile robot in a multi-agent system through a fuzzy neural network approach. *Engineering Applications of Artificial Intelligence*, *23*(6), 923–933. doi:10.1016/j.engappai.2010.04.001

Jolly, K. G., Ravindran, K. P., Vijayakumar, R., & Kumar, R. S. (2007). Intelligent decision making in multi-agent robot soccer system through compounded artificial neural networks. *Robotics and Autonomous Systems*, *55*(7), 589–596. doi:10.1016/j.robot.2006.12.011

Kar, A. K. (2016). Bio inspired computing–A review of algorithms and scope of applications. *Expert Systems with Applications*, *59*, 20–32. doi:10.1016/j.eswa.2016.04.018

Kendall, G., Parkes, A. J., & Spoerer, K. (2008). A Survey of NP-Complete Puzzles. *ICGA Journal*, *31*(1), 13–34.

Kiourt, C., & Kalles, D. (2015). Learning in multi agent social environments with opponent models. In M. Rovatsos, G. Vouros, & V. Julian (Eds.), *Multi-Agent Systems and Agreement Technologies* (pp. 137–144). Springer.

Kiourt, C., & Kalles, D. (2016). A platform for large-scale game-playing multi-agent systems on a high performance computing infrastructure. *Multiagent and Grid Systems*, *12*(1), 35–54. doi:10.3233/MGS-160242

Kolomvatsos, K., Panagidi, K., Neokosmidis, I., Varoutas, D., & Hadjiefthymiades, S. (2016). Automated concurrent negotiations: An artificial bee colony approach. *Electronic Commerce Research and Applications*, *19*, 56–69. doi:10.1016/j.elerap.2016.09.002

Kolp, M., Faulkner, S., & Wautelet, Y. (2008). Social Structure Based Design Patterns for Agent-Oriented Software Engineering. *International Journal of Intelligent Information Technologies*, *4*(2), 1–23. doi:10.4018/jiit.2008040101

Król, D., & Drożdżowski, M. (2010). Use of MaSE methodology for designing a swarm-based multiagent system. *Journal of Intelligent & Fuzzy Systems*, *21*(3), 221–231.

Leu, G., Tang, J., & Abbass, H. (2014). On The Role of Working Memory in Trading-Off Skills and Situation Awareness in Sudoku. In *International Conference on Neural Information Processing* (pp. 571-578). Springer. 10.1007/978-3-319-12643-2_69

Lian, F. L., Yook, J. K., Tilbury, D. M., & Moyne, J. (2006). Network architecture and communication modules for guaranteeing acceptable control and communication performance for networked multi-agent systems. *IEEE Transactions on Industrial Informatics*, *2*(1), 12–24. doi:10.1109/TII.2005.857611

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Liu, H., Zhang, P., Hu, B., & Moore, P. (2015). A novel approach to task assignment in a cooperative multi-agent design system. *Applied Intelligence*, *43*(1), 162–175. doi:10.100710489-014-0640-z

Liu, Y., & Passino, K. M. (2006). Cohesive Behaviors of Multiagent Systems with Information Flow Constraints. *IEEE Transactions on Automatic Control*, *51*(11), 1734–1748. doi:10.1109/TAC.2006.884948

Martin, D. L., Cheyer, A. J., & Moran, D. B. (1999). The open agent architecture: A framework for building distributed software systems. *Applied Artificial Intelligence*, *13*(1-2), 91–128. doi:10.1080/088395199117504

McCarthy, J. (1990). Chess as the Drosophila of AI. In T. A. Marsland & J. Schaeffer (Eds.), *Computers, Chess, and Cognition* (pp. 227–237). New York: Springer. doi:10.1007/978-1-4613-9080-0_14

Minsky, M. (1986). *The Society of Mind*. New York: Simon and Schuster.

Mishra, N., Kumar, V., & Chan, F. T. (2012). A multi-agent architecture for reverse logistics in a green supply chain. *International Journal of Production Research*, *50*(9), 2396–2406. doi:10.1080/00207543.2011.581003

Mullaney, D. (2006). *Using Ant Systems to Solve Sudoku Problems*. University College Dublin.

Neto, H. C., Julia, R. M. S., Caexeta, G. S., & Barcelos, A. R. A. (2014). LS-VisionDraughts: Improving the performance of an agent for checkers by integrating computational intelligence, reinforcement learning and a powerful search method. *Applied Intelligence*, *41*(2), 525–550. doi:10.100710489-014-0536-y

Nicolai, G., & Hilderman, R. J. (2009, September). No-limit Texas hold'em poker agents created with evolutionary neural networks. In *IEEE Symposium on Computational Intelligence and Games* (pp. 125-131). IEEE. 10.1109/CIG.2009.5286485

Object Management Group. (1995). *The Common Object Request Broker (CORBA): Architecture and Specification*. Object Management Group.

Pacurib, J. A., Seno, G. M. M., & Yusiong, J. P. T. (2009). Solving sudoku puzzles using improved artificial bee colony algorithm. In *Fourth International Conference on Innovative Computing, Information and Control* (pp. 885-888). IEEE. 10.1109/ICICIC.2009.334

Reeke, G. N. Jr. (1991). Marvin Minsky, *The Society of Mind. Artificial Intelligence*, *48*(3), 341–348. doi:10.1016/0004-3702(91)90034-H

Russell, S., & Yoon, V. Y. (2009). Agents, Availability Awareness, and Decision Making. *International Journal of Intelligent Information Technologies*, *5*(4), 53–70. doi:10.4018/jiit.2009080704

Saint Germain, B., Valckenaers, P., Verstraete, P., & Van Brussel, H. (2007). A multi-agent supply network control framework. *Control Engineering Practice*, *15*(11), 1394–1402. doi:10.1016/j.conengprac.2006.12.003

Samothrakis, S., Lucas, S., Runarsson, T., & Robles, D. (2013). Coevolving game-playing agents: Measuring performance and intransitivities. *IEEE Transactions on Evolutionary Computation*, *17*(2), 213–226. doi:10.1109/TEVC.2012.2208755

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sardinha, J. A. R. P., Milidiú, R. L., Lucena, C. J., & Paranhos, P. M. (2003). An OO Framework for building Intelligence and Learning properties in Software Agents. In *Proceedings of the 2nd International Workshop on Software Engineering for Large-Scale Multi-Agent Systems at ICSE*.

Sekara, M., Kowalski, M., Byrski, A., Indurkhya, B., Kisiel-Dorohinicki, M., Samson, D., & Lenaerts, T. (2015). Multi-pheromone ant colony optimization for socio-cognitive simulation purposes. *Procedia Computer Science*, *51*, 954–963. doi:10.1016/j.procs.2015.05.234

Siddique, N., & Adeli, H. (2015). Nature inspired computing: An overview and some future directions. *Cognitive Computation*, *7*(6), 706–714. doi:10.100712559-015-9370-8 PMID:26693257

Singh, R. (2007). A Multi-Agent Decision Support Architecture for Knowledge Representation and Exchange. *Journal of Intelligent Information Technologies*, *3*(1), 37–59. doi:10.4018/jiit.2007010103

Stone, P. (2007). Learning and Multiagent Reasoning for Autonomous Agents. In *Proceedings of the International Joint Conference on Artificial Intelligence*, (pp. 12-30). Morgan Kaufmann: San Francisco.

Tan, K. L., Tan, C. H., Tan, K. C., & Tay, A. (2009). Adaptive game AI for Gomoku. In *4th International Conference on Autonomous Robots and Agents* (pp. 507-512). IEEE.

Wang, L., & Shen, J. (2016). Multi-phase ant colony system for multi-party data-intensive service provision. *IEEE Transactions on Services Computing*, *9*(2), 264–276. doi:10.1109/TSC.2014.2358213

Wang, X., Zeng, Z., & Cong, Y. (2016). Multi-agent distributed coordination control: Developments and directions via graph viewpoint. *Neurocomputing*, *199*, 204–218. doi:10.1016/j.neucom.2016.03.021

West, M. J., & Gundersen, H. J. G. (1990). Unbiased stereological estimation of the number of neurons in the human hippocampus. *The Journal of Comparative Neurology*, *296*(1), 1–22. doi:10.1002/cne.902960102 PMID:2358525

Xiang, W., & Lee, H. P. (2008). Ant colony intelligence in multi-agent dynamic manufacturing scheduling. *Engineering Applications of Artificial Intelligence*, *21*(1), 73–85. doi:10.1016/j.engappai.2007.03.008

Yu, H., Shen, Z., & Miao, C. (2008). A goal-oriented development tool to automate the incorporation of intelligent agents into interactive digital media applications. *Computers in Entertainment*, *6*(2), 24. doi:10.1145/1371216.1371227

Zambonelli, F., Omicini, A., Anzengruber, B., Castelli, G., De Angelis, F., Di Marzo Serugendo, G., … Ye, J. (2014). Developing pervasive multi-agent systems with nature-inspired coordination. *Pervasive and Mobile Computing, 17(B),* 236-252.

This research was previously published in the International Journal of Intelligent Information Technologies (IJIIT), 14(4); edited by Vijayan Sugumaran; pages 1-23; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 10
Applying Optoelectronic Devices Fusion in Machine Vision:
Spatial Coordinate Measurement

Wendy Flores-Fuentes
Autonomous University of Baja California, Mexico

Julio C. Rodríguez-Quiñonez
Autonomous University of Baja California, Mexico

Moises Rivas-Lopez
Autonomous University of Baja California, Mexico

Javier Rivera-Castillo
Autonomous University of Baja California, Mexico

Daniel Hernandez-Balbuena
Autonomous University of Baja California, Mexico

Lars Lindner
Autonomous University of Baja California, Mexico

Oleg Sergiyenko
Autonomous University of Baja California, Mexico

Luis C. Basaca-Preciado
Center of Excellence in Innovation & Design – CETYS University, Mexico

ABSTRACT

Machine vision is supported and enhanced by optoelectronic devices, the output from a machine vision system is information about the content of the optoelectronic signal, it is the process whereby a machine, usually a digital computer and/or electronic hardware automatically processes an optoelectronic signal and reports what it means. Machine vision methods to provide spatial coordinates measurement has developed in a wide range of technologies for multiples fields of applications such as robot navigation, medical scanning, and structural monitoring. Each technology with specified properties that could be categorized as advantage and disadvantage according its utility to the application purpose. This chapter presents the application of optoelectronic devices fusion as the base for those systems with non-lineal behavior supported by artificial intelligence techniques, which require the use of information from various sensors for pattern recognition to produce an enhanced output.

DOI: 10.4018/978-1-7998-0951-7.ch010

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The present chapter surged in the research continuity of a 3D Vision System for mobile robot navigation application, a 3D medical laser scanner, and a structural health monitoring system. With the objective of increasing the accuracy of the systems, digital and analog processing signals methodologies have been developed in order to find the energetic center of the optoelectronic signal handled by these systems. Into the task of systems overall robustness, its measurement data has been submitted to statistical analysis, finding a non-linear behavior of the systems, leading to the need of artificial intelligence applications such as neuronal network (NN) and support vector machine regression (SVMR), in a modern approach to the prediction of the non-linear measurement error of the systems to compensate it. In the process of obtaining enough information from a measurement system to extract from it a model to predict its measurement error. It has been done a search of attributes to build the training dataset and test dataset. Ha been found that the pattern recognition can be enhanced by the sensor fusion and redundancy theory. This theory refers to the synergistic use of information from various sensors to achieve the task required by the system. Input data (attributes) are combined, fused and grouped for proper quality and integrity of the decisions to be taken by the intelligent algorithm. Besides, the benefits can be extracted from the redundant data, the reduction of uncertainty and the increasing of precision reliability. By these reasons, the photodiodes and charge coupled devices (CCD) are fusion in the task of robust systems building for machine vision by Spatial Coordinates Measurement (Weckenmann, 2009; Elfes, 1992; Zhang, 2008; Shih, 2015). The specific properties of both, their advantages and limitations have been considered, since, the photodiode is the sensor who gives place to the laser-scanning and the CCD is the sensors who gives place to the close-range photogrammetry. The energetic center of the laser optoelectronic signals from the photodiode and the energetic center of the image signal from the CCD sensor are detected to combine these sensors outputs, and to exploit their natural synergy new experimental results are presented to demonstrate the increase of systems accuracy.

BACKGROUND

Optoelectronics is the study of any devices that produce an electrically-induced optical output or an optically-induced electrical output and the techniques for controlling such devices (Marston, 1999), it includes generation, transmission, routing, and detection of optoelectronic signals in a widespread of applications (Dagenais, 1995). Wherever light is used to transmit information, tiny semiconductor devices are needed to transfer electrical current into optical signals and vice versa. Examples include light-emitting diodes, photodetectors and laser diodes (Piprek, 2003).

Most optoelectronics devices applications have focused on single sensors and relatively simple processes to extract specific information from the sensor, however the use of multiple sensors by an optoelectronic device fusion technology deliver more advanced information and enable to develop intelligent and sophisticated optoelectronic systems, in special for machine vision applications (Yallup, 2014). More than one optoelectronic sensor may be needed to fully monitor the observation space at all times. Methods of combining multiple sensor data are in developing due to the availability and computational power of communications devices that support algorithms needed to reduce the raw sensor data from multiple sensors to convert it to the information needed by the system user (Klein, 2003).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Optoelectronic systems with only one sensor are not recommended for those optoelectronics signals with non-lineal behavior, the observations made by one sensor could be uncertain and occasionally incorrect, a single sensor could not detect all the critical characteristics of the optoelectronic signal as could be done with multiple sensors, which could provide different information each one, in special if they are from different technologies, they can provide different information regarding the target under observation. Besides, if a unique sensor is used, and it fails, the optoelectronic system is shutdown, while this sensor it is not working. Several sensors deliver redundant information to increase reliability in case of errors or failure of a sensor. The redundant information can reduce uncertainty and increases the accuracy with which the characteristics are perceived by the optoelectronic system. Further information from various sensors is obtained to characterize the environment, perceived in a way that would be impossible using only the information for each sensor separately (Castanedo, 2013).

Machine vision is supported and enhanced by optoelectronic devices, the output from a machine vision system is information about the content of the optoelectronic signal, it is the process whereby a machine, usually a digital computer and/or electronic hardware automatically processes an optoelectronic signal and reports what it means. That is, it recognizes the content, it not only locates the content, but to inspect it as well. Machine vision includes two components, measurement of features and pattern classification based on those features. Defining the pattern classification as the process of making decision about a measurement, with a set of measurements knowledge about the possible classes could be obtained, leading to make a decision (Snyder, 2010).

The machine vision technology has a widespread of applications, as is the automatic target recognition, 3D vision, vehicle vision and the industrial inspection, that make use of it for replace or complement manual inspections and measurements, to automate the production, in order to increase production speed and yield, as well as to improve product quality.

In special, machine vision methods to provide spatial coordinates measurement has developed in a wide range of technologies for multiples fields of applications such as robot navigation, medical scanning, and structural health monitoring. Each technology has specific properties that could be categorized as advantage and disadvantage according its utility to the application purpose.

The computer vision has guide the machine vision to the tendency of duplicate the abilities of human vision by electronically perceiving and understanding an image with a high a dimensional data, increasing the data storage requirement, and the time processing due to the complexity of algorithms to extract important patterns and trends and understand what data says (learning data). For some applications it is not necessary to acquire the whole vision of a scene as makes the human vision to extract the features of interest, it is possible to only extract the significant characteristics from the scene by the use of a positioning laser (PL) (Rodriguez-Quiñonez, 2014), a rotatory mirror scanner and dynamic triangulation and/or by the extraction of important features from images captured through the use of camera with CCD (Charge Coupled Device) (Hou, 2011).

Researches previously published has found that the 3D modelling of objects and complex scenes constitutes a field of multi-disciplinary challenges and difficulties, in the going-over accuracy, reliability, quality, portability, low cost and automatization of the whole procedure. By these reasons although there are a wide variety of sensors, the optical scanning and the digital cameras play the main role in this application, because even though these two types of sensors can work in a separate fashion, it is when are fusion together when the best results are attained. The sensor fusion, in particular concerning these both technologies appears as a promising possibility to improve the data acquisition and the geometric and radiometric processing of these data (Gonzalez-Aguilera, 2010).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Optoelectronic Device Fusion in Machine Vision is expected to play an increasingly important role in the near future, by its desired characteristics, such as: optimization, deduction, robustness, flexibility, inference, accuracy, reliability, speed and cost-effectiveness.

By this motivation it is presented a review of modern optoelectronic devices, including light-emitting diodes, edge-emitting lasers, vertical-cavity lasers, electro absorption modulators, a novel combination of amplifier and photodetector (Piprek, 2003), and CCD, to know the diverse optoelectronic devices, and which are the most suitable according to the needs in numerous machine vision applications.

The available sensor fusion techniques are classified into three nonexclusive categories:

1. Data association,
2. State estimation, and
3. Decision fusion.

Also architectures and algorithms are introduced.

Machine vision operation are described, an introduction to the methods for imaging, processing, analysis, communication and actions are developed.

The current optoelectronic scanning technology and image processing for 3D reconstruction techniques are fusion, and a briefly introduction to the optimal hardware architecture is reviewed.

OPTOELECTRONIC SENSORS FUSION IN MACHINE VISION

Spatial Coordinate Measurement

Machine vision methods to provide spatial coordinates measurement has developed in a wide range of technologies, the computer vision has guide the machine vision to the tendency of duplicate the abilities of human vision by electronically perceiving and understanding an image with a high a dimensional data, increasing the data storage requirement, and the time processing due to the complexity of algorithms to extract important patterns and trends and understand what data says (learning data). For some applications it is not necessary to acquire the whole vision of a scene as makes the human vision to extract the features of interest (Sonka, 2014), it is possible to only extract the significant characteristics from the scene by the use of a positioning laser (PL), a rotatory mirror scanner and dynamic triangulation, for detailed system overall see (Basaca-Preciado, 2014; Rodriguez-Quiñonez, Sergiyenko, Preciado, et al, 2014; Flores-Fuentes, 2014a; Flores-Fuentes, 2014b).

Optoelectronic Devices

Optoelectronic devices take advantage of sophisticated interactions between electrons and light, it combines electronics and photonics. Photons are the smallest energy packets of light waves and their interaction with electrons, energy exchange between them is the key physical mechanism in optoelectronic devices. Optoelectronics brings together optics and electronics within a single device, a single material, semiconductors designed to allow for the transformation of light into current and vice versa.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. **Photodetector Devices:** The purpose of any photodetector is to convert light (photons) into electric current (electrons). Photodetectors are among the most common optoelectronic devices; they are used to automatic open doors, detect the signal from infrared remote controls, and record pictures in modern cameras. Photodetectors are fabricated from various semiconductor materials, since the band gap needs to be smaller than the energy of the photons detected. Photon absorption generates electron–hole pairs which are subsequently separated by the applied electrical field. Depending on the desired use, photodetectors are designed in many different ways, and can be classified in photodiodes, infrared-detectors and charge coupled devices (CCD), as shown in Figure 1. Modern photodetectors include an amplifier; the amplification photodetector simultaneously amplifies and absorbs the incoming light.
    a. **Photodiode:** Photodiodes are light detecting devices fabricated in semiconductor materials, when a photon of sufficient energy strikes a diode, it creates an electron-hole pair. This mechanism is also known as the inner photoelectric effect. If the absorption occurs in the junction's depletion region, or one diffusion length away from it, these carriers are swept from the junction by the built-in electric field of the depletion region. Thus holes move toward the anode, and electrons toward the cathode, and a photocurrent is produced. The total current through the photodiode is the sum of the dark current (current that is generated in the absence of light) and the photocurrent, so the dark current must be minimized to maximize the sensitivity of the device. Photodiodes are similar to regular semiconductor diodes except that they may be either exposed or packaged with a window or optical fiber connection to allow light to reach the sensitive part of the device (Tavernier, 2011). The most common photodiodes are the P-N photodiode, PIN photodiode, avalanche photodiode, waveguide photodiode and amplification photodiode (Syms, 1992; Piprek, 2003).
    b. **Infrared-Detectors:** This detector is sensible to infrared radiation (IR). Thermal and photonic photodetectors are the main types of infrared-detectors. The effects of the incident IR radiation can be followed through many temperature dependent phenomena; a thermal detector type has no wavelength dependence. Major characteristics indicating infrared detector performance are photo sensitivity, the photodetector is wavelength dependent. In comparison with visible and ultraviolet rays, infrared radiation has small energy. To increase infrared detection efficiency, the detector should be cooled. The signal output from a detector is generally quite small and needs to be amplified. When designing a preamplifier, it is considered the impedance match between the detector and the amplifier, if the incident light is modulated by a chopper, the use of a tuned amplifier is necessary to reduce noise. If the detector is cooled, it is also practical to cool the amplifier. IR photodetectors are classified in intrinsic type and extrinsic type (Ge:Au, Ge:Hg, Ge:Cu, Ge:Zn, Si:Ga, Si:As), the intrinsic type is classified on Photoconductive type (PbS, PbSe, InSb, HgCdTe) and Photovoltaic type (Ge, InGaAs, Ex. InGaAs, InAs, InSb, HgCdTe) (Rogalski, 2002). The most common infrared detectors are InGaAs PIN photodiode and linear image sensor, PbS and PbSe photoconductive detector, InAs and InSb photovoltaic detector, MCT (HgCdTe) and InSb photoconductive detector, and two color detector (McClure, 2003; Tribolet, 2008).
    c. **Charge Coupled Devices (CCD):** CCD is the abbreviation for charge-coupled device, also known as CCD image sensor. CCD sensors are typically built with silicon technology. Silicon based integrated circuits (ICs), consisting of a dense matrix of photodiodes or photogates that

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

operate by converting light energy, in the form of photons, into electronic charges. When a photon strikes a silicon atom in or near a CCD photosite, the photon usually produces a free electron and a hole via the photoelectric effect.

The primary function of the CCD is to collect the photogenerated electrons in its potential wells (or pixels) during the CCD's exposure to radiation, and the hole is then forced away from the potential well and is eventually displaced into the Si substrate. While more light that is incident on a particular pixel, the higher the number of electrons that accumulate on that pixel. The output signal is then transferred to the computer for image regeneration or image processing (Li, 2006).

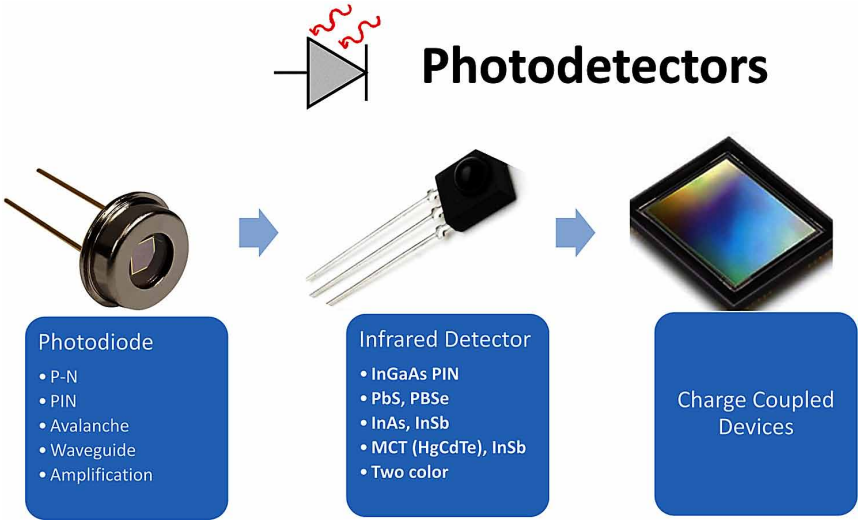
The pixels in a CCD sensor can be arranged in various configurations, the linear and area array are the classical architectures. A linear or linescan CCD sensor consists of a single line of pixels, adjacent to a CCD shift register that is required for the read-out of the charge packets. The isolation between the pixels and the CCD register is achieved by a transfer gate. Typically, the pixels of a linear CCD are formed by the photodiodes. Two-dimensional imagining is possible with area array CCD sensors; the entire image is captured with one exposure, eliminating the need for any movement by the sensor or the scene. As charge coupled device (CCD) technology matures it dominates the digital imaging, for commercial and industrial manufacturing applications (Janesick, 2001).

Conventional CCDs are designed for the front illuminated mode of operation. Front illuminated CCDs are quite economical to manufacture by using standard wafer fabrication procedures, and are popular in consumer imagining applications, as well as industrial-grade applications. However, front-illuminated CCDs are inefficient at short wavelengths (e.g., blue and UV) due to the absorption of photons by the polysilicon layers if photogates are used as the pixel element. Typically, conventional front-illuminated CCDs are adequate for low-end applications and for consumer electronics. But, for large, professional observatories and high-end industrial inspection systems that demand that require extremely sensitive detectors, conventional thick frontside-illuminated chips are rarely used. Thinned back-illuminated CCDs offer a more compatible solution for these applications. Backside-thinned back-illuminated chips are rarely used. Thinned back-illuminated CCDs offer a more compatible solution for these applications. Backside-thinned back-illuminated CCDs exhibit a superior responsively, remarkably in the shorter wavelength region (Li, 2003).

2. **Light Emitting Devices:** Other application of the p-n junction is as a light source, known as a light emitting diode (LED). It is based in the electroluminescence optical-electrical phenomenon, electrical generation of light through spontaneous emission. It is when a material emits light in response to the passage of an electric current, a radiative recombination of electrons and holes are done in the material, usually a semiconductor. In a semiconductor electrons and holes are separated by doping the material to form a p-n junction, then a recombination of electrons and holes are done, and it is when the excited electrons release their energy as photons (light). As an isolated lump of material will not emit significant quantities of light, in thermal equilibrium (room temperature) the number of downward electron transitions is extremely small. To improve the optical output, the material should be moved far from equilibrium, so that the rate of spontaneous emission is considerably increased. This might be done by taking (for example) a p-type material, which already contains a large holes density, and pouring electrons into it by a forward-biased p-n junction. Spontaneous emission generates photons that travel in random directions, so the emission is isotropic, the emission is also un-polarized, and the output is incoherent, with a spectrum consisting of a broad range

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Photodetectors classification



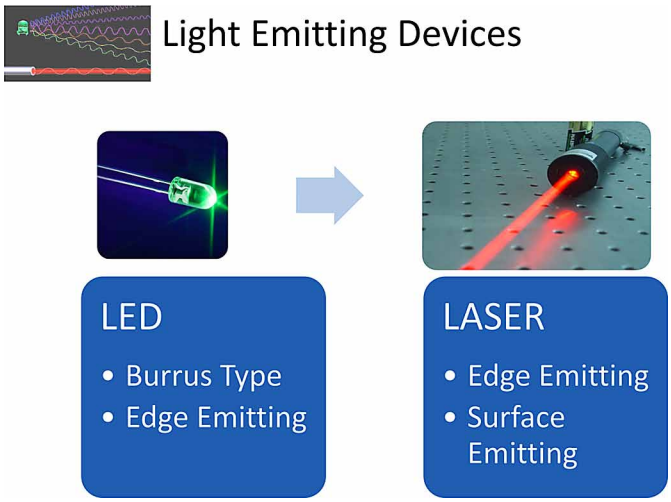
of wavelengths, consequently, the LED cannot be used as a source for high-speed, long-distance optical communications; the dispersion caused by such an extended spectrum would be far too large. LED transmitters are therefore restricted to short-haul applications.

A single photon is able to generate an identical second photon by stimulating the recombination of an electron-hole pair. This photon multiplication is the key physical mechanism of lasing. The second photon exhibits the same wavelength and the same phase as the first photon, doubling the amplitude of their monochromatic wave. Subsequent repetition of this process leads to strong light amplification. However, the competing process is the absorption of photons by the generation of new electron-hole pairs. Stimulated emission prevails when more electrons are present at the conduction band level than at the valence band level. This carrier inversion is the first requirement of lasing and it is achieved at pn-junctions by providing conduction band electrons from the n-doped side and valence band holes from the p-doped side. The photon energy is given by the band gap, which depends on the semiconductor material. Continuous current injection into the device leads to continuous stimulated emission of photons, but only if enough photons are constantly present in the device to trigger this process. Therefore, optical feedback and the confinement of photons in an optical resonator is the second basic requirement of lasing.

The absorption and simulated emission process produced in a LED is the principle of a light amplification by stimulated emission of radiation (LASER), the exponentially growing wave produce an amplification ought to offer a method of light generation. To achieve this, the electron density is raised far above the equilibrium level by the injection of electrons across a forward-biased diode. The laser is a device based on the stimulated emission of electromagnetic radiation. It emits light through a process of optical amplification; it emits spatial coherently light as a tight spot, achieving a very high irradiance, or they can have very low divergence in order to concentrate their power at a great distance. The laser enables applications such as cutting, welding and lithography for optical disk drives, laser printers, barcode scanners, fiber-optic, optical communication, surgery, skin treatments, device enforcement for marking targets and measuring range and speed, and lighting display.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Light emitting devices classification



The first semiconductor lasers consisted of heavily doped p-n junctions with cleaved facets. This construction is inefficient, as there is no defined region in which recombination can take place. Carriers can be lost to diffusion before recombination occurs. As a result, these early devices required a lot of current to reach threshold. Threshold currents were reduced with the developments of the double-heterostructure laser, which has a thin region of semiconductor with a smaller energy gap sandwiched between two oppositely doped semiconductors with a wide bandgap energy. When forward biased, carriers flow into the active region and recombine more efficiently because of the potential barriers of the heterostructure confine the carriers to the active region. There is an added advantage of guided the laser light, because the refractive index of the cladding layers is less than the active region. See Figure 2.

a. **Burrus-Type LED:** In place of the simple p-n junction, a more complicated structure is used for the Burrus-type LED. Typically, this might be a P-n-N double heterostructure, containing three layers: a p-type layer with a wide energy gap, a narrow-gap n-type layer, and a wide-gap n-type layer. A single heterojunction provides potential barriers of different heights for electrons and holes; the double heterostructure provide a high barrier at different positions for electrons and holes. In this way, the recombination region may be limited to a defined region of space, resulting in an increase in the radiative recombination efficiency. This is a high-efficiency device, suitable for use with multimode fiber. To fabricate the device, a double heterostructure is first grown on a substrate, and a SiO_2 isolation layer is then deposited on the lower layer of the heterojunction. This layer is etched, to expose the hetero- junction over a small window. Metallization layers are then added to the upper and lower surfaces; clearly, the lower of these makes contact with the heterojunction only in the region of the window. A deep well, aligned with this window, is then etched right through the substrate to the upper layer of the heterojunction, and a multimode optical fiber is epoxied into the well in contact with the LED surface. When a forward current is passed through the LED, it flows mainly through the region immediately below the fiber. The recombination region (or active region) is therefore confined vertically by the double heterojunction and laterally by the distribution of current flow.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

b. **Edge Emitting LED:** This Edge Emmiting LED also called superluminescent LED presents an alternative geometry to the Burrus-type LED (Fukuda, 1999). This is also based on a double heterostructure, but now the emission is taken from the edge of the junction rather than its surface. Generally, the current is forced to flow through a narrow strip down the device center line by the introduction of a current-blocking silica layer. As a result, the active volume is constrained vertically by the double heterostructure and laterally by the current flow. Most importantly, light generated by spontaneous emission is also confined vertically to a certain extent; the refractive index difference at the top and bottom of the active layer results in total internal reflection. This tends to channel a large fraction of the light towards the emission windows at relatively shallow angles.
c. **Edge Emitting Laser:** In this Laser the light propagates in a direction along the wafer surface of the semiconductor chip and is reflected or coupled out at a cleaved edge. These are the original and still very widely used form of semiconductor lasers. Within the edge-emitting laser structure, the laser beam is guided in a waveguide structure. Typically, one uses a double heterostructure, which restricts the generated carriers to a narrow region and at the same time serves as a waveguide for the optical field. This arrangement leads to a low threshold pump power and a high efficiency. Depending on the waveguide properties, particularly its transverse dimensions, it is possible either to obtain an output with high beam quality but limited output power, or an output with high output power but with poor beam quality.
d. **Surface Emitting Laser:** In the surface-emitting laser, the light propagates in the direction perpendicular to the semiconductor wafer surface. These are further subdivided into monolithic and external-cavity devices: Monolithic means that the laser resonator is realized in the form of two semiconductor Bragg mirrors with the quantum well section in between. Such a device is called VCSEL (vertical cavity surface-emitting laser) and is electrically pumped in most cases (Lorenser, 2003).

Sensors and Data Fusion

Data fusion refers to the synergistic use of information from different sensors for proper quality and integrity of the decision making to accomplish a task required by a system. All tasks that demand any type of parameter estimation from multiple sources can benefit from the use of data/information fusion methods. The terms information fusion and data fusion are typically employed as synonyms; but in some scenarios, the term data fusion is used for raw data (obtained directly from the sensors) and the term information fusion is employed to define already processed data. In this sense, the term information fusion implies a higher semantic level than data fusion. Other terms associated with sensor fusion that typically appear in the literature include decision fusion, data combination, data aggregation, multi-sensor data fusion, and data fusion (Castanedo, 2013).

Systems with only one sensors are not recommended for those with non-lineal behavior, the observations made by one sensor could be uncertain and occasionally incorrect, a single sensor could not detect all the critical characteristics of the system as could be done with multiple sensors, which could provide different information each one, in special if they are from different technologies and can provide different information regarding the system. Besides if a unique sensor is used, and it fails the system is shutdown while this sensor it is not working. The redundant information can reduce uncertainty and increases the accuracy with which the characteristics are perceived by the system. Several sensors delivering redundant

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

information increase reliability in case of errors or failure of a sensor. Further information from various sensors to characterize the environment perceived in a way that would be impossible using only receive the information for each sensor separately.

The goal of using sensor fusion in multi-sensor environments is to obtain a lower detection error probability and a higher reliability by using data and information from multiple distributed sources, as shown in Figure 3.

1. **Sensors and Data Fusion Classification:** Fusion techniques and methods are difficult to classify due to they involve a big variety of application fields. A general classification is based on data criteria, as described at Figure 4.
    a. **Relation between the Input Data Sources (Sensors):** Complementary: Each sensor provides complementary information that can be combined together, the information provided by the sensors represents different parts of the scene and could thus be used to obtain more complete global information (Vauhkonen, 2014).
        i. **Redundant:** Various sensors providing information about the same target or scene (Abidi, 1992).
        ii. **Cooperative:** Two or more sensors are independently used to perform detection, providing information that is combined into new information that is typically more complex than the original information (Labayrade, 2005).
    b. **Input/Output Data Sensors Types and Their Nature:** Low Level Data in-Data out (DAI-DAO): The raw data are directly provided as an input to the data fusion process. This type of data fusion process inputs and outputs raw data; the results are typically more reliable or accurate. Data fusion at this low level is conducted immediately after the data are gathered from the sensors. The algorithms employed at this level are based on signal and image processing algorithms.
        i. **Low Level Data In-Feature Out (DAI-FEO):** At this low level, the data fusion process employs raw data from the sources to extract features or characteristics that describe an entity in the environment;
        ii. **Medium Level Feature In-Feature Out (FEI-FEO):** At this medium level, both the input and output of the data fusion process are features or characteristics (shape, texture, and position), the data fusion process addresses a set of features with to improve, refine or obtain new features. This process is also known as feature fusion, symbolic fusion, information fusion or intermediate-level fusion.
        iii. **High Level Feature In-Decision Out (FEI-DEO):** Takes symbolic representations as sources and combines them to obtain a more accurate decision, this high level obtains a set of features as input and provides a set of decisions as output. Most of the classification systems that perform a decision based on a sensor's inputs fall into this category of classification.
        iv. **High Level Decision in-Decision out (DEI-DEO):** This type of classification is also known as decision fusion. It fuses input decisions to obtain better or new decisions (Dasarathy, 1997).
        v. **Multiple Level:** It is a combination of data provided from different levels of abstraction.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Sensors and data fusion applied to systems with non-lineal behavior

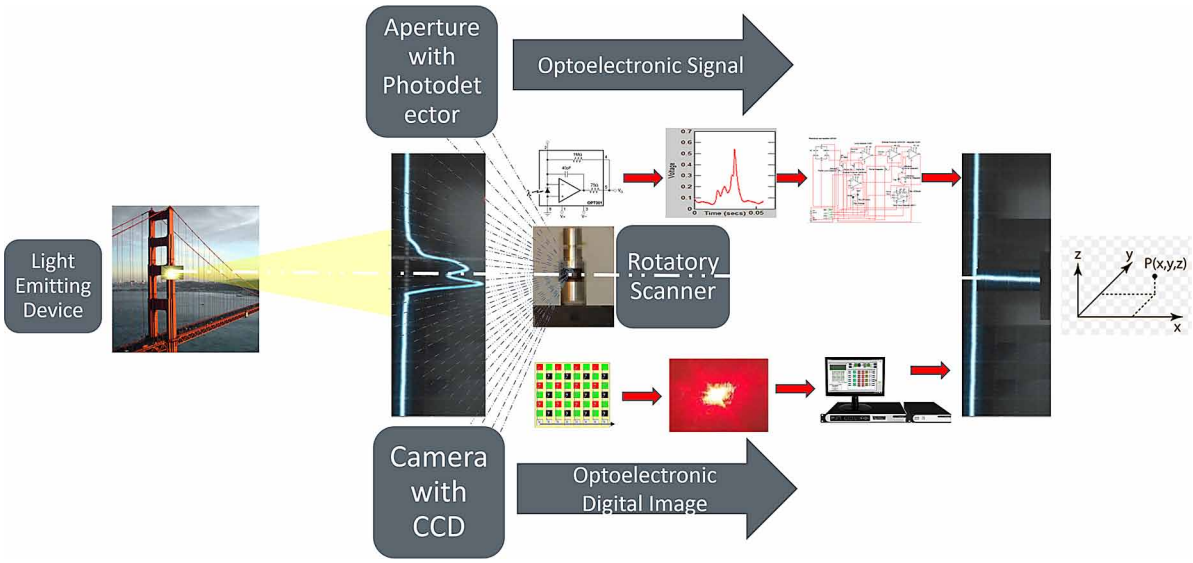
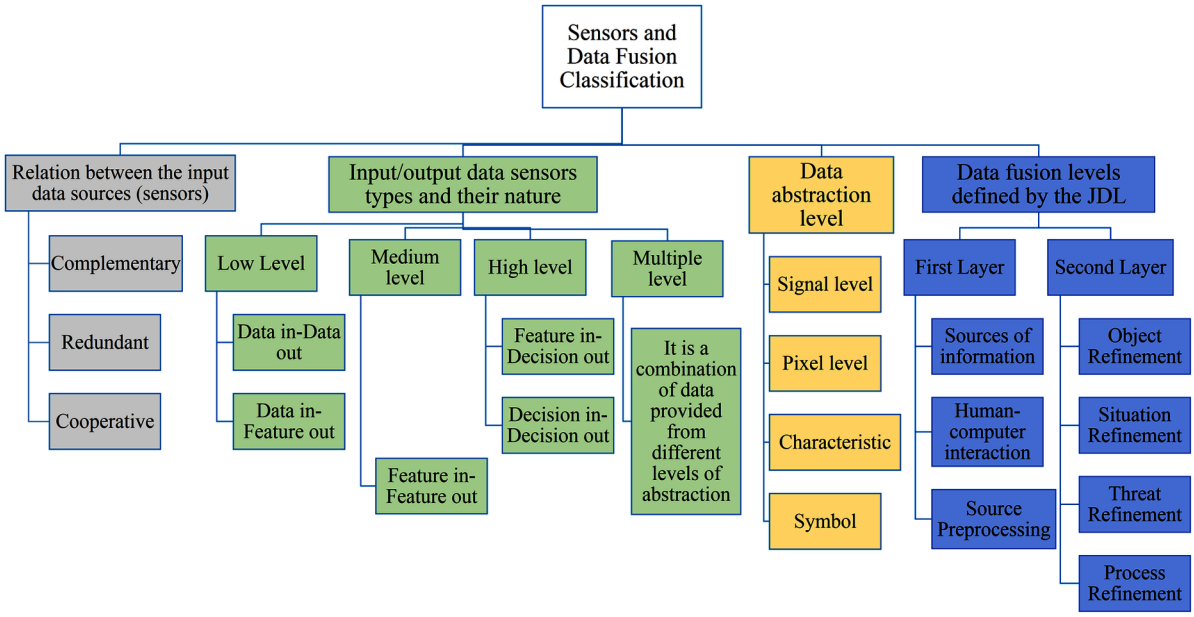


Figure 4. Sensors and data fusion classification



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

c. **Data Abstraction Level:** It is typically classified in three levels of abstraction: measurements, characteristics, and decisions. Other possible classifications of data fusion based on the abstraction levels are as follows:
   i. **Signal Level:** Directly addresses the signals that are acquired from the sensors.
   ii. **Pixel Level:** Operates at the image level and could be used to improve image processing tasks.
   iii. **Characteristic:** Employs features that are extracted from the images or signals (i.e., shape or velocity),
   iv. **Symbol:** At this level, information is represented as symbols; this level is also known as the decision level (Luo, 2002).

d. Data fusion levels defined by the JDL (Joint Directors of Laboratories).
   i. **First Layer Sources of Information:** Local and distributed sensors linked to a data fusion system, and or a priori information, as databases.
   ii. **First Layer Human-Computer Interaction (HCI):** It is an interface that allows inputs to the system from the operators and produces outputs to the operators. HCI includes queries, commands, and information on the obtained results and alarms. It incorporates not only multimedia methods for human interaction (graphics, sound, tactile interface, etc.), but also methods to assist humans in direction of attention, and overcoming human cognitive limitations.
   iii. **First Layer Source Preprocessing:** An initial process allocates data to appropriate processes and performs data pre-screening. Source preprocessing reduces the data fusion system load by allocating data to appropriate processes. The database management system stores the provided information and the fused results. This system is a critical component because of the large amount of highly diverse information that is stored.
   iv. **Second Layer Level 1 Processing (Object Refinement):** Transforms sensor data into a consistent set of units and coordinates, refines and extends in time estimates of an object's position, kinematics, or attributes, assigns data to objects to allow the application of statistical estimation techniques, and refines the estimation of an object's identity or classification.
   v. **Second Layer Level 2 Processing (Situation Refinement):** Formal and heuristic techniques are used to examine, in a conditional sense, the meaning of Level 1 processing results.
   vi. **Second Layer Level 3 Processing (Threat Refinement):** It is an impact assessment, at this level the impact of techniques used in level 2 are evaluated to obtain a proper perspective. And a future projection is performed to identify possible risks, vulnerabilities, and operational opportunities. This level includes an evaluation of the risk or threat and a prediction of the logical outcome.
   vii. **Second Layer Level 4 Processing (Process Refinement):** It is a process concerned about other processes, monitors the data fusion process performance to provide information about real-time control and long-term performance, identifies what information is needed to improve the multilevel fusion product (inferences, positions, identities, etc.), determines the source specific requirements to collect relevant information (i.e., which sensor type, which specific sensor, which database), and allocates and directs the sources to achieve mission goals (Luo, 2002).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. **Sensors and Data Fusion Methods:** Most current sensors data fusion methods algorithms can be broadly classified as illustrate at Figure 5.
    a. **Estimation Methods:** These methods take the weighted average of redundant information coming from a group of sensors and use it as the value of the fusion sensors. These methods determine the state of the target under movement (typically the position) given the observation or measurements. State estimation techniques are also known as tracking techniques with a final goal of obtain a global target state from the observations. The estimation problem involves finding the values of the vector state (e.g., position, velocity, and size) that fits as much as possible with the observed data. From a mathematical perspective, we have a set of redundant observations, and the goal is to find the set of parameters that provides the best fit to the observed data. In general, these observations are corrupted by errors and the propagation of noise in the measurement process. State estimation methods fall under level 1 of the JDL classification and could be divided into two broader groups:
        i. **Linear Dynamics and Measurements:** Here, the estimation problem has a standard solution. Specifically, when the equations of the object state and the measurements are linear, the noise follows the Gaussian distribution, do not refer to it as a clutter environment; in this case, the optimal theoretical solution is based on the Kalman filter;
        ii. **Nonlinear Dynamics:** The state estimation problem becomes difficult, and there is not an analytical solution to solve the problem in a general manner. In principle, there are no practical algorithms available to solve this problem satisfactorily. This method allows real-time processing of data. Some of them are: Maximum Likelihood and Maximum Posterior, Particle Filter, The Kalman Filter, The Distributed Kalman Filter, Distributed Particle Filter, and Covariance Consistency Methods: Covariance Intersection/Union. Being The Kalman filter the most popular.
    b. **Classification Methods:** A data association is performed to determine the set of measurements that correspond to each target. The multidimensional feature space can be partitioned into distinct regions, each representing an identification or identity class. The location of a feature vector is compared to pre-specified locations in feature space. A similarity measure must be computed, and each observation is compared to a priori classes. A feature space may be partitioned by geometrical or statistical boundaries. Therefore, the templating approach may declare a unique identity or an identity with an associated uncertainty. The implementation of parametric templates is computationally efficient for multisensor fusion systems. The most common methods are Nearest Neighbors and K-Means.
    c. **Inference Methods:** Bayesian inference allows multisensor information to be combined according to the rules of probability theory. Bayes' formula provides a relationship between the a priori probability of a hypothesis, the conditional probability of an observation given a hypothesis, and a posteriori probability of the hypothesis. Bayesian inference updates the probabilities of alternative hypotheses, based on observational evidence. New information is used to update the a priori probability of the hypothesis.
    d. **Artificial Intelligence Methods:** In these methods decisions are typically taken based on the knowledge of the perceived situation, a high-level inference is required, as the human reasoning, for pattern recognition, planning, induction, deduction, and learning.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The artificial intelligence methods are based on an initial data set and the rule-based knowledge base comprising rules, frames and scripts. The inference process uses the a priori data set, and searches the complete set of rules to identify applicable rule. The rule selection strategies from among multiple applicable rules include refraction, actuation, rule ordering, specificity, and random choice. Examples of these methods are Neuronal Network, Fuzzy Logic, Support Vector Machines, etc.

Artificial intelligence methods can be seen as the advanced versions of the estimation, the classification, and the inference methods. Artificial intelligence methods can be model-free, rather than model-specific, and have sufficient degree of freedom to fit complex nonlinear relationships, with the necessary precautions to properly generalize. As a result, artificial intelligence methods can effectively conduct sensor fusion at different levels (Hall, 1997).

3. **Sensors and Data Fusion Architectures:** This describes where the sensors and data fusion process will be performed, as shown in Figure 6.
    a. **Centralized Architecture:** All of the fusion processes are executed in a central processor that uses the provided raw measurements from the sources.
    b. **Decentralized Architecture:** It is composed of a network of nodes in which each node has its own processing capabilities and there is no single point of data fusion. Therefore, each node fuses its local information with the information that is received from its peers. Data fusion is performed autonomously, with each node accounting for its local information and the information received from its peers.
    c. **Distributed Architecture:** Measurements from each source node are processed independently before the information is sent to the fusion node; the fusion node accounts for the information that is received from the other nodes. Therefore, each node provides an estimation of the object state based on only their local views, and this information is the input to the fusion process, which provides a fused global view.
    d. **Hierarchical Architecture:** Other architectures comprise a combination of decentralized and distributed nodes, generating hierarchical schemes in which the data fusion process is performed at different levels in the hierarchy.

MACHINE VISION FOR SPATIAL COORDINATE MEASUREMENT

Machine vision methods to provide spatial coordinates measurement has developed in a wide range of technologies, the computer vision has guide the machine vision to the tendency of duplicate the abilities of human vision by electronically perceiving and understanding an image with a high a dimensional data, increasing the data storage requirement, and the time processing due to the complexity of algorithms to extract important patterns and trends and understand what data says (learning data). For some applications it is not necessary to acquire the whole vision of a scene as makes the human vision to extract the features of interest (Luo, 2002), it is possible to only extract the significant characteristics from the scene by the use of a positioning laser (PL), a rotatory mirror scanner and dynamic triangulation as shown at Figure 7, for detailed system overall see (Basaca-Preciado, 2014; Rodriguez-Quiñonez, Sergiyenko, Preciado, et al, 2014; Flores-Fuentes, 2014a).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Sensors and data fusion methods

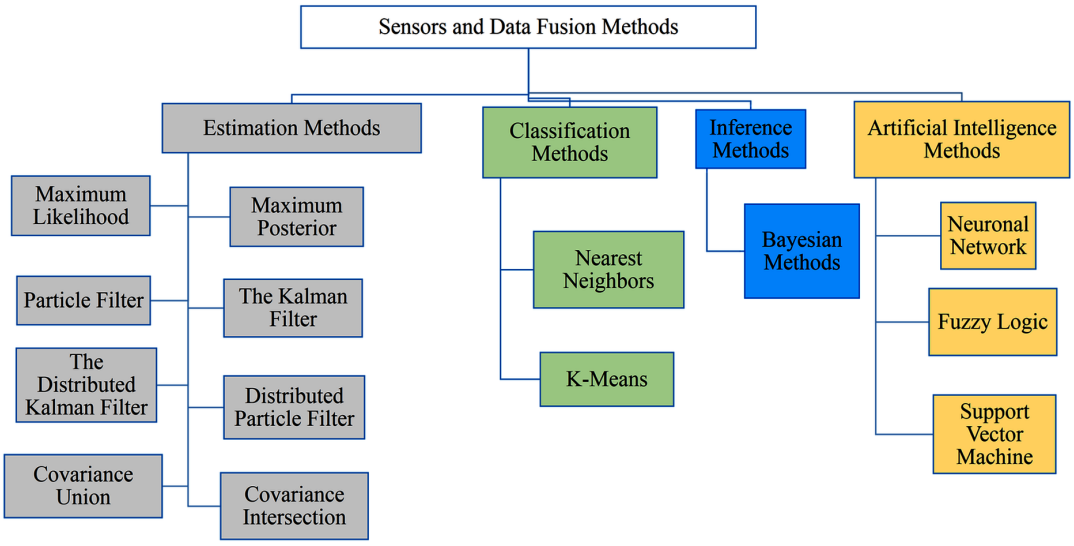
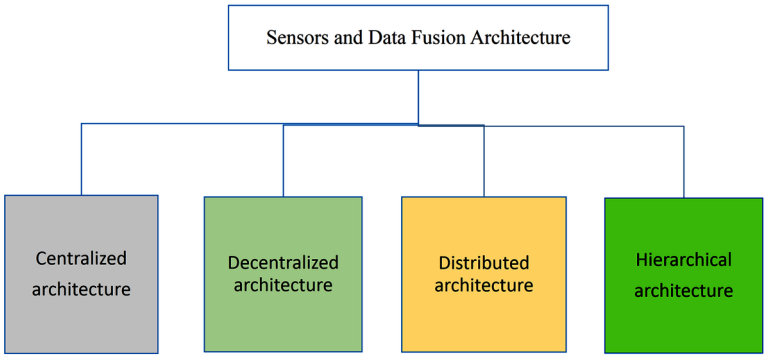


Figure 6. Sensors and data fusion architectures



Machine Vision System Overall

The main sensor is a light scanning system, it can be a coherent light emitter source such as a laser or an incoherent light source such as a bulb like the ones used in motor vehicles. We assume that for any light emitter source there is only one energy centre that represents its point position.

A rotating optical aperture, designed as a 45° slanted mirror surface attached to a cylindrical rod that deflects the light beam into a double convex lens provided with an interference filter and a photodiode. The cylindrical rod is mounted onto the shaft of a dc electrical motor, and as it rotates, an electronic signal is generated. Figure 8 shows the main elements of the optical scanning aperture. When the mirror starts spinning, the sensor "s" is synchronized when the origin generates a pulse indicating the 0° position and the starting of a cycle of 360° that finishes immediately before the "s" sensor generates the next beginning pulse. These pulses are used to calculate the scanning frequency and the zero reference, which is used to measure the angle where the light emitter source is positioned.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. 3D laser scanning system for spatial coordinates

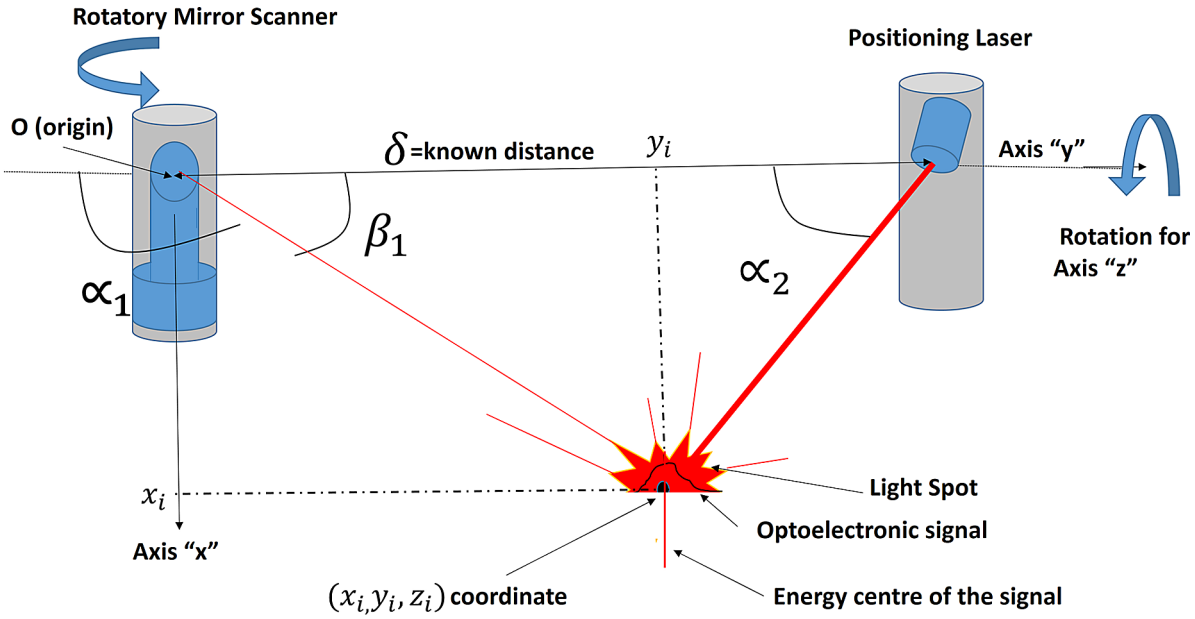
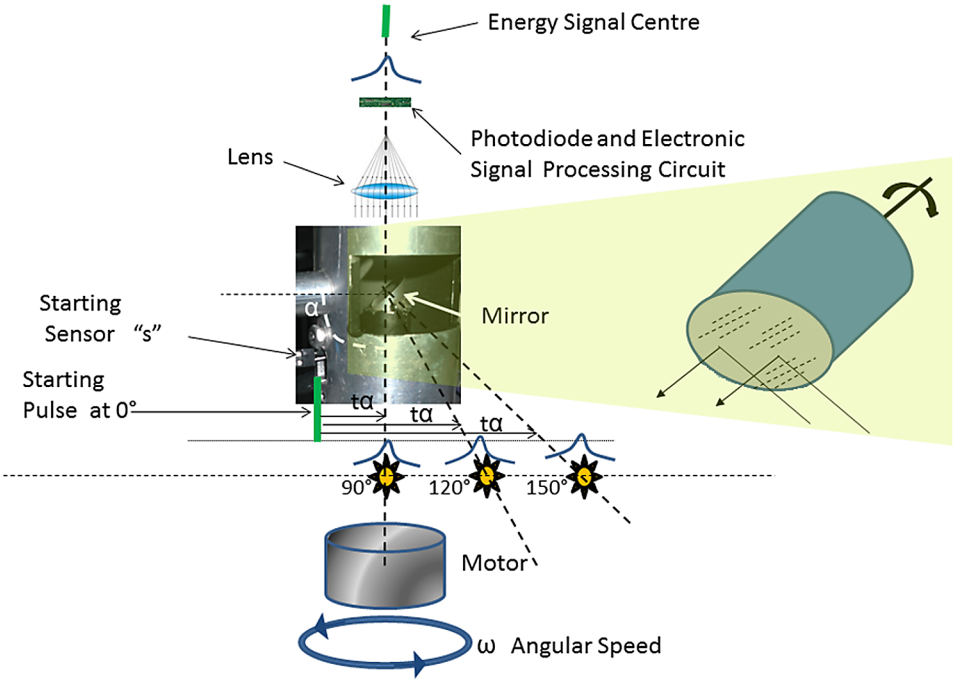


Figure 8. Optical scanning aperture



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The signal timing diagram in Figure 9 shows the starting pulse and the optoelectronic signal relation used to calculate the light emitter energy signal centre position as described in the equations below. The interval $T_{2\pi}$ is the time for one motor revolution, defined as the time between m_1 and m_1 as in (1), which are expressed by the code $I_{2\pi}$ as defined in (2). As for m_1, it is the reference of the starting of one motor revolution (0°).

$$T_{2\pi} = \frac{2\pi}{\omega} = \frac{2\pi}{2\pi f} = \frac{1}{f} \tag{1}$$

$$I_{2\pi} = T_{2\pi} f \tag{2}$$

The time t_α is the time between the reference of the starting of one motor revolution and the energy signal centre, that is, the interval between m_1 and m_2. This can be expressed by the code I_α as defined in (3). As for m_2, it is the energy signal centre. The angle under measurement is then calculated by (4).

$$I_\alpha = t_\alpha f \tag{3}$$

$$\alpha = \frac{2\pi I_\alpha}{I_{2\pi}} = \frac{2\pi t_\alpha f}{T_{2\pi} f} = \frac{2\pi t_\alpha f}{\frac{2\pi}{\omega} f} = t_\alpha \omega = t_\alpha 2\pi f \tag{4}$$

where:

m_1 is the reference of the starting of one motor revolution.
m_2 is the energy signal centre.
$T_{2\pi}$ is the time interval of one cycle, from m_1 to m_1.
ω is the angular speed
f is the scanning frequency (cycles per second).
$I_{2\pi}$ is the interval code of one cycle, from m_1 to m_1.
I_α is the range code from the starting of one cycle to the energy signal centre.
t_α is the interval from the starting of the signal to the energy signal centre, from m_1 to m_2.

Applications of Machine Vision for Spatial Coordinate Measurement.

The present chapter surged in the research continuity of a 3D Vision System for mobile robot navigation application, a 3D medical laser scanner, and a structural health monitoring system, shown at Figure 10.

1. **Mobile Robot Navigation:** Mobile Robot Navigation, also known as Navigation System (NS) utilizes Technical Vision System (TVS) to obtain the characteristics of obstacles appearing in the field of view (FOV) of a mobile robot; with that information the NS calculates a new trajectory to reach its programmed goal. The mobile robot needs to have a goal programmed to begin operation, then an initial trajectory is calculated by calculating the shortest path between current position and goal. The mobile robot will follow the initial trajectory until an obstacle is detected by an array of

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Optical scanning aperture timing

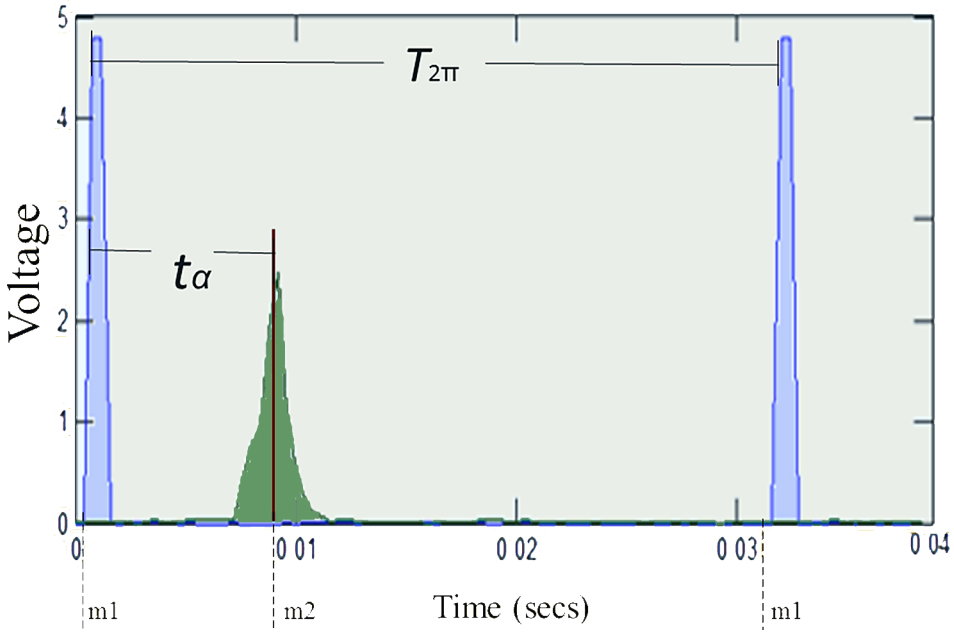
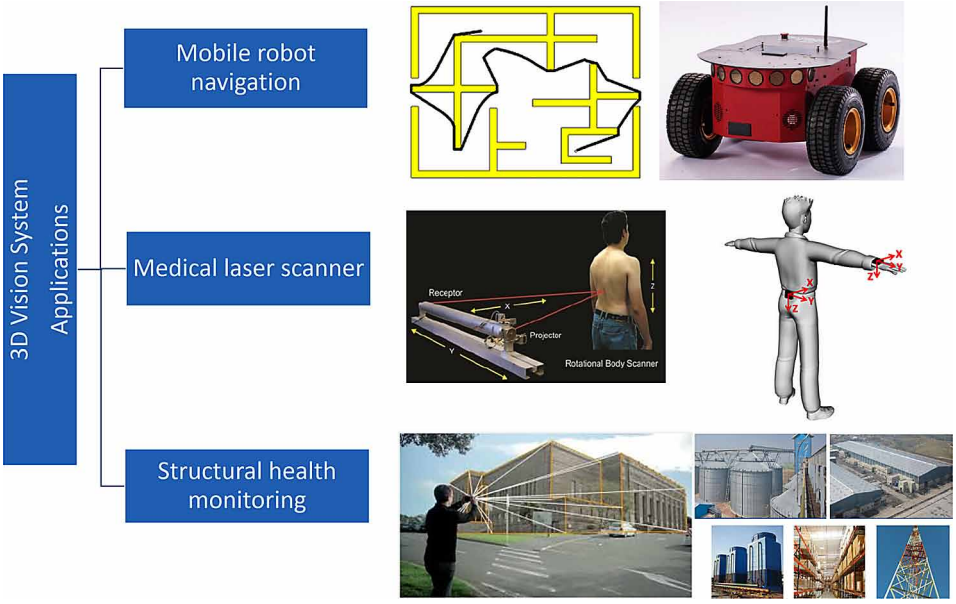


Figure 10. Applications of machine vision for spatial coordinate measurement



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

infrared (IR) sensors. The IR sensors signal will trigger the TVS to perform a scan and obtain the characteristics (3D coordinates) of the detected obstacles. If no obstacles are detected, the mobile robot will reach its goal and stop advancing.

This NS is based in the Pioneer 3-AT (Figure 11) which is a small, four-wheel, four-motor skid-steer intelligent robotics platform, with all the characteristics to be a very suitable research development platform to implement the navigation and 3D vision system for autonomous mobile robots (Basaca-Preciado, 2014).

2. **3D Medical Laser Scanning:** For Medical Scanning System, the 3D measurements of the human body surface or anatomical areas have gained importance in many medical applications. Medical professionals widely use size and shape to assess nutritional status and developmental normality and to calculate the requirements of drug, radiotherapy, and chemotherapy dosages, as well as the production of prostheses. This system has the capability to realize precise measurements by 3D point clouds sampled from the surface of human body, used to obtain the 3D coordinates of the body that are in the FOV in front of the system as shown in Figure 12. This system is based in medical professional traditionally manual measurement of the body size and shape to assess health status and guide treatment (Rodriguez-Quiñonez, Sergiyenko, Preciado, et al, 2014).
3. **Structural Health Monitoring System:** Important engineering constructions as shown at Figure 13, require geometrical monitoring to predict their structural health during their lifetime. Structure experiences deterioration and damage due to environmental conditions and excessive load conditions. Optical Scanning Systems (OSS) provides position measurements for structural health monitoring (SHM) task by a spatial coordinates measurements. This system is based on an initial prototype consisting of an incoherent light source emitter (non-rotating) mounted on the structure under monitoring, a passive rotating optical aperture sensor designed with a 45°-sloping mirror, and embedded into a cylindrical micro rod. A photodiode to capture the beam of light while the cylindrical micro rod mounted on a dc electrical motor shaft is rotating, and a signal energetic centre detector by digital and analog processing (Flores-Fuentes, 2014a), (Flores-Fuentes, 2014b).

Optoelectronic Scanning and Digital Image Processing through the Energetic Center of Laser Spot

Optoelectronic scanning systems generate different shapes of signals, depending on the light source and the sensor of the scanner. A typical position measuring process includes an emitter source of light, as a laser diode or an incoherent light lamp and the position sensitive detector like photodiode and CCD as a receiving device, which collects a portion of the back-reflected light from the target.

When a photodiode is used as a sensor on a scanner with rotating mirror it can also originate a similar Gaussian-like shape with some noise and deformation, the spatial distribution function of light has an Airy-function-like shape, by the other side the CCD produces output currents related to the “centre of mass” of light incident on the surface of the device, CCD use the light quantity distribution of the entire beam spot entering the light receiving element to determine the beam spot centre or centroid and identifies this as the target position.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Pioneer 3-AT mobile robotic platform



Figure 12. 3D rotational body scanner

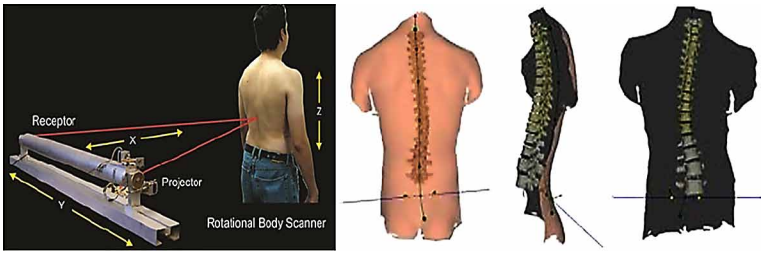


Figure 13. Structures (historical center, skyscraper, suspension bridge, dam water retention, antenna, and industrial warehouses)



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

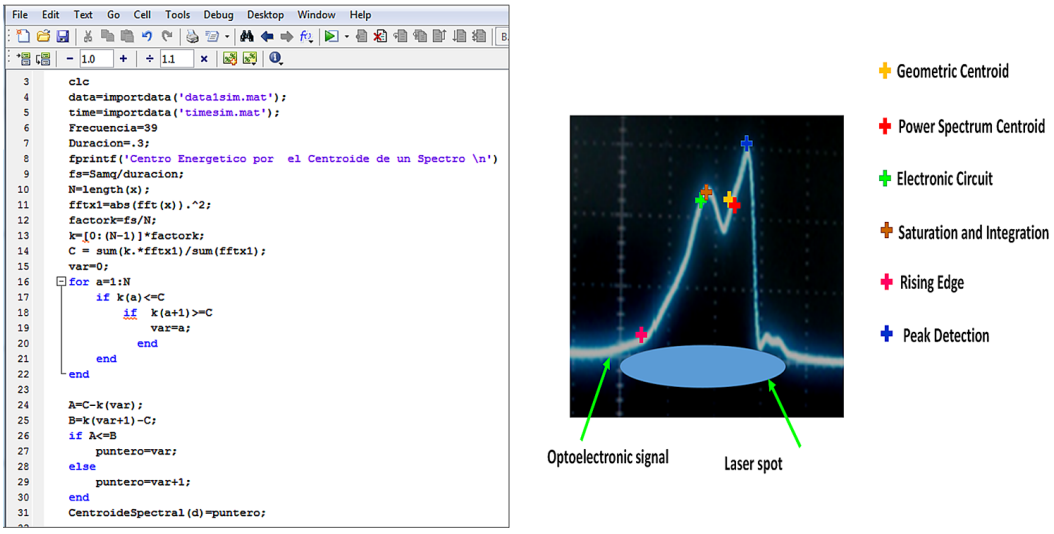
Concluding that, the photodiode signal originates a similar function to a CCD, consequently, it is possible to enhance the accuracy measurements in optoelectronic scanners with a rotating mirror, using a method for improving centroid accuracy by taking measurement in the energy centre of the signal (Sonka, 2014), (Li, 2014).

1. **Energetic Center of Laser Optoelectronic Signal Detected by Photodiode:** The photodiode is the sensor, which is used for laser-scanning (Beiser, 1995). As SHM is an upcoming tendency of determining the integrity of structures and development of strategies to prevent undesirable damage, the OSS was designed, it has the necessity off detect a light emitter mounted on the structure under monitoring and calculate its energy centre localization. The OSS generates the targeted signals to be analyzed by the proposed methods previously assessed (Rivas-Lopez, 2014).

These six Energy Centre Localization Methods (ECLM), (Geometric Centroid, Power Spectrum Centroid, Analog Processing by Electronic Circuit, Saturation and Integration, Rising Edge, and Peak Detection), Figure 14 shows MATLAB code for Power Spectrum Centroid y graphical results of six ECLM. ECLM were assessed based on the assumption that the OSS signal from light emitter scanning is a Gaussian-like shape signal, the light emitter is an incoherent light, considered a punctual light source and due to the fact that with distance the light source expands its radius a cone-like or an even more complex shape is formed depending on the properties of the medium through which the light is travelling. And to reduce errors in position measurements, the best solution is taking the measurement in the energy centre of the signal generated by the OSS.

According with the results from previous assessment made with incoherent light source the most accurately ECLM are the Geometric Centroid and the Power spectrum centroid, by this reason the Geometric Centroid ECLM was selected to perform the energetic centre localization of a laser spot to measure laser beam deviation.

Figure 14. ECLM graphic representation



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. **Energetic Center of Laser Image Signal Detected by CCD:** The CCD is the sensor who gives place to the close-range photogrammetry (Franca). Photogrammetry gives information about the geometry of an object or surface through the use of photographs, provides less dense surface information but with high quality, especially along object space discontinuities, edges and borders, it can be defined as the automation of information extraction from digital images, based on image analysis methods.

A camera is located parallel to the OSS photodiode to obtain a second source of information regarding angle β_1, by using a CCD sensor to acquire an image of the FOV at the same time as the OSS photodiode. The use of two methods has been proposed, one focused in the determination of edges and the shape centroid and other focused in the density of points measured on the surface of laser reflection. Both of them calculate the energetic center (image matrix centroid) applying morphological operations. Mathematical morphology is a tool for extracting image components that are useful in representation and description of region shape, it works directly on spatial domain, is a non-linear approach for detection of edges, by image transformations based on simple expanding and shrinking operations. Edge detection identifies and locates discontinuities in an image. An edge may be regarded as boundary between two dissimilar regions in an image, classical methods of edge detection involve the convolution of an image with an operator (like Canny and Prewitt), which is developed to be sensitive to gradients in the image and returning values of zero in uniform areas.

In the method focused in the determination of edge and the shape centroid, a laser light beam is projected in a desired coordinate to create an edge, then the laser light scattered reflection is used to generate a minimum illumination difference (MID), to isolate only the laser light reflection from the image and to detect the edges between the two images (with/without laser light beam), a continuity the edge location is performed by an operator like Canny or Prewitt. Second, the laser light reflection energetic center is localized by the application of several morphological operations as dilation, erosion, boundaries location, fill image regions and remove open area (Li, 2013). MATLAB code and results in Figure 15.

In the method focused in the density of points of laser reflection, the first step the image is captured, and the red color spectrum is obtained, a Gaussian blur filter (5x5 averaging matrix) is applied to smooth the captured image. By calculating the convolution of the original image and a 5x5 averaging matrix, is obtained the desired smoother image which allows to find the center of the laser spot more precisely. Straightforward formal implementation of the two-dimensional convolution equation in spatial form is used, calculated for finite intervals, defined by image frame limits.

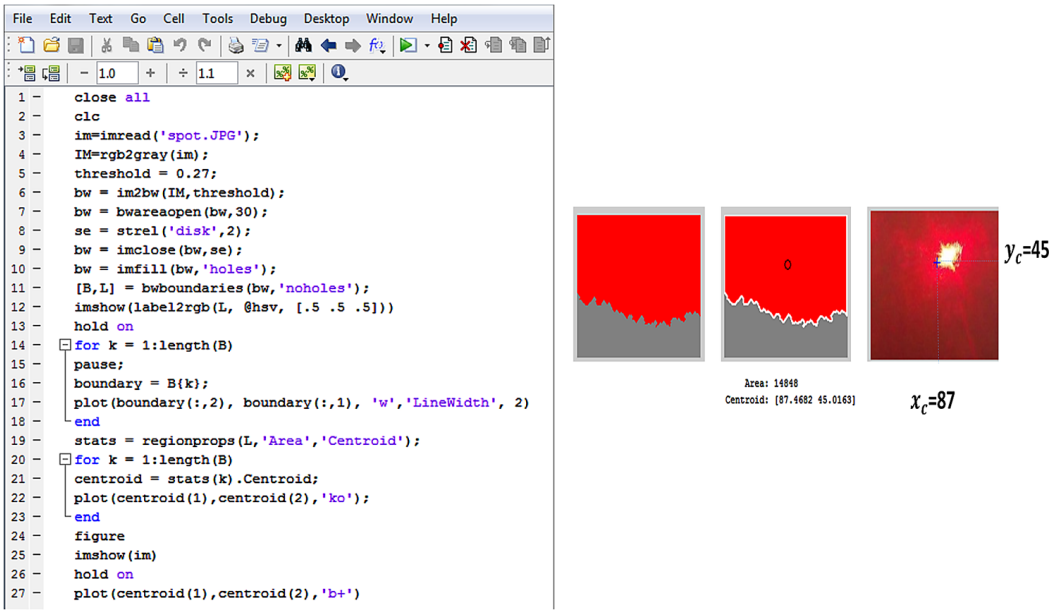
After the image has been blurred the maximum values of red color are calculated for each row (x) and column (y) of the 2D matrix representing the image; the brighter the red color in the image, the higher the red value.

The final results are the x and y coordinates of the center of the laser spot inside the captured image (i.e. these coordinates instantly represented in pixels and further, converted to X and Y) (Damelin, 2012; Zhang, 2012). MATLAB code and graphical results are shown in Figure 16.

These two Energy Centre Localization Methods (ECLM) for CCD Images have not been formerly evaluated to compare them, due to the measurement quality depends on surface reflection properties and lighting conditions, the surface reflection properties are dictated by a number of factors, as a) angle of the laser ray hitting, b) surface material, and c) roughness. Although for this fusion sensor purpose, the method focused in the density of points of laser reflection has been selected to be used during the experimentation.

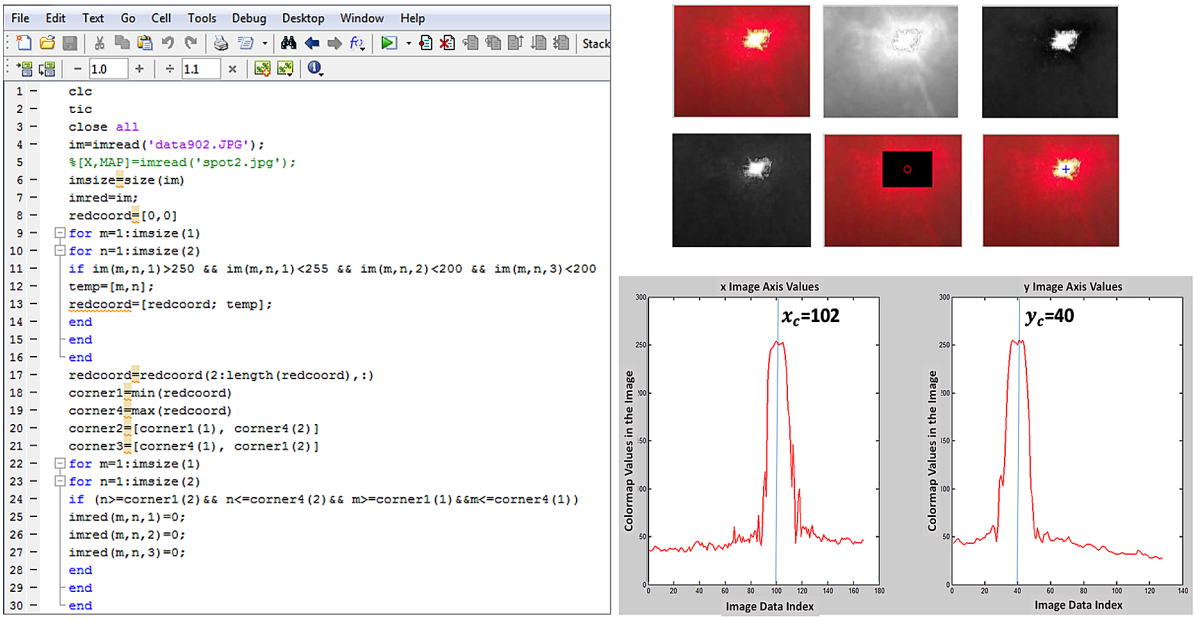
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 15. Graphic image processing stages by method focused in the determination of edge and the shape centroid



```
close all
clc
im=imread('spot.JPG');
IM=rgb2gray(im);
threshold = 0.27;
bw = im2bw(IM,threshold);
bw = bwareaopen(bw,30);
se = strel('disk',2);
bw = imclose(bw,se);
bw = imfill(bw,'holes');
[B,L] = bwboundaries(bw,'noholes');
imshow(label2rgb(L, @hsv, [.5 .5 .5]))
hold on
for k = 1:length(B)
pause;
boundary = B{k};
plot(boundary(:,2), boundary(:,1), 'w','LineWidth', 2)
end
stats = regionprops(L,'Area','Centroid');
for k = 1:length(B)
centroid = stats(k).Centroid;
plot(centroid(1),centroid(2),'ko');
end
figure
imshow(im)
hold on
plot(centroid(1),centroid(2),'b+')
```


Figure 16. Graphic image processing stages by method focused in the density of points of laser reflection



```
clc
tic
close all
im=imread('data902.JPG');
%[X,MAP]=imread('spot2.jpg');
imsize=size(im)
imred=im;
redcoord=[0,0]
for m=1:imsize(1)
for n=1:imsize(2)
if im(m,n,1)>250 && im(m,n,1)<255 && im(m,n,2)<200 && im(m,n,3)<200
temp=[m,n];
redcoord=[redcoord; temp];
end
end
end
redcoord=redcoord(2:length(redcoord),:)
corner1=min(redcoord)
corner4=max(redcoord)
corner2=[corner1(1), corner4(2)]
corner3=[corner4(1), corner1(2)]
for m=1:imsize(1)
for n=1:imsize(2)
if (n>=corner1(2)&& n<=corner4(2)&& m>=corner1(1)&&m<=corner4(1))
imred(m,n,1)=0;
imred(m,n,2)=0;
imred(m,n,3)=0;
end
end
end
```


EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Photodiode and CCD Fusion for Machine Vision

The photodiode and CCD sensors fusion is the base for those systems with non-lineal behavior that require the use of information from various sensors for pattern recognition to produce an enhanced output. The goal of using data fusion in multisensor environments is to obtain a lower detection error probability and a higher reliability by using data from multiple distributed sources. Artificial intelligence methods for sensor and data fusion require human reasoning such as pattern recognition, planning, deduction and learning. The processes of inference used by expert systems begin with a group of initial data (data prior) and some basic rules. Machine Learning algorithms are investigated to fuse the sensors data, such as Support Vector Machine (SVM). The best model generated by each algorithm is called estimator. It is shown that the employment of estimators based on artificial intelligence can improve significantly the performance achieved by each sensor alone (Faceli, 2004; Banerjee, 2012; Waske, 2007).

In previous OSS experimentation angle β_1 measurements with photodiode has been done through an optical table, to observe the measurement error through the FOV, a total of 6020 measurements built the dataset that was used to train and to test an SVM algorithm in the error prediction, to measurement corrections of angle β_1, obtaining a satisfactory improvement, increasing from 53.39% to 97.60% the percentage of measurements with error lower than 2.9°. Corrected angles β_1 were used to calculate the spatial coordinates by dynamic triangulation. At the same time Navigation and Medical Scanning Systems experimentations were applying the Feed Forward back propagation neural networks (NN) Levenberg−Marquardt algorithm with Widrow-Hoff Height/bias learning function to predict the error of spatial coordinates previously calculated with measurement of angle β_1, NN provided an acceptable approximation of the measurement error up to 99.97%, however, the validation of the method shows only a 96% of performance.

For sensors and data fusion experimentation the designed process explained at Figure 17 was followed. A total of 80 measurements built the new experimentation database of 80 instances (*n* rows) and 6 variables (*p* columns), the database values corresponding for column *p*=1 are the real objective angle under measurement (β_{1R}), for *p*=2 are the real objective distance under measurement, for *p*=3 are the system scanning motor frequency, for *p*=4 are the angle calculation from photodiode output ($\beta_{1photodiode}$), for *p*=5 are the angle calculation from CCD output (β_{1CCD}), and for *p*=6 are the error measurement (*E*) calculated by (5).

$$E = \beta_{1R} - \beta_{1M} \tag{5}$$

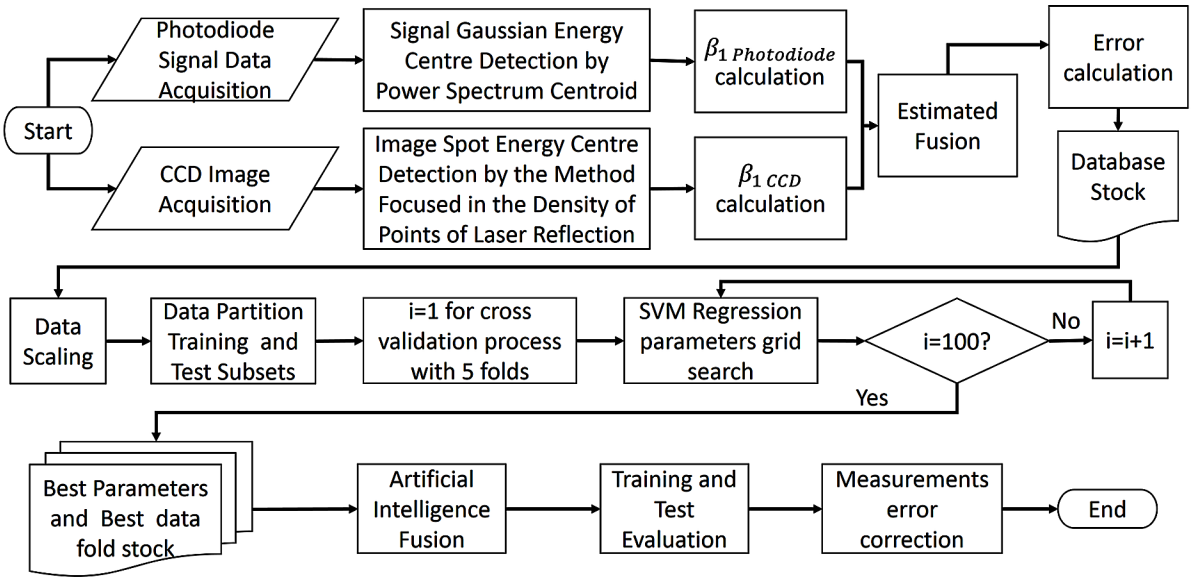
where:

E is the measurement error, representing how far the measurement from the real value is.
β_{1R} is the angle real value.
β_{1M} is the angle measured by the system (estimated fusion of photodiode and CCD by weighted average).

A scaling step on each column took place (from -1 to 1). The scaled database was partitioned in two sets, 80% training dataset and 20% test dataset, obtaining a training dataset of 64 rows with 6 columns and a test dataset of 16 rows with 6 columns. The 5 first columns are indicated as attributes and the last column as target.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 17. Sensors and data fusion experimentation flowchart



In order to fusion data by artificial intelligence, a properly validate the SVM regression performance was executed, the well-known k-fold cross-validation method was executed, repeating 100 times a 5-fold and in each run performing grid search parameters, as illustrated at Figure 18.

The following kernel parameters were analyzed:

- **Kernel Type:** [Polynomial, RBF],
- **Gamma:** [2^{-5}, 2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}, 2^{0}, 2^{1}, 2^{2}, 2^{3}, 2^{4}, 2^{5}],
- **C:** [2^{-5}, 2^{-4}, 2^{-3}, 2^{-2}, 2^{-1}, 2^{0}, 2^{1}, 2^{2}, 2^{3}, 2^{4}, 2^{5}],
- **NU:** [0.015, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9],
- **Polynomial Degree in Polynomial Kernel:** [1, 2, 3].

EXPERIMENTAL RESULTS

By implementing the method of finding the energetic center of the laser spot by CCD, a second value of the β_1 angle was obtained, the average of that value and the previously corrected angle by artificial intelligence algorithm generated a new value for the β_1 angle, being the first fusion sensor experimentation.

Since the first experimentations for β_1 angle error correction included the known values of frequency (scanner rotation), angle and distance of the scanning aperture (SA) and the angle of incidence β_1 of the laser light on the phototransistor as input data (attributes) for the intelligence algorithm, this new experimentation, utilizes the intelligence algorithm SVMR as a fusion sensor technique, to include as attributes for the training process the β_1 angle measured for both sensors, photodiode and CCD in order to obtain the best sensors characteristics fusion. Showing results illustrated on Table 1.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 18. Flowchart critical steps illustration

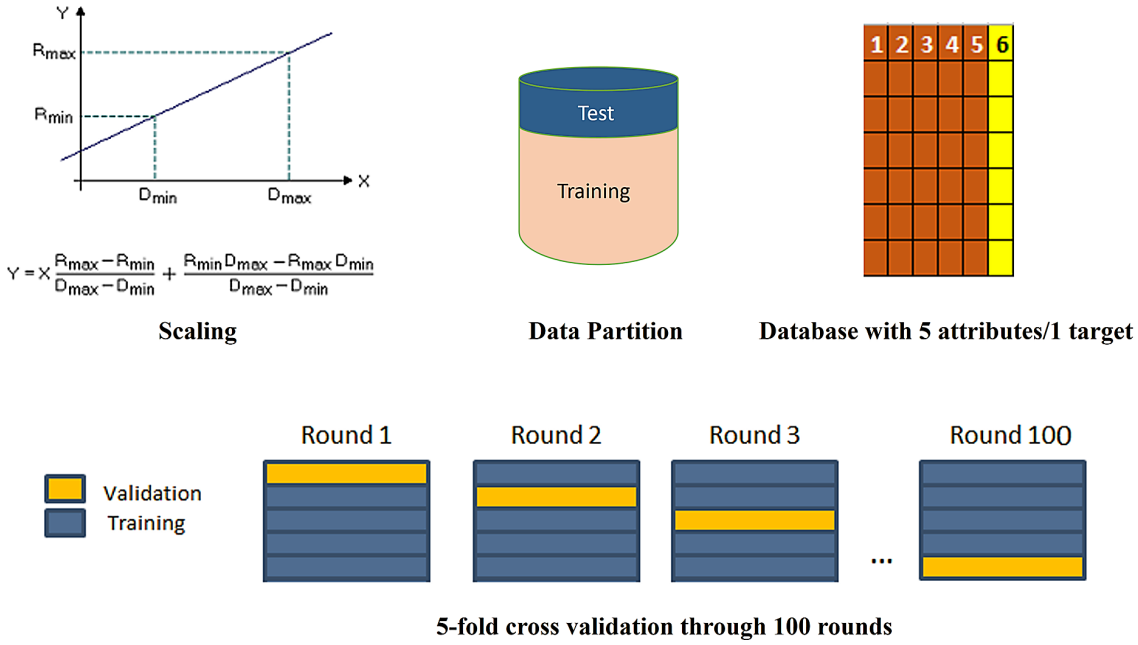


Table 1. Photodiode and CCD fusion sensor

Target	Photodiode Result	CCD Result	SVM Prediction	Fusion Result Error
90°	89.4°	91.1°	90.35°	0.35°
95°	98.1°	99.9	95.60°	0.60°
100°	99.4°	101.0°	100.08°	0.08°
⋮	⋮	⋮	⋮	⋮
105°	103.2°	104.0°	104.60°	0.40°
110°	108.0°	108.0°	109.80°	0.20°

The SVMR algorithm has demostrated been an efficient intelligent tool for prediction, the change of attributes for the target prediction to include the CCD sensor, has improved the β_1 angle measurement accuracy (with error lower than 1°), for the spatial coordinate dynamic triangulation calculation.

CONCLUSION

The goal of using sensor fusion in Machine Vision for Spatial Coordinates Measurement is to obtain a lower measurement error probability and a higher reliability by using data from multiple distributed sources (signal energetic center of sensors, previously corrected by artificial intelligence). This method implements the use of a second value of a measured variable to be used as an atributte in SVMR prediction, which directly affects the overall system accuracy and consequently the monitoring, navigation and scanning systems performance.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Due to the measurement quality depends on surface reflection properties and lighting conditions, the surface reflection properties are dictated by angle of the laser ray hitting, and surface roughness material, further experimentation will be realized scanning over different surface materials like wood, skin, paper, paperboard, egg carton, etc, to build a database to train the system to be able to operate in almost any environment, installed in an optimal hardware architecture in order to further increase the system accuracy and expand the optical scanning applications.

The author's project consists of three main research stages (Laboratory, Outdoor, and Field). The design of experiment and measurements in this chapter describes the results of the Laboratory research stage, and takes them as background for Outdoor research stage, where the fusion of 2D/3D Image and Range Scanner Datasets will be done with data from real environment measurements, while the Field research stage summarize future challenges for accomplish the project scope of SHM for heritage buildings in Autonomous University of Baja California at Mexicali city. In the Laboratory Research Stage, we started with the build of the one OSS for 3D measurements, the use of one OSS provides of polar coordinates in a plane, the use of two OSS separated at known distances, provide of a spatial coordinate in a plane by dynamic triangulation calculations, while the use of three OSS provides the full scope of 3D spatial coordinates measurement. In this stage one OSS was built and characterized at laboratory environment. During tests has been observed that it presents a non-linear behavior throughout all scanning scene, and in order to obtain an accuracy increase the requirement of computational intelligence techniques support surged, for pattern recognition to produce an enhanced output in a modern approach to the prediction of the non-linear measurement error of the OSS to compensate it. The outcome of our project has been a prototypal system for supporting an expert user in conducting structural analyses on building by extracting spatial coordinate information from light emitters mounted over a structure as indicator, in order to evaluate any structural displacement. Users may thus monitor the state of previous interventions, reason about the stability and integrity of the structure, identify potential risks, and plan reinforcement activities accordingly. Besides the specific scenario considered here, such a system may find application in several areas, as for instance in a museum scenario, bridge, damp, industrial structures or any other historical building monitoring projects. Concluding that this chapter outlines the proposed strategy by Artificial Intelligent Algorithms, and the capability to understand images and range scanner data through their fusion.

Future challenges for accomplish the project scope of SHM for buildings are:

- Incorporate a third aperture to the OSS, the use of three apertures in the OSS provides the full scope of 3D spatial coordinates measurement.
- Perform real time measurements by the integration of the signal processing stages (acquisition, energy signal centre calculations, error compensations, fusion and triangulation).
- Integrate the OSS signal processing by hardware, such as Field Programmed Gate Array (FPGA) to satisfy the monitoring system requirements to continue their tasks with safety and high accuracy on temporary occlusion condition.
- Include adaptive elements to the signal processing task in order to enhance the system to be capable of performance at an unknown environment.
- Implement a communication system for measurements record by wireless.
- Evaluate the requirement of time processing reduction.
- Reduce the energy consumption of the OSS.
- Perform measurement analysis to identify structures natural patterns from damages.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Abidi, M. A. (1992). *Data fusion in robotics and machine intelligence*. Academic Press Professional, Inc.

Banerjee, T. P., & Das, S. (2012). Multi-sensor data fusion using support vector machine for motor fault detection. *Information Sciences*, *217*, 96–107. doi:10.1016/j.ins.2012.06.016

Basaca-Preciado, L. C.-Q.-L.-B., Sergiyenko, O. Y., Rodríguez-Quinonez, J. C., García, X., Tyrsa, V. V., Rivas-Lopez, M., ... Starostenko, O. (2014). Optical 3D laser measurement system for navigation of autonomous mobile robot. *Optics and Lasers in Engineering*, *54*, 159–169. doi:10.1016/j.optlaseng.2013.08.005

Beiser, L. (1995). Fundamental architecture of optical scanning systems. *Applied Optics*, *31*(34), 7307–7317. doi:10.1364/AO.34.007307 PMID:21060601

Castanedo, F. (2013). A review of data fusion techniques. *The Scientific World Journal*, 1–19. PMID:24288502

Dagenais, M. a. (1995). *Integrated optoelectronics*. Academic Press.

Damelin, S. B. (2012). *The mathematics of signal processing*. Cambridge University Press.

Dasarathy, B. V. (1997). Sensor fusion potential exploitation-innovative architectures and illustrative applications. *Proceedings of the IEEE*, *1*(85), 24–38. doi:10.1109/5.554206

Elfes, A. (1992). Multi-source spatial data fusion using Bayesian reasoning. *Data fusion in robotics and machine intelligence*, 137-163.

Faceli, K., de Carvalho, A. C. P. L. F., & Rezende, S. O. (2004). Combining intelligent techniques for sensor fusion. *Applied Intelligence*, *3*(20), 199–213. doi:10.1023/B:APIN.0000021413.05467.20

Flores-Fuentes, W. L.-N.-C.-B.-Q., Rivas-Lopez, M., Sergiyenko, O., Gonzalez-Navarro, F. F., Rivera-Castillo, J., Hernandez-Balbuena, D., & Rodríguez-Quiñonez, J. C. (2014a). Combined application of power spectrum centroid and support vector machines for measurement improvement in optical scanning systems. *Signal Processing*, *98*, 37–51. doi:10.1016/j.sigpro.2013.11.008

Flores-Fuentes, W. L.-Q.-B.-C., Rivas-Lopez, M., Sergiyenko, O., Rodriguez-Quinonez, J. C., Hernandez-Balbuena, D., & Rivera-Castillo, J. (2014b). Energy Center Detection in Light Scanning Sensors for Structural Health Monitoring Accuracy Enhancement. *Sensors Journal, IEEE*, *7*(14), 2355–2361. doi:10.1109/JSEN.2014.2310224

Franca, J. G. (2005). A 3D scanning system based on laser triangulation and variable field of view. *Image Processing, 2005. ICIP 2005. IEEE International Conference on*. IEEE.

Fukuda, M. (1999). *Optical semiconductor devices*. John Wiley and Sons.

Gonzalez-Aguilera, D. L.-G. (2010). *Camera and laser robust integration in engineering and architecture applications*. INTECH Open Access Publisher. doi:10.5772/9959

Hall, D. L., & Llinas, J. (1997). An introduction to multisensor data fusion. *Proceedings of the IEEE*, *1*(85), 6–23. doi:10.1109/5.554205

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hou, B. (2011). Charge-coupled devices combined with centroid algorithm for laser beam deviation measurements compared to a position-sensitive device. *Optical Engineering (Redondo Beach, Calif.)*, *50*(3), 033603–033603. doi:10.1117/1.3554379

Janesick, J. R. (2001). *Scientific charge-coupled devices.* SPIE Press.

Klein, L. A. (2003). Sensor and data fusion: a tool for information assessment and decision making. Bellingham, WA: SPIE Press.

Labayrade, R., Royere, C., Gruyer, D., & Aubert, D. (2005). Cooperative fusion for multi-obstacles detection with use of stereovision and laser scanner. *Autonomous Robots*, *2*(19), 117–140. doi:10.100710514-005-0611-7

Li, F. a. (2006). *CCD image sensors in deep-ultraviolet: degradation behavior and damage mechanisms*. Springer Science and Business Media.

Li, F. M. (2003). Degradation behavior and damage mechanisms of CCD image sensor with deep-UV laser radiation. *Electron Devices. IEEE Transactions on*, *12*(51), 2229–2236.

Li, L., Ma, G., & Du, X. (2013). Edge detection in potential-field data by enhanced mathematical morphology filter. *Pure and Applied Geophysics*, *4*(179), 645–653. doi:10.100700024-012-0545-x

Li, X., Zhao, H., Liu, Y., Jiang, H., & Bian, Y. (2014). Laser scanning based three dimensional measurement of vegetation canopy structure. *Optics and Lasers in Engineering*, *54*, 152–158. doi:10.1016/j.optlaseng.2013.08.010

Lorenser, D. A. (2003). Towards wafer-scale integration of high repetition rate passively mode-locked surface-emitting semiconductor lasers. *Applied Physics. B, Lasers and Optics*, *8*(79), 927–932.

Luo, R. C.-C., & Yih. (2002). Multisensor fusion and integration: Approaches, applications, and future research directions. *Sensors Journal, IEEE*, *2*(2), 107–119. doi:10.1109/JSEN.2002.1000251

Marston, R. M. (1999). *Optoelectronics circuits manual*. Butterworth-Heinemann.

McClure, W. F. (2003). 204 years of near infrared technology: 1800-2003. *Journal of Near Infrared Spectroscopy*, *6*(11), 487–518. doi:10.1255/jnirs.399

Piprek, J. (2003). *Semiconductor optoelectronic devices: introduction to physics and simulation*. Academic Press.

Rivas-Lopez, M. C.-F.-Q.-B.-B. (2014). Scanning for light detection and Energy Centre Localization Methods assesment in vision systems for SHM. *2014 IEEE 23rd International Symposium on* (pp. 1955-1960). Industrial Electronics (ISIE).

Rodriguez-Quiñonez, J. B.-L.-F.-P. (2014). Improve 3D laser scanner measurements accuracy using a FFBP neural network with Widrow-Hoff weight/bias learning function. *Opto-Electronics Review*, *4*(22), 224–235.

Rodriguez-Quiñonez, J. C., Sergiyenko, O. Y., Preciado, L. C. B., Tyrsa, V. V., Gurko, A. G., Podrygalo, M. A., ... Balbuena, D. H. (2014). Optical monitoring of scoliosis by 3D medical laser scanner. *Optics and Lasers in Engineering*, *54*, 175–186. doi:10.1016/j.optlaseng.2013.07.026

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Rogalski, A. (2002). Infrared detectors: An overview. *Infrared Physics & Technology*, *3*(43), 187–210. doi:10.1016/S1350-4495(02)00140-8

Shih, H.-C. P. (2015). High-resolution gravity and geoid models in Tahiti obtained from new airborne and land gravity observations: Data fusion by spectral combination. *Earth, Planets, and Space*, *1*(67), 1–16.

Snyder, W. E. (2010). *Machine vision*. Cambridge University Press. doi:10.1017/CBO9781139168229

Sonka, M. A. (2014). *Image processing, analysis, and machine vision*. Cengage Learning.

Syms, R. R. (1992). *Optical guided waves and devices*. McGraw-Hill.

Tavernier, F. A. (2011). *High-speed optical receivers with integrated photodiode in nanoscale CMOS*. Springer Science and Business Media. doi:10.1007/978-1-4419-9925-2

Tribolet, P. A. (2008). Advanced HgCdTe technologies and dual-band developments. *SPIE Defense and Security Symposium*. International Society for Optics and Photonics.

Vauhkonen, J. a. (2014). Tree species recognition based on airborne laser scanning and complementary data sources. In Forestry applications of airborne laser scanning (pp. 135-156). Springer. doi:10.1007/978-94-017-8663-8_7

Waske, B. a. (2007). Fusion of support vector machines for classification of multisensor data. *Geoscience and Remote Sensing*. *IEEE Transactions on*, *122*(45), 3858–3866.

Weckenmann, A. D.-R., Jiang, X., Sommer, K.-D., Neuschaefer-Rube, U., Seewig, J., Shaw, L., & Estler, T. (2009). Multisensor data fusion in dimensional metrology. *CIRP Annals-Manufacturing Technology*, *2*(58), 701–721. doi:10.1016/j.cirp.2009.09.008

Yallup, K. A. (2014). *Technologies for smart sensors and sensor fusion*. CRC Press.

Zhang, F.-M. A.-H.-H. (2008). Multiple sensor fusion in large scale measurement. *Optics and Precision Engineering*, (7), 18.

Zhang, Q., Su, X., Xiang, L., & Sun, X. (2012). 3-D shape measurement based on complementary Gray-code light. *Optics and Lasers in Engineering*, *4*(50), 574–579. doi:10.1016/j.optlaseng.2011.06.024

This research was previously published in Developing and Applying Optoelectronics in Machine Vision edited by Oleg Sergiyenko and Julio C. Rodriguez-Quiñonez; pages 1-37; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 11
Play in the Museum:
Design and Development of a Game-Based Learning Exhibit for Informal Science Education

Jonathan P. Rowe
North Carolina State University, USA

Bradford W. Mott
North Carolina State University, USA

Eleni V. Lobene
Aon Hewitt, USA

James C. Lester
North Carolina State University, USA

ABSTRACT

Digital games have been found to yield effective and engaging learning experiences across a broad range of subjects. Much of this research has been conducted in laboratory and K-12 classrooms. Recent advances in game technologies are expanding the range of educational contexts where game-based learning environments can be deployed, including informal settings such as museums and science centers. In this article, the authors describe the design, development, and formative evaluation of Future Worlds, a prototype game-based exhibit for collaborative explorations of sustainability in science museums. They report findings from a museum pilot study that investigated the influence of visitors' individual differences on learning and engagement. Results indicate that visitors showed significant gains in sustainability knowledge as well as high levels of engagement in a free-choice learning environment with Future Worlds. These findings point toward the importance of designing game-based learning exhibits that address the distinctive design challenges presented by museum settings.

INTRODUCTION

Over the past decade, growing evidence has emerged that games are effective learning tools for a broad range of subjects and student populations (Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012; Clark, Tanner-Smith, & Killingsworth, 2016; Wouters, van Nimwegen, van Oostendorp, van der Spek, 2013). Much of the work has focused on two categories of game-based learning: 1) games for formal educa-

DOI: 10.4018/978-1-7998-0951-7.ch011

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

tion settings such as schools (Habgood & Ainsworth, 2011; Wouters et al., 2013), and 2) serious games, which typically investigate game technologies for training (Johnson, 2010) or increasing awareness of social, geopolitical, or economic issues (Mitgutsch & Alvarado, 2012). While important, these research directions do not address a notable class of educational contexts that stands to benefit as much, or perhaps even more, from the introduction of well-designed educational games: informal education settings, such as science museums.

Formal education settings, and particularly schools, differ from science museums in several important respects (National Research Council, 2009). Schools and museums serve different populations: Schools are responsible for educating school-age children, whereas museums are visited by learners of all ages. Museums have a particular emphasis on affective outcomes, such as sparking interest and excitement, and they rarely use tests or grades. In schools, learners typically sit at designated desks for prescribed class periods. Museums are free-choice learning environments; visitors come and go as they please, and there is no teacher to provide instruction. Museum learning, quite literally, *looks different* from classroom learning. These distinctions have significant implications for the design of educational games—including pedagogical and gameplay designs—and they merit careful consideration to ensure effective and engaging learning experiences.

In this article, we explore the design of game-based learning environments for museums by presenting a case study of Future Worlds, a game-based learning environment for collaborative explorations of environmental sustainability. Future Worlds is an interactive exhibit designed to enrich museum visitors' understanding of sustainability, promote collaboration and scientific reasoning, and foster engagement in environmental science. We outline the design principles that drove the iterative development of Future Worlds, and we illustrate the methods utilized to investigate Future Worlds' effectiveness for fostering learning and engagement.

BACKGROUND AND RELATED WORK

The affordances of digital games align naturally with the goals of museum education, such as fostering engagement in science, and enabling learners to manipulate, test, and explore hypotheses about the natural world (National Research Council, 2009). However, designing game-based learning environments presents several challenges. Game design is multidisciplinary, requiring close collaboration between software developers, educators, artists, testers, and other specialists. Games are complex, requiring myriad design decisions with uncertain impacts on learners' experiences. Most notably, there is a dearth of evidence-based research on the design principles and methods necessary for creating effective game-based learning environments.

Reviews of the game-based learning literature have broadly concluded that games can yield positive learning outcomes across a range of educational subjects (Connolly et al., 2012). In recent years, a pair of prominent meta-analyses independently concluded that, in general, digital game technologies are often more effective than traditional instructional methods in fostering learning and retention (Clark et al., 2016; Wouters et al., 2013). Expanding on this conclusion, Wouters et al. (2013) advise, "the next step is more value-added research on specific game features that determine ... effectiveness" (p. 262). Clark et al. (2016) echo this argument, concluding, "[Research on game-based learning] should thus shift emphasis … to cognitive-consequences and value-added studies exploring how theoretically driven design decisions influence situated learning outcomes" (p. 116).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

With the aim of identifying evidence-based design principles for game-based learning, Isbister, Flanagan, and Hash (2010) conducted interviews with experienced game developers to identify key design practices used within the professional game industry. The interviewees described themes such as emphasizing fun as a central design value, requiring high levels of polish and well-tuned end-user experiences, emphasizing deep learning content rather than "bolted on" learning materials, supporting collaboration and specialization, designing for role-playing and emotional engagement, and including affordances for exploring complex systems. While the identified themes are high-level and abstract, they do describe characteristics often lacking in game-based learning environments. Isbister et al. (2010) account for this by pointing to problems in the communication pathways between practitioners and theorists, describing situations where design choices are not well articulated or discussed critically.

In other work, Linehan, Kirman, Lawson, and Chan (2011) describe methods for educational game design rooted in applied behavior analysis. Their framework emphasizes personalized instruction and mastery learning, describing a four-step design pattern that involves 1) defining and measuring learners' behavior, 2) recording and analyzing changes in learners' behavior, 3) providing immediate corrective feedback, and 4) tailoring game events based on learners' performance. It is notable that the guidelines from Linehan et al. are theoretically grounded, but unfortunately they are not examined in the context of an actual educational game.

A handful of empirical studies have been conducted to investigate the learning impacts of specific game design decisions. Habgood and Ainsworth (2011) found that tight integration of subject matter and game mechanics in a game for elementary mathematics yielded enhanced learning outcomes compared to versions that separated content and gameplay. Other work has investigated the impacts of narratives in games (Adams, Mayer, MacNamara, Koenig, & Wainess, 2012; Rowe, McQuiggan, Robison, & Lester, 2009). Results suggest that narratives may be effective in fostering student motivation, but narrative designs should be crafted sparingly in order to avoid distracting learners or imposing extraneous cognitive load.

Much of this research has taken place in the context of formal education settings, such as K-12 classrooms. However, the design requirements for game-based learning in museums differ from games in schools because of the distinct characteristics of museum-based learning. As outlined in a report from the National Research Council (2009), the informal science education community has identified six interrelated strands of informal science learning: sparking interest and excitement in science; enabling learners to understand, remember, generate, and use science concepts; encouraging learners to test, explore, question, and observe the natural and physical world; fostering reflection on the scientific process; creating opportunities for learners to participate in scientific activities; and encouraging learners to think of themselves as scientists.

Museums and schools share many of these objectives, but they prioritize the objectives differently. Schools are responsible for preparing students to meet academic standards focused on student knowledge, skills, and abilities. Attendance is often compulsory, and instructional practices are aligned with school objectives. In contrast, the voluntary nature of many museum visits leads toward an increased emphasis on non-cognitive outcomes, such as visitor interest and engagement. Museum visitors typically spend brief amounts of time at each exhibit—a few minutes or shorter—which limit expectations for deep knowledge gains. Museums also have different educational resources (e.g., panoramic theaters, fossil collections) and physical layouts (e.g., open floor plans) than schools, which shape the learning experiences that are available to visitors. Discerning the extent to which a game-based learning environment embodies the six strands of informal science learning, as well as the extent to which it accounts for the distinct capabilities and constraints of museum settings, provides a framework for evaluating a game-based learning exhibit's design.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Recent years have witnessed growing interest in the application of digital game technologies in informal settings, including animated pedagogical agents (Lane et al. 2013; Schulman, Sharma, & Bickmore, 2008), interactive tabletop-based games (Horn et al., 2012), and mobile games (Greenspan & Whitson, 2013). For example, advanced learning technologies in museums now utilize the high-end rendering capabilities of game engines so that virtual humans serve as simulated docents. These docents offer personalized guidance and feedback to visitors through natural language and affective expression (Lane et al., 2013). Game-based learning exhibits have also been developed for a range of subjects, including evolution (Horn et al., 2012), sustainability (Antle, Bevans, Tanenbaum, Seaborn, & Wang, 2011), and history (Greenspan & Whitson, 2013). These exhibits have shown promise for promoting engagement and visitor enthusiasm by leveraging game designs centered on brief visitor interactions, multi-touch interfaces, and multi-user hardware technologies (Antle et al., 2011; Horn et al., 2012;). However, there has been relatively little empirical research on visitor learning and individual differences with game-based exhibits in science museums. Further, few projects have explicitly examined the design factors that distinguish game-based learning in museums from game-based learning in schools. In the next section, we outline these design factors, which shaped the development of the Future Worlds exhibit.

DESIGN FACTORS FOR GAME-BASED LEARNING IN MUSEUMS

During the early stages of the project, we held a series of discussions with partners at the North Carolina Museum of Natural Sciences about interactive science exhibits and the dynamics of visitor learning. Further, we reviewed several interactive exhibits in prominent science museums across the eastern United States to inform our efforts to gather software requirements for the game's design. From this analysis, we identified five overarching design principles for Future Worlds to address the objectives, affordances, and constraints of game-based learning in museums (Table 1).

1. **Low barrier to entry**. Because interactions with museum exhibits can be either short or long in duration, embedding game mechanics that are accessible and quick-to-learn is critical. For this reason, we adopted natural user interfaces—in our case multi-touch interaction—and highly

Table 1. Alignment between museum-based learning attributes and game design principles

Museum-Based Learning Attributes	Design Principles for Game-Based Exhibits
Emphasize interest and excitement	• Foster exploration and curiosity • Present melodramatic feedback • Leverage novel hardware
Feature short dwell times	• Prioritize low barriers to entry
Appeal to diverse learners (i.e., age, gender)	• Prioritize low barriers to entry • Utilize inviting aesthetics with broad appeal
Support solo and collaborative interactions	• Leverage novel hardware • Provide flexible user experiences
Support active and passive engagement	• Leverage novel hardware • Provide flexible user experiences
Account for free-choice movement between exhibits	• Foster exploration and curiosity • Present melodramatic feedback

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

streamlined heads-up displays (HUD) to guide learner interactions. Similarly, we sought to leverage learners' prior knowledge whenever possible, adopting multi-touch controls that are widely adopted in modern computing (e.g., tap, swipe, pinch-to-zoom, two-finger rotation), as well as game paradigms that are widely recognized by the public (e.g., SimCity-style visual appearance).

2. **Exploration and curiosity**. Effective interactive exhibits typically avoid didactic instructional methods, such as lengthy text presentations and non-interactive animations. Our design, therefore, emphasizes learners' direct manipulation of the virtual environment, as well as explorations of cause-and-effect and interrelations between disparate environmental elements. We sought to minimize the amount of text in the game and avoid direct instruction about environmental science concepts.
3. **Immediate and dramatic feedback**. Immediate feedback is important for any educational setting, but discussions with museum partners yielded recommendations for incorporating melodramatic feedback (in their words, "use fireworks") to convey the effects of learners' decisions. This advice was specifically targeted at young learners, with the implication that subtle effects are often missed or ignored in museum-based learning. Museum spaces often house many competing exhibits, and learner attention is easily diverted. Therefore, we sought to integrate rich combinations of animation, color, and sound to make the effects of young learners' actions clear.
4. **Inviting visual aesthetics with broad appeal**. Although the project targeted young learners, parents and guardians often accompany these visitors. Given the broad range of potential participants found in science museums, we sought to devise a visual style that was inviting and had broad family appeal; we chose not to adopt a visual style that was edgy or dystopian in conveying the impacts of environmental decisions. While labor-intensive, we also sought to emphasize visual "polish," consistent with recommendations from the educational game design literature (Isbister et al., 2010). The importance of visual polish was particularly salient in our review of exhibits in prominent science museums. High-quality visual presentations are the standard, not the exception, in museum-based learning. This emphasis on visual quality is a notable differentiator from many game-based learning environments for classroom learning.
5. **Novel hardware platforms**. Given the leisure-based nature of many museum visits, and relatively high bar for learners' attention[1], we determined that using a novel hardware platform was an important opportunity for fostering learner interest and engagement. We decided to utilize surface computing hardware to create interactive learning experiences that were not feasible in most school or home settings. This criterion is contrasted with formal education settings, where even rudimentary games are often viewed as being more engaging than conventional instructional methods, such as lectures. Surface computing tables simultaneously integrate multi-touch interfaces, which are immediately familiar to many learners, with novel multi-user collaboration possibilities. We then extended the platform by integrating the surface computing table with a second, vertically mounted high-definition digital display, doubling the available screen real estate for the exhibit (Figure 1). One collaborator described the platform as akin to a "giant Nintendo DS."

In addition to these principles, we came to recognize several interaction patterns that distinguish museum-based learning from other settings, particularly classroom learning. These point to the need for a sixth design principle: providing flexible user experiences that accommodate the broad range of visitor interactions typified by museum-based learning.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Early artist rendering of Future Worlds multi-display exhibit



1. **Solo vs. collaborative interactions**. Learners approach museum exhibits in many different types of configurations. Learners may explore museums individually, engaging in solo learning experiences involving limited conversation and social interaction. Young learners may explore exhibits while accompanied by siblings, parents or guardians. Adults, who may or may not directly interact with the exhibit themselves, will often scaffold young learner's interactions by asking questions or giving advice. In another interaction pattern, learners may approach exhibits that are already partially occupied, engaging in collaborative or cooperative learning with "strangers" with whom they share no prior relationship. Alternatively, learners may approach exhibits as part of large field trip groups, exploring exhibits collaboratively with large groups of peers with whom they have strong social bonds. Given these distinct types of configurations, game-based learning designs for museums should be compatible with a broad range of solo and multi-player interaction paradigms.
2. **Short vs. long durations**. Learner interactions may be shallow and brief—5 second interactions that are quickly abandoned—or they may be deep and extensive—30 minute explorations involving detailed conversations and reflection. Paradoxically, learner interactions may also be long in duration but shallow in terms of learning. In these cases, learners play with surface features of the exhibit but fail to deeply engage with the subject matter. Gameplay and content in a museum-centric game-based learning environment should be devised to accommodate these different extremes of engagement, providing positive educational value in as many scenarios as possible.
3. **Active vs. passive engagement**. Active prolonged engagement is an oft-cited goal in museum exhibit design (Horn et al., 2012). However, in practice, learners may engage with an exhibit in many different ways. Some learners confidently approach an exhibit and begin physically interacting immediately. Others initially hold back, passively watching or deliberating whether to explore the exhibit more actively. Learners may also passingly engage with an exhibit while walking en route to another destination, judging whether to explore the exhibit at a later point in time. A salient indicator of these differences in engagement is physical proximity; active engagers are often

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

physically close to the exhibit, while passive engagers are often farther away. Another important indicator is dwell time, which denotes the amount of time that a visitor spends at an exhibit. In our work, we attend to both factors together, investigating how long visitors spend at several different tiers of proximity to the exhibit. Given individual differences in engagement, game-based learning exhibits should be designed to support learning across different physical distances and durations, gradually easing learners toward higher levels of active engagement as they are comfortable.

Having identified these design principles and museum-centric interaction patterns, we embarked on the design and development of a game-based learning environment for sustainability education in museums, Future Worlds.

DESIGN AND DEVELOPMENT OF FUTURE WORLDS

To create Future Worlds, we adopted an iterative design process that involved close interdisciplinary collaboration between developers (game designers, software engineers, digital artists), educators (informal science educators, elementary science educators), and subject matter experts (environmental engineers, environmental scientists). The game design model involved several iterations of designing, developing, evaluating, and refining prototypes of Future Worlds, including both paper and software-based prototypes.

Future Worlds

Future Worlds is a prototype game-based learning exhibit about environmental sustainability for children ages 9–12. The exhibit integrates turn-based strategy game mechanics, pedagogical agents, and surface computing tables to support collaborative explorations of environmental sustainability. In Future Worlds, learners solve sustainability-centered problems by investigating the impacts of alternate environmental decisions in a 3D simulated environment. Learners explore environmental decisions—such as modifying a region's electricity portfolio or a farm's waste management practices—through collaborative, multi-touch interactions on the exhibit's interactive tabletop display. The effects of learners' environmental decisions are realized in real-time through rich 3D virtual environments, and they are accompanied by narrated explanations from a virtual docent who observes the learners' actions.

The design of Future Worlds is intentionally aligned with several strands of informal science learning (National Research Council, 2009). Specifically, Future Worlds offers opportunities for enriching visitors' understanding of environmental sustainability; it fosters high levels of visitor engagement; it enables visitors to test, manipulate, and observe changes in the simulated environment; and it fosters collaboration and conversation about environmental science. Future Worlds also accounts for the distinctive capabilities and constraints of science museums. The game accommodates different levels of visitor engagement, ranging from brief passive encounters to extended hands-on experiences. Future Worlds' multi-touch gameplay is designed to be easily accessible, particularly for learners with minimal game-playing experience. Further, the game's 3D graphics are designed to appeal to a broad range of visitors.

The prototype exhibit's physical design is comprised of two adjacent digital displays: a horizontally oriented Samsung SUR40 interactive tabletop, and a vertically oriented non-interactive 50" high-definition television (Figure 1). The configuration provides an integrated two-screen setup; visitors congregate

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

around the horizontal display to explore the science simulation through multi-touch interactions, and the vertical display provides additional screen real estate for explanations of sustainability concepts, which are also accessible to learners standing farther away from the exhibit.

Future Worlds focuses on three integrated themes of sustainability: water, food, and energy. Visitors' objective during learning interactions with Future Worlds is to use the interactive tabletop display to collaboratively (or individually) reconfigure an unsustainable virtual environment into a sustainable environment. Learners can engage in deep, extended interactions with Future Worlds —solving sustainability problem scenarios and identifying complex relationships between alternate environmental decisions—or they can engage in shallow interactions with Future Worlds, tapping on the virtual environment and briefly listening to a narrated explanation before moving on.

As an illustration of a typical learning interaction with Future Worlds, consider the following scenario. A small group of learners approaches the exhibit. Tapping on the start screen causes a 3D model of Earth to rotate into view. From the globe, a small environmental region from the United States' eastern coast emerges. The virtual environment represents the first problem-solving scenario: a portion of a simulated watershed that is in an unsustainable state.

The virtual environment is divided into several discrete, hexagon-shaped locations (Figure 2). Each hexagon encompasses an atomic geographic region that can be acted upon by learners as part of solving the environmental sustainability problem scenario. The virtual environment's three-dimensional appearance is stylized, and at first it appears to be brown and lifeless. The environment's appearance is a melodramatic visualization illustrating that the virtual environment is "unhealthy," that is, the environment is in an unsustainable state. Shortly afterward, a short description of the scenario's sustainability-centric objective is presented. The scenario focuses on issues related to farming, particularly waste management and electricity sources. Learners are instructed to explore alternate choices for improving the sustainability of the farms. At this stage, learners are free to begin manipulating the virtual environment to explore the effects of alternate environmental decisions.

Learners can manipulate the virtual environment using multi-touch gestures, such as pinch-to-zoom and rotations. Additionally, learners can tap on various hexagon-shaped locations in the virtual environment, request information about the different factors that bear on the environment's sustainability, and change the behavior of the farms. For example, one of the farms appears to be contaminating a nearby river through fertilizer runoff from a crop field. A learner taps on the farm, and a menu appears offering several options for changing the farm's fertilizer runoff management practices. In order to learn more about the farm's current management practices, the learner taps an "Info" button on the menu. In response, a virtual docent appears on the second vertical display providing a short narration about fertilizer runoff and waste management.

Another example of the use of Future Worlds is that a learner encounters the option to add a riparian buffer to the hog farm. The docent explains that a riparian buffer consists of plants and tree roots that can absorb hazardous nitrogen levels, which leach from synthetic fertilizers used by many farms into nearby water bodies. The learner decides that installing a riparian buffer is worth exploring, and she taps on the menu to enact it. In response, the virtual environment begins to change in appearance: the farm's 3D models transform as the brown-colored fertilizer runoff disappears and a stretch of small bushes, shrubs, and tall grass appears between the farm and adjacent riverbank (Figure 2). Nearby terrain grows greener, and portions of the environment begin to animate. The 3D virtual environment—replete with aesthetic lighting, detailed 3D models, high-resolution textures, and believable animations—provides

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Successive stages of a virtual environment in Future Worlds



a dynamic visualization of the learners' environmental actions through immediate feedback, in effect making learners' sustainability choices "come to life." Using this feedback, learners explore the sustainability simulation and solve the problem scenario (Figure 2).

Paper Prototyping

After identifying design requirements to guide the development of Future Worlds, we undertook an iterative paper prototyping process that involved developing, testing, and refining several paper-based versions of the system (Figure 3). Paper prototypes emulate the visual appearance and interaction design of the final game environment while providing a low-cost method for rapidly creating, deploying, and refining alternate interaction designs. Paper prototypes resemble a "board game" version of the final system—they are typically made using paper and cardboard cutouts—and they source difficult-to-model aspects of the system's functionality to a human "game master." Pilot testing paper prototypes with young learners shares similarities with Wizard of Oz studies (Kelly, 1983). In both cases a human emulates functionality that will eventually be provided by software, such as simulation systems or artificial intelligence functionalities.

During the project's first phase, the team conducted two studies to test early paper prototypes of Future Worlds. The first study was held during a teen-focused event at the North Carolina Museum of Art, and it focused on Future Worlds' interaction design and visual style. The second study was held at the North Carolina Museum of Natural Sciences, and it focused on a refined version of the Future Worlds paper prototype, investigating the complexity of the environmental problem-solving tasks.

During the first study, 11 learners ranging in age from 13-17 interacted with the paper prototype. The prototype focused almost entirely on tradeoffs between energy generation and pollution. Its main purpose was to support investigations of Future Worlds' core gameplay dynamics. In addition to testing basic interaction design, the study provided an opportunity to investigate methods for presenting scientific information, and compare alternate visual styles for the virtual agents. Learners worked individually and collaboratively to solve two problem scenarios with assistance from a project team member. All learners solved the scenarios, and several learners identified alternative solutions to the scenarios. However, the research team observed that learners did not appear to engage deeply with the descriptions of alternate energy sources or consider a broad range of environmental factors in their decision making. The project team reviewed the paper prototype following this first focus group in order to generate feedback about the science content and presentation. The paper prototype was modified in order to enhance the science content, refine the task descriptions, introduce an additional problem scenario, and modify the input mechanics.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Future Worlds paper prototype



The second study was held in coordination with a science summer camp program hosted by the North Carolina Museum of Natural Sciences. The study was designed to examine the revised implementation of the Future Worlds paper prototype. The focus group involved 17 learners in total. The learners were divided into groups so that learners could work on the paper prototypes in pairs while a designated researcher oversaw each paper prototype. Each group interacted with the paper prototypes for approximately 15–20 minutes.

During the focus group's pre- and post-assessments, learners completed a drawing task that involved creating a picture of what is an environment based on their understanding. Participants were asked to label parts of the drawing and write an explanation of what makes their drawing a representation of an environment. Learners produced images of forests, towns, water bodies, animals, and the sky. As part of the post-assessment that was administered after the Future Worlds experience, learners were given the opportunity revise their original drawings. Of those learners, 16 elected to make changes and improve their representation of the environment. All of these learners chose to make additions and not remove aspects of their original drawings. The most common new inclusions were representations of energy sources such as wind (windmills), solar (the sun), petroleum (oil power plants), hydroelectric power (water sources) and human beings, providing preliminary evidence of positive gains in sustainability understanding.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Iterative Development

Although agile-based software development processes are common practice in the game industry, only in recent years have iterative development processes begun to gain attention in the research literature on educational game design (Chamberlin, Trespalacios, & Gallagher, 2012). Our process further contributes to this growth: Upon establishing a benchmark design through several iterations of paper prototyping and pilot tests, we began software production on the Future Worlds game-based learning environment. Future Worlds was built with the Unity game engine. An agile-based development process was employed, which involved identifying, prioritizing, and developing targeted sets of features for 2-4 week development "sprints." During each sprint, software development and art production activities focused only on the features identified for that sprint. At the conclusion of each sprint, an incremental build of the system was demoed internally to the project team, as well as outside collaborators, for constructive feedback. Software production proceeded in this manner for approximately six months.

MUSEUM PILOT TEST

A study was conducted at the North Carolina Museum of Natural Sciences to investigate visitors' learning and engagement with the Future Worlds prototype game-based learning exhibit. Participants were summer campers enrolled in existing summer camp programs at the museum and visited the exhibit in small groups. The study was designed to investigate the prototype exhibit's impacts in a free-choice learning environment that approximated real-world use—this included several competing co-located exhibits—as well as the influence of learners' individual characteristics on sustainability learning and dwell time.

Participants

A total of 43 participants interacted with Future Worlds during the study. The campers were grouped into separate cohorts divided across three sessions ($N = 14, 16, 13$). Of the sample of 43 campers, a majority was male (57.4%). Participants ranged from ages 8-14 with a mean age of 10.7 years. The majority of participants identified as Caucasian (75.9%), followed by Asian/Pacific Islander (7.4%), Black (3.7%), and American Indian/Alaskan Native, Hispanic, and Other (1.9%).

Materials and Apparatus

Learners completed three complementary measures of sustainability concept understanding: a personal meaning map, a sustainability identification task, and a sustainability image-sorting task. These measures were devised to assess visitors' ability to understand and remember sustainability concepts, while remaining consistent with a spirit of voluntary participation and leisure-based learning in museums. To investigate learner engagement, video recordings of participant behavior were recorded and analyzed to determine dwell times.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Personal Meaning Maps

In order to investigate learners' conceptual understanding of sustainability, participants completed personal meaning maps (PMMs). Each PMM consisted of a blank piece of paper with a brief set of instructions and a prompt phrase: sustainability. Participants used a blue ink pen to write or draw words, phrases, and pictures about what they thought and knew about the prompt word on the blank paper. After leaving the exhibit space, participants were given the opportunity to revise their PMM using a red ink pen. After the study, two raters scored each PMM based on the relevance and accuracy of each element included on the page. Typical features that were considered *sustainability-relevant* included drawings of windmills, trees, and solar panels, as well as descriptive words such as "renewable energy" and "ecosystem." Irrelevant features included drawings of pizza and other foods—some learners appeared to confuse "sustenance" with "sustainability"—as well as words that indicated misconceptions or guessing, such as "strong." The raters summed the total number of relevant and correct items listed and subtracted the total number of irrelevant and incorrect items included. The inter-rater reliability for the pre-test ($r = .84$) and post-test ($r = .88$) grades achieved acceptable levels.

Identification Task

Learners were instructed to inspect an illustrated picture of an environment—which depicted both sustainable and unsustainable environmental practices—and annotate the picture by circling "good" environmental practices and crossing out "bad" practices. Learners' pictures were returned during the post-test, and participants revised their annotations. Two independent raters scored learners' annotations using a rubric vetted by subject matter experts. The raters obtained high scoring agreement on both the pre-test ($r = .97$) and post-test ($r = .95$) versions.

Image Sorting Task

Learners were given paper copies of ten images depicting various environmental practices (e.g., recycling, solar panels, heavy traffic congestion, off-shore drilling). Participants were asked to organize the images into two categories of their choosing and were instructed to select the two categories such that they contained as similar a number of images as possible. An expert-based categorization of the images into "sustainable" and "unsustainable" categories was considered the most disciplinarily sophisticated response, and this benchmark was used to grade students' responses. This activity was completed as part of both the pre- and post-test.

Dwell Times

Sessions were video recorded by two cameras strategically placed in the study room to collect data on visitor dwell behavior. Through post-study analyses of the video data, dwell times were determined for each participant in order to assess durations of learner engagement. For each learner, two coders determined total dwell time (time spent interacting with the exhibit to any extent) as well as time spent in each of three possible "tiers of proximity" relative to the Future Worlds exhibit. The first tier is defined as touching the table or standing directly against the interactive tabletop. The second tier is defined as

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

standing immediately behind the first tier participants. The third tier is one layer behind the second tier, within an approximately 5-foot radius of the exhibit. Inter-rater reliability was established with a subset of the study data and then the remainder was coded independently ($k = .70$).

Study Design and Procedure

Pre-test measures were administered on the first day of each cohort's week-long summer camp. The pre-test included the PMM, identification task, sorting task, and demographic questionnaire. Later in the week, learners spent time exploring various parts of the museum, including a room with a thematic focus of "citizen science." This area served as the study room. The Future Worlds exhibit (Figure 4) was kept on-site and made available for use by the summer campers when exploring the citizen science area. During each study session, a human docent stood next to Future Worlds to resolve technical issues and answer questions. Conversation between the docent and participants was kept to a minimum. As part of their summer camp experience, learners were allowed to spend up to 40 minutes in the study room. For each cohort, all participants entered the room at the same time. Learners were told that they could leave the space at any time, and were free to explore the space as they saw fit. Once they left the room, they were not permitted to return.

Including Future Worlds, there were 13 exhibits in the study room. No other museum visitors had access to the citizen science area during the study. Eleven of the exhibits in the study room were permanent exhibits at the museum. In addition to Future Worlds, one temporary exhibit was added to serve as a control. This temporary exhibit was the only other exhibit with a human docent, and it consisted of

Figure 4. Visitors exploring the prototype Future Worlds game-based learning exhibit



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

a white board with a sign asking learners to write down the most interesting thing they learned in the citizen science area. Of the 12 distractor exhibits within the space, half of them ($N = 6$) were interactive and more than half were digital ($N = 7$). It should be noted that none of the content in the distractor exhibits overlapped with the content included in Future Worlds.

A post-test was conducted immediately following each participant's exit of the citizen science area. It included the same sustainability learning assessments as the pre-test.

FINDINGS

To investigate the learning impact of the Future Worlds exhibit, statistical analyses of the pre- and post-test measures, as well as coded video recordings, were conducted. Paired t-tests indicated that learners showed significant gains in PMM score from pre-test ($M = 0.8$, $SD = 1.8$) to post-test ($M = 1.2$, $SD = 2.3$), $t(37) = 2.5$, $p<.05$. There were also significant increases on the identification task from pre-test ($M = 6.0$, $SD = 2.5$) to post-test ($M = 6.4$, $SD = 2.6$), $t(37) = 3.3$, $p<.05$, as well as significant gains on the image sorting task from pre-test ($M = 7.1$, $SD = 3.8$) to post-test ($M = 8.7$, $SD = 2.7$), $t(37) = 2.6$, $p<.05$. Correlation analyses were conducted to investigate whether learners' individual characteristics—including age and gender—showed significant relationships with learning outcomes, but none were observed.

For each learner, total dwell time, as well as time spent in each of three proximity tiers, was determined using video recording data. Across all participants, the average dwell time at Future Worlds was 12.5 minutes. This is promising, given dwell times typically reported in other informal contexts, such as 5.0 minutes (Horn, Solovey, Crouser, & Jacob, 2009) and 4.9 minutes (Lane et al., 2013). However, it should be noted that Future Worlds dwell times were recorded in a semi-controlled study setting, whereas the above cited dwell times were recorded from observations of the general public in non-controlled settings.

We next investigated the effects of individual differences—specifically gender and age—on dwell time. A two-tailed independent samples t-test revealed a significant effect of gender on dwell time, where girls (M = 8m:46s, SD = 5m:21s) spent roughly half the time as boys (M = 16m:23s, SD = 10m:33s) engaging with Future Worlds, $t(35) = 2.9$, $p<.05$. To follow up on this analysis, tier-specific dwell time was examined by gender. Results indicated that males (M = 14m:12s, SD = 7m:34s) spent significantly more time in the first tier than females (M = 8m:56s, SD = 3m:56s), which is the tier of closest proximity to the exhibit, $t(29) = 2.3$, $p<.05$.

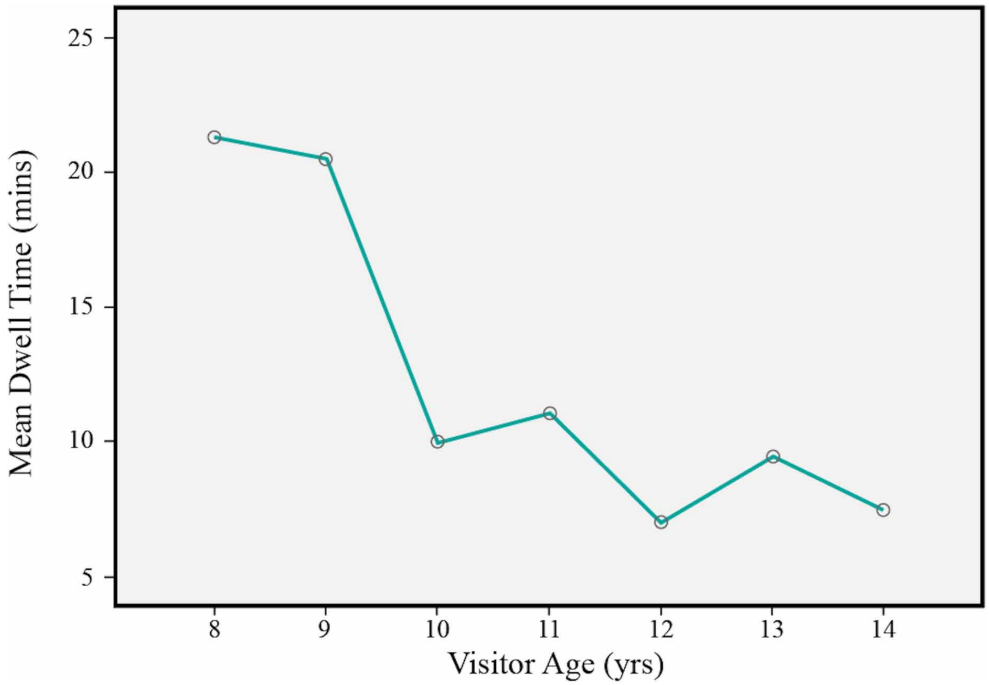
Additional findings about the influence of learners' individual characteristics emerged from analyses of engagement based on learner age. Results indicated that younger children spent more time at the Future Worlds exhibit than older children, $t(44) = 3.5$, $p<.01$. In fact, children under age 10 spent twice as much time (M = 20m:50s, SD = 12m:33s) as children age 10 and older (M = 9m:56s, SD = 6m:45s). Observations of mean dwell time by age are shown in Figure 5.

DISCUSSION

Our results suggest that learners' sustainability understanding improves from interactions with Future Worlds. Furthermore, evidence of extended dwell time, compared to the existing literature, suggests that learners are highly engaged with the exhibit. In combination, these two sets of findings suggest that learner

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Mean dwell time (in minutes) with Future Worlds by visitor age (in years)



engagement with Future Worlds is not superficial; learners are actively engaged for prolonged periods at sufficient depth to yield significant learning gains across three distinct measures of sustainability knowledge.

Our observation that gender and age have significant effects on dwell time point toward engagement-centric design implications for future iterations of Future Worlds. Regarding gender, several possible mediating factors could explain why girls spent less time with Future Worlds than boys, such as video game interest or affinity for the game's visual aesthetic style. Gender is widely recognized for its importance in shaping the design and use of entertainment-focused games (Jenson & de Castell, 2010). Our findings support that gender is likely to be similarly important in the design and use of game-based learning exhibits in museums. Additional studies are necessary to isolate what design factors are responsible for the observed gender differences in visitor experiences with Future Worlds and to determine how Future Worlds' design should be augmented to appeal to females to an extent that is comparable to males.

Regarding age, results suggest that future iterations of Future Worlds should incorporate problem-solving scenarios that span a broader range of content and complexity. It is possible that the implemented problem scenarios in the Future Worlds prototype were sufficiently challenging for young children but were not difficult enough for older children, and thus did not sustain their engagement for extended periods. Older visitors may have solved all of the game's embedded sustainability challenges—in other words, they may have *finished* the game—more quickly than younger learners, leading to shorter dwell times. Incorporating intelligent tutoring system capabilities to dynamically adapt the difficulty of problem scenarios to individual learners, or groups of learners, is a promising way to match scenarios' content complexity to the capabilities of diverse visitors. Adaptive pedagogical planning will require models for automatically detecting learners' individual characteristics as they approach and use exhibits, since administering lengthy pre-tests is not a viable design choice for most informal settings.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Notably, we did not find a relationship between dwell time and learning. We expect that adding curricular material beyond the prototype's current proof-of-concept scope—creating opportunities for more variance in content exposure—could reveal relationships between dwell time and learning that have not thus far been observed.

CONCLUSION AND FUTURE WORK

Game-based learning environments show considerable promise for informal settings such as museums and science centers. Creating game-based learning environments for science museums raises challenges that are distinct from those raised by games for classrooms. Museums and classrooms emphasize different aspects of learning, serve different learner populations, and afford different levels of learner autonomy. To address these challenges, we have identified several design factors for creating game-based learning exhibits in science museums: (1) providing a low barrier to entry, (2) fostering exploration and curiosity, (3) delivering immediate and dramatic feedback, (4) prioritizing visual aesthetics with broad appeal, (5) leveraging novel hardware platforms, and (6) providing flexible user experiences. To investigate these design factors, we pilot tested a game-based learning exhibit called Future Worlds, which was created using these design criteria as a guide. The pilot study found that visitors achieve significant gains in sustainability knowledge, as well as high levels of engagement, while learning with the game. Boys were observed to actively engage with Future Worlds for significantly longer durations than girls, and young children engaged with the exhibit for longer periods than did older children. These individual differences underscore the importance of future work investigating adaptive pedagogical functionalities, as can be provided by intelligent tutoring system techniques, as well as automated detectors for diagnosing learners' individual and group characteristics.

ACKNOWLEDGMENT

The authors would like to thank the staff and visitors of the North Carolina Museum of Natural Sciences as well as the IntelliMedia Group at North Carolina State University. This research was supported by the National Science Foundation under Grant DRL-1114655. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Adams, D., Mayer, R., MacNamara, A., Koenig, A., & Wainess, R. (2012). Narrative games for learning: Testing the discovery and narrative hypotheses. *Journal of Educational Psychology*, *104*(1), 235–249. doi:10.1037/a0025595

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Antle, A., Bevans, A., Tanenbaum, J., Seaborn, K., & Wang, S. (2011). Futura: Design for collaborative learning and game play on a multi-touch digital tabletop. In M. Gross, N. Nunes, E. Do, S. Brewster, & I. Oakley (Eds.), *Proceedings of the Fifth International Conference on Tangible, Embedded, and Embodied Interaction* (pp. 93–100). New York, NY: ACM. 10.1145/1935701.1935721

Chamberlin, B., Trespalacios, J., & Gallagher, R. (2012). The Learning Games Design Model: Immersion, Collaboration, and Outcomes-Driven Development. *International Journal of Game-Based Learning*, *2*(3), 87–110. doi:10.4018/ijgbl.2012070106

Clark, D. B., Tanner-Smith, E., & Killingsworth, S. (2016). Digital games, design, and learning: A systematic review and meta-analysis. *Review of Educational Research*, *86*(1), 79–122. doi:10.3102/0034654315582065 PMID:26937054

Connolly, T., Boyle, E., MacArthur, E., Hainey, T., & Boyle, J. (2012). A systematic literature review of empirical evidence on computer games and serious games. *Computers & Education*, *59*(2), 661–686. doi:10.1016/j.compedu.2012.03.004

Greenspan, B., & Whitson, J. (2013). Whether to play or preserve the past? Creating the forgotten worker quest. In G. Yannakakis, E. Aarseth, K. Jorgenson, & J. Lester (Eds.), *Proceedings of the 8th International Conference on Foundations of Digital Games* (pp. 298–305). New York, NY: ACM.

Habgood, M. P. J., & Ainsworth, S. E. (2011). Motivating children to learn effectively: Exploring the value of intrinsic integration in educational games. *Journal of the Learning Sciences*, *20*(2), 169–206. doi:10.1080/10508406.2010.508029

Horn, M., Leong, Z., Block, F., Diamond, J., Evans, E. M., Phillips, B., & Shen, C. (2012). Of BATs and APEs: An interactive tabletop game for natural history museums. In J. Konstan, E. Chi, & K. Hook (Eds.), *Proceedings of the 2012 ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 2059–2068). New York, NY: ACM. 10.1145/2207676.2208355

Horn, M. S., Solovey, E. T., Crouser, R. J., & Jacob, R. J. K. (2009). Comparing the use of tangible and graphical programming languages for informal science education. In D. Olsen, R. Arthur, K. Hinckley, M. Morris, S. Hudson, & S. Greenberg (Eds.), *Proceedings of the 2009 ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 975–984). New York, NY: ACM. 10.1145/1518701.1518851

Isbister, K., Flanagan, M., & Hash, C. (2010). Designing games for learning: Insights from conversations with designers. In E. Mynatt, G. Fitzpatrick, S. Hudson, K. Edwards, & T. Rodden (Eds.), *Proceedings of the 2010 ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 2041–2044). New York, NY: ACM. 10.1145/1753326.1753637

Jenson, J., & De Castell, S. (2010). Gender, simulation, and gaming: Research review and redirections. *Simulation & Gaming*, *41*(1), 51–71. doi:10.1177/1046878109353473

Johnson, W. L. (2010). Serious Use of a Serious Game for Language Learning. *International Journal of Artificial Intelligence in Education*, *20*(2), 175–195.

Kelley, J. F. (1983). An empirical methodology for writing user-friendly natural language computer applications. In *Proceedings of the 1983 ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 193–196). New York, NY: ACM. 10.1145/800045.801609

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lane, H. C., Cahill, C., Foutz, S., Auerbach, D., Noren, D., Lussenhop, C., & Swartout, W. (2013). The Effects of a Pedagogical Agent for Informal Science Education on Learner Behaviors and Self-efficacy. In H.C. Lane, K. Yacef, J. Mostow, & P. Pavlik (Eds.), *Proceedings of the 16th International Conference on Artificial Intelligence in Education* (pp. 309–318). Heidelberg, Germany: Springer. 10.1007/978-3-642-39112-5_32

Linehan, C., Kirman, B., Lawson, S., & Chan, G. (2011). Practical, appropriate, empirically-validated guidelines for designing educational games. In D. Tan, G. Fitzpatrick, C. Gutwin, B. Begole, & W. Kellogg (Eds.), *Proceedings of the 2011 ACM SIGCHI Conference on Human Factors in Computing Systems* (pp. 1979–1988). New York, NY: ACM. 10.1145/1978942.1979229

Mitgutsch, K., & Alvarado, N. (2012). Purposeful by design? A serious game design assessment framework. In M. El-Nasr, M. Consalvo, & S. Feiner (Eds.), *Proceedings of the 7th International Conference on Foundations of Digital Games*, (pp. 121–128). New York, NY: ACM.

National Research Council. (2009). *Learning science in informal environments: People, places, and pursuits*. Washington, DC: The National Academies Press.

Rowe, J., McQuiggan, S., Robison, J., & Lester, J. (2009). Off-Task Behavior in Narrative-Centered Learning Environments. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. Graesser (Eds.), *Proceedings of the 14th International Conference on Artificial Intelligence in Education* (pp. 99–106). Amsterdam, the Netherlands: IOS Press.

Schulman, D., Sharma, M., & Bickmore, T. (2008). The identification of users by relational agents. In L. Padgham & D. Parkes (Eds.), *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multi Agent Systems*, (pp. 105–111). Richland, SC: IFAAMAS.

Wouters, P., van Nimwegen, C., van Oostendorp, H., & van der Spek, E. (2013). A Meta-Analysis of the Cognitive and Motivational Effects of Serious Games. *Journal of Educational Psychology*, *105*(2), 249–265. doi:10.1037/a0031311

ENDNOTE

[1] In our partner museum, competing exhibits include panoramic theaters, dinosaur skeletons, live animals, and open labs where the public interact with real scientists.

This research was previously published in the International Journal of Gaming and Computer-Mediated Simulations (IJGCMS), 9(3); edited by William Bart; pages 96-113; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 12
Using Physics Inspired Wave Agents in a Virtual Environment:
Longitudinal Distance Control in Robots Platoon

Baudouin Dafflon
Université de Lyon, France

Maxime Guériau
Univ Lyon, ENTPE, IFSTTA, LICIT UMR_T 9401, F-69675, France

Franck Gechter
LE2I, UTBM, University of Bourgogne Franche Comté, France

ABSTRACT

The monitoring and the surveillance of industrial and agricultural sites have become first order tasks mainly for security or the safety reasons. The main issues of this activity is tied to the size of the sites and to their accessibility. Thus, it seems nowadays relevant to tackle with this problem with robots, which can detect potential issues with a low operational cost. To that purpose, in addition to individual patrolling behavior, robots need coordination schemes. The goal of this paper is to explore the possibility of using interference fringes and waves properties in a virtual environment to tackle with the longitudinal distance regulation in the platoon control issue. The proposed model, based on a multi-agent paradigm, is considering each vehicle as an agent wave generator in the virtual environment.

DOI: 10.4018/978-1-7998-0951-7.ch012

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. INTRODUCTION

The monitoring and the surveillance of industrial and agricultural sites have become first order tasks for several reasons such as the security or the safety. For agricultural sites, the monitoring can also be used to adapt the feeding in fertilizer and the prevention of the expansion of diseases. The main issues of this activity is tied to the size of the sites and to their accessibility. Thus, it seems nowadays relevant to tackle with this problem with robots (wheeled or aerial) which can detect potential issues with a low operational cost.

To that purpose, in addition to individual patrolling behavior, robots need coordination schemes and strategies for surrounding a target or for moving collectively as a platoon to one specific location. Among these coordination schemes, platoon control algorithms are mostly used especially when one want to coordinate vehicles. Platoon control has become a wide spread research area for collective robot control but also for autonomous vehicles. A platoon system can be defined as a set of vehicles that move together without any material coupling while maintaining a specific geometric configuration that can be a column, a line, a diagonal or any combination of these. A lot of research projects have been dealing with platoon control for robots moves synchronization or vehicles on highways (Path (\cite{shladover2007path}), SARTRE European project[1]), in urban areas (SafePlatoon[2]) or for freight transportation (FurBot[3]). The column configuration (robots one behind another) is the most spread in literature even if other spatial configurations can also be useful such as in military or agricultural contexts for harvesting for instance.

Basically, there are two main trends for performing platoon control depending on whether the localization reference frame is local (tied to the robot we want to control) or global (common to all the robots of the platoon). In the global approaches, the reference frame can be a shared predefined trajectory, or the trajectory of the leader of the formation sent directly to all its mates. This approach relies on specific methods and sensors aimed at precisely determining each robots position relatively to the global reference frame. Besides, in the case of a transmission of the trajectory of one specific robot in run time, each of them must have efficient and reliable wireless communication devices. The global point is the most spread in literature (Guillet, Lenain, Thuilot, & Martinet, 2014), (Bom, Thuilot, Marmoiton, & Martinet, 2005), (Avanzini, Thuilot, & Martinet, Manual convoying of automated urban vehicles relying on monocular vision, 2012), (Avanzini, Royer, Thuilot, & Derutin, 2013).

On the other side, the local approaches are based on a reference frame local to each robot of the platoon. Thus, each of them is only capable of following one of the robots perceived in its neighborhood, which is thus considered as its local leader. For instance, in a column configuration (i.e. vehicle forming a row), the local leader is generally taken as the preceding robot in the row. Among the papers found in literature based on this approach, one can cite: (El-Zaher, Gechter, Gruer, & Hajjar, 2011), (Scheuer, Simonin, & Charpillet, 2009), (Bouchaala, Marouf, Abualhoul, Pollard, Shagdar, & Nashashibi, 2013), (Abualhoul, Marouf, Shagdar, & Nashashibi, 2013), (Dafflon, Gechter, Gruer, & Koukam, 2013), and (Yazbeck, Scheuer, & Charpillet, 2014).

Each of these solutions is linked to interesting properties. The local approaches are tied to a "individual" standpoint of the platoon function. Generally, they are more simple to put into practice but provide less precise results especially for lateral distance. However, some solutions (El zaher, Gechter, Hajjar, & Gruer, 2016), based on a virtual link with dynamic properties, can compensate this problem. By contrast, due to their "system" point of view, global approaches have better results but requires to design algorithms according to the specific required devices aimed at feeding them in data such as precise differential GPS, communication protocols...

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Among the solutions that rely on the local point of view some of them are based on multi agent paradigm, (Dafflon, Vilca, Gechter, & Adouane, 2015). Multi-Agent Systems (MAS) have become a widespread solution to tackle different kinds of issues in various application fields such as pedestrian simulation (Teknomo, 2006), robot control (Posadas, Poza, Simó, Benet, & Blanes, 2008), complex system description (Bonabeau, 2002), battery management systems... Among the two main trends in multi-agent approaches (cognitive *vs.* reactive), the reactive one is particularly interesting for applications that require both fast response time, adaptability and robustness (Nolfi, 2002). Reactive Multi-Agent Systems (RMAS) rely on simple interaction schemes between entities that can be inspired whether by Biology or Physics. Even if the biological approaches are the most represented in the literature (Leitão, Barbosa, & Trentesaux, 2012), (Biswas, Mishra, Tiwari, & Misra, 2013) the physical ones also allow to obtain good properties especially when they are applied to Cyber-Physical Systems (CPS) control and/or modeling (Sanislav & Miclea, 2012). Most of the models used are inspired by particles which interact through forces (conservative or not) and/or potential fields (Hagelbäck & Johansson, 2008). Few of them are dealing with waves propagation and their associated interaction pattern, known as interference fringes. The goal of this paper is to explore the possibility of using interference fringes and waves properties within a multi-agent standpoint so as to tackle with the issue of platoon longitudinal distance control. This issue has been already explored by different strategies including particle agents approaches or standard PID controller (Hallé, Chaib-draa, & Laumonier, 2003) but has never been studied using waves interaction. This paper is an extension of the paper published in the international conference on tools with artificial intelligence in 2016 (Dafflon et al. (2016)).

The paper is structured as follows: Section 2 presents a review of existing multi-agent approaches. Section 3 exposes the core of our contribution by describing the wave agent framework. Finally, simulations results, in the case study of platoon control, (Section 4) highlight the ability of the approach to cope with such complex problems. Then, in the conclusion, we provide some future development of the proposed framework.

2. RELATIVE WORK

2.1. From Distributed Artificial Intelligence to Multi-Agent Systems

The distributed artificial intelligence can have several forms. Basically, two main trends can be encountered. First, one can see distributed artificial intelligence as a collective behavior that stem from interaction between intelligent entities. These models, which comprise cognitive agents' models, are based on symbolic and explicit knowledge representation. These are then able to exhibit high-level capabilities such as deliberation, consensus or elections (Castelfranchi, Dignum, Jonker, & Treur, 1999). The second trends is relying on the fact that the intelligence is a collective property that emerges from numerous interactions between entities that are not necessary having intelligent individual behaviors. This last kind of models is based on stimulus/reaction behaviors tied to environment representation (Hagelbäck & Johansson, 2008).

In this paper, only the second type of models is considered since the environment is at the center of the problem. In this context, platoon system could be modeled as a classical attraction/repulsion problem between agents each representing a robot as in (El-Zaher, Dafflon, & Gechter, 2015).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

MAS emergent properties are one of the main expected features by multi-agent designers. In (Drogoul, Vanbergue, & Meurisse, 2002), the authors provide a definition of emergence focused on the requirements needed to generate an emerging phenomenon in a dynamic system. In this type of system, emergence refers to the apparition of a phenomenon (the process, steady state or invariant) perceived by global entities, and monitored by an observer or by entities themselves. The aim of Multi-Agent System design is to achieve global objectives through local agent's behavior specifications. Then, its evolution directly depends on the environment modeling choices. Emergent behaviors can be exploited in the same way as convergence properties of classical systems. The study of emergent properties has to cope with some classical issues such as: stability, convergence, and robustness. Moreover, due the generally huge number of agents, the computational cost can be important especially when one need to deal with a high number of high-level agent perceptions. One of the possible solutions to overcome this issue is to use the environment as an active entity, thus not limiting its definition to a place where agents are living, and then giving the environment an active role for the global system behavior (Simonin & Gechter, 2005). For this reason, using waves to propagate information through the environment can be an interesting way to reduce the perception computational cost issues.

Others critical points are linked to the control of the system entities and to the characterization of emergent phenomena (Serugendo Di Marzo, Gleizes, & Karageorgos, 2005). Generally, emergent properties are observed or measured by an external entity. This observation can then be used to design control laws or can be interpreted as a solution to the targeted problem. The main challenge is to characterize the emergent property as a stable global feature of the system, whereas it stems from numerous dynamical interactions. In this context, we argue that waves can overcome this apparent contradiction by providing emergent global properties (interference fringes) arising from waves dynamical propagation. Thus, in opposition to classical MAS approaches based on a bottom-up modeling (*i.e.* from agents' behavior to system properties) we propose to minimize the non-deterministic aspects of multi-agent systems by transferring a part of the interactions in the environment model and dynamics using waves.

2.2. Biological and Physics Inspired Interaction Models

In literature, we can find two main trends for interaction models. The most commonly used is the biological inspiration which tends to reproduce behaviors and interaction that are existing in nature and for which we already know the expected global results. The inspiration sources are numerous for algorithms inspired by social insects (ants, termites, bees, spiders... (Hoolimath, Kiran, & Reddy, 2012), (Zhai, Xu, & Yang, 2015), (Chandran, Reddy, & Janet, 2016) to the use of bigger animals' behaviors inspired by wolfs or birds for instance. The second type of inspiration source is based on Physics. In this case, agents are generally considered as particles that can generate forces on other particles or are moving according to force fields. This kind of interaction can be expressed through forces and can be modeled by gravitation laws (\cite{Sampaio10}), spring and damper forces or Coulomb's laws (Gechter, Beresford, & Rice, Reconstruction of Battery Level Curves Based on User Data Collected from a Smartphone, 2016).

This "particles" point of view is nearly the only one used especially due to the situated properties of agents. This point of view has been used widely in many application fields such as robots foraging or patrolling strategies (Sampaio, Ramalho, & Tedesco, 2010), vehicle driving assistance (Gechter, Contet, Gruer, & Koukam, 2010), platoon control (El-Zaher, Dafflon, & Gechter, 2015) or optimization algorithms such as Particle Swarm Optimization (PSO).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

From a computer execution standpoint, particle agent simulations are made through a scheduler. At each loop, this scheduler computes for each particle agent, the forces that are applied to it. These are collected depending on the perception properties of the considered agent. Once the forces are collected, they are applied to the agent using the second law of motion. Then, through two discrete integrations, the speed and the position of the agent can be determined.

In some ways, the particle inspired agent systems are similar to other particle based methods. The main difference is tied to the fact that each particle agent decides which force has got an influence on it. In classical particle systems, each force is applied to each particle without any estimation of the perception capabilities. This can be seen as a disruption of the interaction symmetry which allows the system to be generally more stable (Fass & Gechter, 2015).

3. WAVE AGENTS IN A VIRTUAL ENVIRONMENT

In this paper is introduced a first model of a multi-agent system where interaction relies on the superposition of waves. This model, called WAVE (Wave Agents in a Virtual Environment) can be considered as a generic approach developed to tackle non-linear decision-making problems. The originality of the approach relies on the use of concepts from wave mechanics to generate emergent stable behaviors. The role of the environment is the cornerstone of the model since it supports agents' interactions that are modelled through waves propagation. WAVE agents are mobile virtual entities, situated in this waves-built environment. They are able to generate and to perceive waves which dynamically modify the environment.

3.1. Objectives and Overview

The main goal of the presented work is to ensure a system high-level stability using low-level interactions based on the reactive multi-agent system paradigm. As described in Section 2, emergent properties of the multi-agent system are somehow difficult to predict or control. However, they offer the ability to model or to generate complex behaviors at the system scale. To enhance the stability of the system, the wave agent model proposes to use an a priori unstable phenomenon (waves propagation) and to exploit emergent stable patterns (interference fringes). To deal with these new kind of interactions, the agents evolve on an adapted environment model. According to these high-level needs, the proposed general framework can be described as the representation given in Figure 1. So as to apply the proposed framework, the first step is to extract the key features of the problem. This step can be done using knowledge on the system's targeted goal. Then the problem can be modeled using WAVE as a physics-inspired MAS environment using waves, distortions and classical Newtons laws. Two types of agents are then defined: The Surfers and the Emitters.

Surfer agents, that can be considered as elementary particles, evolve freely following the waves produced by the Emitter agents, the characteristics of which are tied to the problem. The final decision-making process relies on a dynamic and statistical evaluation of the agent's population.

Figure 2 shows the waves generated by two vehicles (preceding vehicle represented by a vehicle shape and follower vehicle represented by a big dot) and the surfer agents (light color small dots).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Global overview

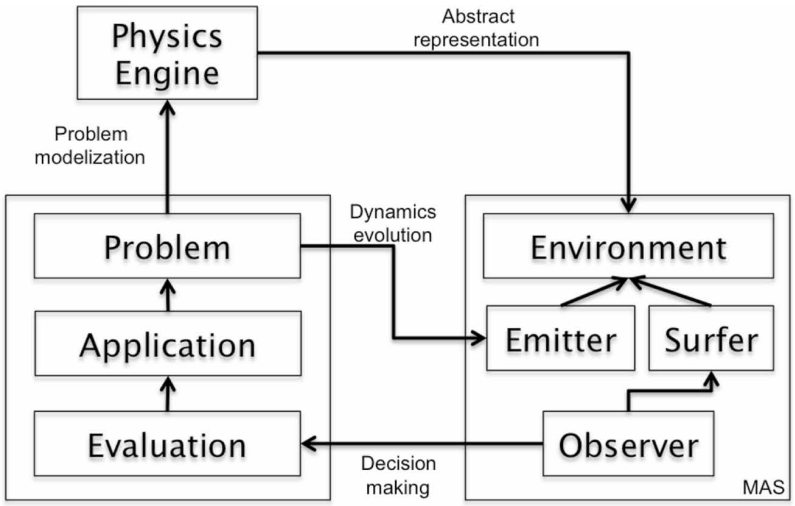
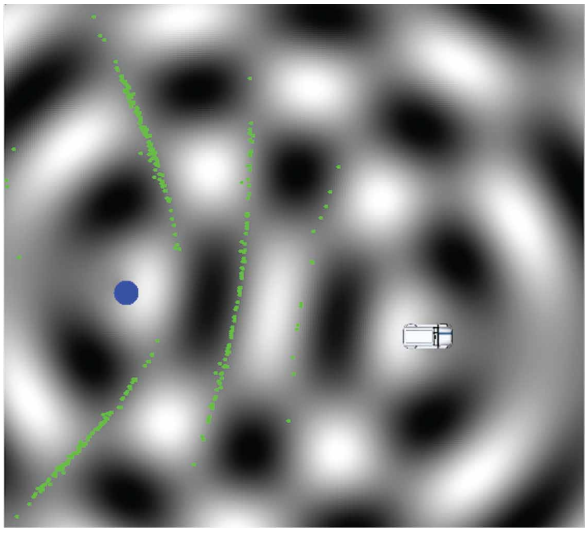


Figure 2. Car waves generated by two vehicles (big dot and vehicle shape) and Surfer agents (light dots)



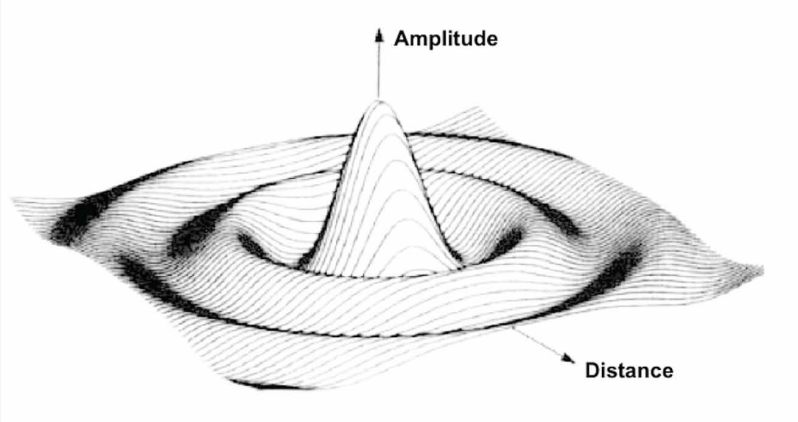
3.2. The Virtual WAVE Environment

The Wave agent framework relies on virtual environment in which situated mobile agents are evolving. The environment is modelled as a continuous 2-dimensional surface described in a 3-D space (Figure 3), the third dimension corresponding to the amplitude of the wave.

The purpose of the WAVE environment is to support, model and propagate waves emitted by agents and related to the problem one want to tackle with. In this work, we consider that the environment is homogeneous and isotropic (i.e. its properties are the same independently from the source position and from the direction of the wave propagation). This implies that waves velocity is fixed and constant, thus allowing the use of a traditional propagation model such as:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. The WAVE environment



$$s(t) = A\cos(\omega t + \varphi)$$

where $s(t)$ denotes the amplitude at time t for given position; the maximum amplitude of the emitting source is denoted as A; ω is the source frequency and φ corresponds to the phase difference (set to zero since the synchronization is ensured by the modeled environment). The instantaneous wave amplitudes is computed by the environment for the local neighborhood of all agents, using the wave propagation model. The model basically computes the sum of the amplitudes at a given position when several wave sources are emitting at the same time. This requires a low computational cost and is able to dynamically combine a high number of influences from several source agents. Hence, the final goal is to seek, identify and exploit perceived global stable patterns.

3.3. Agent Roles

Through the observation of the resulting amplitude of the WAVE environment in their frustum, the agents are identifying and are following the stable patterns generated by their neighbors. We define two agent roles that are combined to generate a collective emergent stable behavior, observable at the system global level. Due to the generic framework of WAVE, the system can be adapted to different problems during the implementation. Thus, the agent roles introduced in this section can be extended, according to the considered problem. The proposed behaviors, called “emitter” and “surfer”, respectively describe the process of a wave source and of a simple entity reacting to perceived waves.

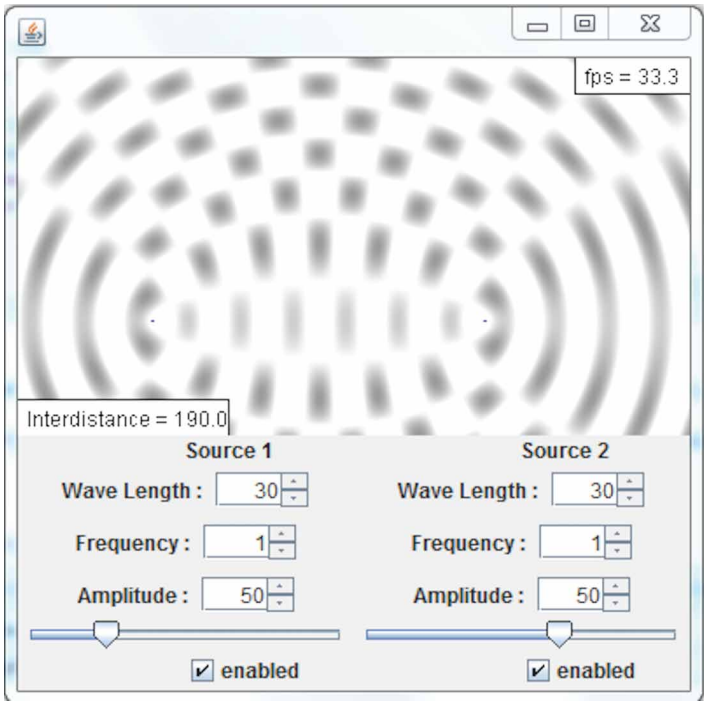
3.3.1. Emitters

Emitter agents are basically wave sources. They generate waves, according to their intrinsic parameters, that are propagated through the WAVE environment. The purpose of the environment is to combine several waves, which can lead to the apparition of an interference fringe pattern, as plotted in Figure 4.

The goal of each emitter agent is to model one of the several goals of the system, depending on the problem addressed. Choosing identical parameters for distinct emitter agents is not mandatory but it contributes to the generation of stable patterns. Otherwise, the fringes generated in the environment

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Screen-shot of the experimental prototype of WAVE featuring the wave propagation model and using two emitter agents (black dots). The selected wave sources parameters generate distinct patterns called interference fringes.



could not remain static (the pattern still occurs but it can move into the environment). We do not consider such a case in this work, although this setting could be exploited to tackle problems where a system has to choose between several goals that are not equivalent. In section 4, we present simulations where two emitter agents are used with the same emission parameters to model a leader and a follower vehicle in a column platoon configuration.

3.3.2. Surfers

A population of moving entities called surfer agents is generated by the WAVE model. The goal of each individual is to reach a stable position in the interference fringes pattern. The key underlying idea is to use a combination of simple reactive agents to observe, as an evaluation, the collective behavior/pattern at the system global scale. The resulting behavior of a surfer agent is a composition of three basic behaviors:

- **Surf:** Surfer agents are attracted by low wave amplitudes (near to 0);
- **Separate:** Surfer agents tend to flee each other, enabling an increase of spatial spreading;
- **Source Attraction:** Agents are trying to stay as close as possible of a given source.

Each behavior is modeled using simple attraction-repulsion equations, based on gravitational laws in $1/(\text{distance}^2)$, involving a mass (set at the same value for all surfer agents). The combination of interactions is performed in the decision process of each surfer agent by applying proportional coefficients to each

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

behavior output. These coefficients are aimed at managing the importance of each behavior relatively to the others. Some behavioral control can also be done by modifying the mass affected to the agents (for both surfers and emitters).

3.4. Decision Making

The decision-making system in WAVE relies on the resulting position of the population of surfer agents. Emitter agents represent the goal of the system while the surfers are using the patterns generated by the emitted waves combination. The decision making relies on an evaluation at the system scale of the positions of the surfer agents relatively to the system goal. Hence, we chose to use the vector which starts from the emitter position and reaches the barycenter (i.e. the weighted average position of agents) of the surfers' positions. This vector can be directly used as an influence vector for the system or can be aggregated (for instance taking into account the distance or the direction only). Thus, the computation of the decision is expected to be robust since it relies on the individual behavior of all the WAVE agents. Next section is dealing with simulation results.

4. SIMULATION RESULTS

In the following simulation experiment, we use the decision-making process of WAVE to provide a longitudinal regulation command to an unmanned vehicle following a leader trajectory. This command must ensure the string stability in a column platoon configuration.

4.1. MAS Platform

Instantiating a multi-agent model requires a platform that aggregates agents, computes interactions and applies influences. To achieve this goal, a tinyMas plug-in was developed using distributed objects and implemented in Java language. tinyMas is a multi-agent platform specifically designed to deal with the implementation and the deployment of continuous environment, reactive agents and holonic multi-agent systems (Fischer, Schillo, & Siekmann, 2003).

It is based on an organizational approach and provides key features such as the support of the concepts of role and organization as first-class entities. This consideration has a significant impact on the agents implementation since it allows them to easily and dynamically change or adapt their behavior and thus contributes to fill the gap between conception and implementation phases in this domain.

4.2. Implementation

As previously said, the first stage of the implementation of the WAVE environment is to extract the main characteristics of the problem. In the case study of platoon control, the goal is to provide a stable longitudinal command to a vehicle which follows a leading vehicle. The leader vehicle is freely controlled by a human driver and the follower is driven by an embedded automated system. The goal of the decision making using WAVE is to compute a command vector for the follower vehicle. Thus, we model the positions of both vehicles in the agent virtual environment as emitter agents and we generate a population

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

of surfer agents executing their behavior, using the follower vehicle as their target for source attraction behavior. The physical behavior of the vehicle then results from a physics-based model which rules the final motion of the vehicle in the real (or simulation) world.

4.2.1. Emitter: Vehicles Model

The emitter agent location is determined by the leader position perceived by the follower sensors. Vehicles used in the simulation represent (graphically and physically) the real experimental vehicles of the laboratory. They are equipped with two 270-degree virtual laser range finders, a replica of LMS SICK 200 sensor and with a simulated GPS-RTK receivers required to follow and study trajectories and to estimate the precision of the platoon function. Simulations are performed on a 3-D geo-localized model of a part of the city of Belfort (France).

4.2.2. Surfer Agent for Platoon

Surfer agents for platoon haven't neither a real representation nor a matching with anything in the real-world problem. They are devoted to decision-making and are based on the generic surfer agent definition given previously. These agents have only one simple behavior, they are attracted by the leader, considered as an emitter, in the virtual environment. Agents motion is determined by gravitational-inspired laws. The environment dynamic topography is updated by waves emission and propagations. Surfing toward the waves hollow, surfer agents are collectively describing a balance point that corresponds to the desired system control.

4.3. Simulation

4.3.1. Scenario

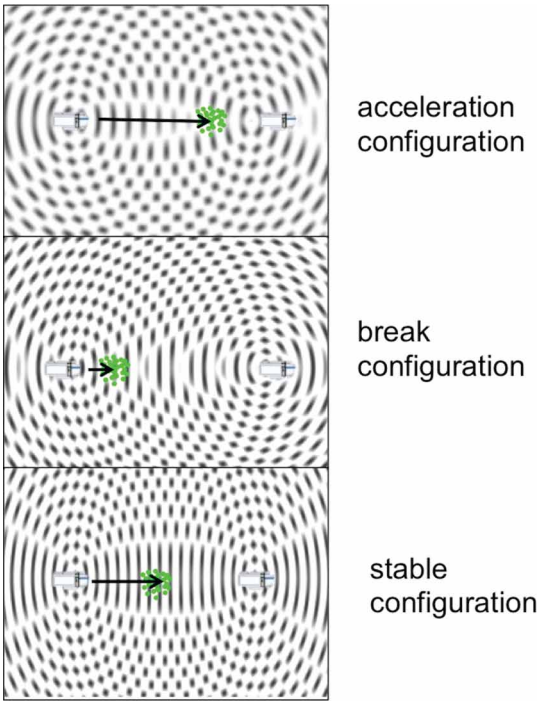
This work is a preliminary step toward a complete platoon application. As a proof of concept, we start the experimentation with the design of a simple scenario. The WAVE environment is used to provide a longitudinal platoon regulation command. To validate the approach, we design a simulation where a leading vehicle and a follower are stopped and aligned on the same lane. The leader starts and quickly reach a speed of 30 km/h, remains at this speed during 2 minutes and then stops. The follower is expected to follow the leader with a very stable inter-distance while the vehicles speeds are the same. Figure 5 depicts a view of the environment where the leader and follower vehicles are modeled as wave sources (i.e. emitters with the same parameters: amplitude = 50, wavelength = 100 and frequency = 3). The surfer's agents are attracted by the source, towards the leader position and generates the longitudinal control.

4.3.2. Result

The behavior resulting from WAVE interaction is the longitudinal behavior of the following vehicle. In order to assess the effect of the model on the stability of the generated behavior, one can observe the inter-distance evolution during the designed scenario (Figure 6 (a)). This figure is presenting both the result obtained with the WAVE framework and results obtained with a classical platoon control local method based on a virtual link made of 2 springs (Figure 6 (b)). This classical method has been studied

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Screen-shot of the environment view. The green dots are the Surfer agents, moving along the interference fringe generated from the addition of the waves emitted by the leader vehicle (white) and the follower (blue). The wave amplitude is represented by a gray-scale (from 0 – white – to 100 – black).



in (El-zaher, Contet, Gechter, and Koukam (2012)) and shows good results especially in the respect of the target longitudinal distance. Using the WAVE framework, the inter-distance, initially set to 2.2 m, is increasing after the start of the leader, then it remains very stable (around 2.7 m) during the cruise period, and finally drops to 2.6 m. While the leader and the follower are driven at full speed, the inter-distance remains quite stable. As opposed to these results, while using the 2-spring method, the inter-distance, initially set to 2.5 m, is better respected but the trajectory is presenting more oscillations around this target position. These oscillations are tied to the 2-spring based model used which is not using a damper as we can sometimes find in other virtual-link based models. We consider the mean error, for the 2-spring model, it is close to zero whereas the WAVE approach is presenting an offset of 0.5 m.

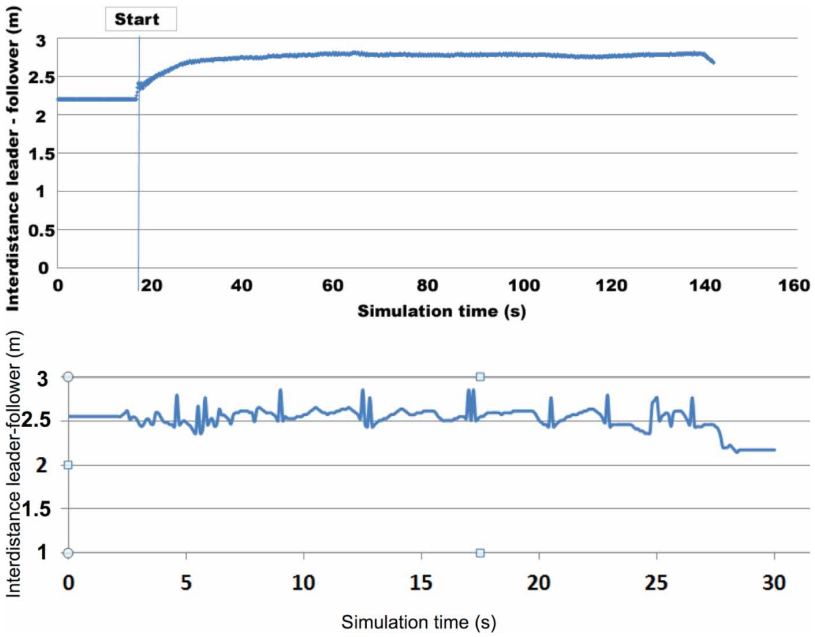
By contrast, the WAVE approach is showing few oscillation making the platoon more comfortable for the passengers. Even if these results are confirming the potential of the use of the pattern generated in the virtual WAVE environment, some improvement are required so as to obtain a good platoon control algorithm comparable to those.

5. CONCLUSION AND FUTURE WORK

Platoon control algorithms are widespread in literature. This platoon function can be considered as a first step towards high automated vehicles that are behaving together as a transportation system. The main issue of the platoon function is tied to the quality of the longitudinal control in terms of respect of the

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Inter-distance between leading and following vehicles in a platoon configuration driven in straight line where the longitudinal control is regulated by a WAVE environment with a target longitudinal distance of 2.2 m (a) compared to the result obtained with a longitudinal regulation made with an interaction model based on 2 springs with a target longitudinal distance of 2.5 m (b)



target distance and in stability. The goal of this paper was to present an alternative to the classical local approaches, which are based whether on regulation loops or on virtual physical links, taking inspiration from wave mechanics. In this model, vehicles are generating waves in a virtual environment that together produce emergent pattern as interference fringes. A first implementation in the WAVE environment for platoon control highlights the potential of the use of physics-based individual agent behavior to generate a complex and stable decision making. As compared to other local approaches based on physics inspired interaction function using for instance virtual springs, the wave approach is providing more stability (i.e. less oscillations) but a less precise respect of the target longitudinal distance. Moreover, as a first step, the experimentation framework has been kept simple and shows some limitation of the model such as the difficulty to provide a matching between agents' environment and real world parameters.

In future work, we plan to integrate more high-level characteristics of the waves patterns as a new source of interaction for the agent. Thus, we can expect to extend the decision taken by the WAVE environment to the lateral distance control this being not dealt yet in the model. We also plan to apply this model to other examples to build up a more generic definition.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Abualhoul, M. Y., Marouf, M., Shagdar, O., & Nashashibi, F. (2013, Oct). Platooning control using visible light communications: A feasibility study. In *Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)* (pp. 1535-1540). 10.1109/ITSC.2013.6728448

Avanzini, P., Royer, E., Thuilot, B., & Derutin, J. P. (2013, May). Using monocular visual SLAM to manually convoy a fleet of automatic urban vehicles. In *Proceedings of the 2013 IEEE International Conference on Robotics and Automation* (pp. 3219-3224). 10.1109/ICRA.2013.6631025

Avanzini, P., Thuilot, B., & Martinet, P. (2012, June). Manual convoying of automated urban vehicles relying on monocular vision. In *Proceedings of the 2012 IEEE Intelligent Vehicles Symposium* (pp. 19-24). 10.1109/IVS.2012.6232128

Biswas, A., Mishra, K. K., Tiwari, S., & Misra, A. K. (2013). Physics-inspired optimization algorithms: a survey. *Journal of Optimization*.

Bom, J., Thuilot, B., Marmoiton, F., & Martinet, P. (2005, Aug). A global control strategy for urban vehicles platooning relying on nonlinear decoupling laws. In *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 2875-2880). 10.1109/IROS.2005.1545270

Bonabeau, E. (2002). Agent-based modeling: Methods and techniques for simulating human systems. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(Suppl. 3), 7280–7287. doi:10.1073/pnas.082080899 PMID:12011407

Bouchaala, Y., Marouf, M., Abualhoul, M., Pollard, E., Shagdar, O., & Nashashibi, F. (2013). Dec#). WAVE Low Latency Video Streaming for Platooning Safety Real-Time Application. In O. Altintas, W. Chen, & A. Yglinski (Eds.), *IEEE Vehicular Networking Conference - VNC 2013* (p. 2). Boston: IEEE.

Castelfranchi, C., Dignum, F., Jonker, C. M., & Treur, J. (1999). Deliberative normative agents: Principles and architecture. In *Proceedings of the International Workshop on Agent Theories, Architectures, and Languages* (pp. 364-378).

Chandran, T. R., Reddy, A. V., & Janet, B. (2016, Feb). A social spider optimization approach for clustering text documents. In *Proceedings of the 2016 2nd International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB)* (pp. 22-26).

Dafflon, B., Gechter, F., Gruer, P., & Koukam, A. (2013, September). Vehicle platoon and obstacle avoidance: A reactive agent approach. *IET Intelligent Transport Systems*, *7*(3), 257–264. doi:10.1049/iet-its.2011.0125

Dafflon, B., Guériau, M., & Gechter, F. (2016, November 6-8). WAVE: wave agents in a virtual environment - A proof of concept applied to longitudinal distance regulation in platoon control. In *Proceedings of the 28th IEEE international conference on tools with artificial intelligence ICTAI 2016*, San Jose, CA (pp. 41–45). doi:10.1109/ICTAI.2016.0017

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dafflon, B., Vilca, J., Gechter, F., & Adouane, L. (2015). Adaptive Autonomous Navigation using Reactive Multi-agent System for Control Law Merging. *Procedia Computer Science*, *51*, 423–432. doi:10.1016/j.procs.2015.05.262

Drogoul, A., Vanbergue, D., & Meurisse, T. (2002). Multi-agent based simulation: Where are the agents? In *Proceedings of the International Workshop on Multi-Agent Systems and Agent-Based Simulation* (pp. 1-15).

El-Zaher, M., Contet, J.-m., Gechter, F., & Koukam, A. (2012, September). Echelon platoon organisation: a distributed approach based on 2-spring virtual links. In *Proceedings of the 15th international conference on artificial intelligence: Methodology, systems, applications (AIMSA),* LNCS (Vol. 7557). Springer-Verlag Berlin Heidelberg. doi:10.1007/978-3-642- 33185-5_28

El-Zaher, M., Dafflon, B., & Gechter, F. (2015). A cyber-physical model for platoon system. In *Proceedings of the 2015 9th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)* (pp. 1-8).

El-Zaher, M., Gechter, F., Gruer, P., & Hajjar, M. (2011, Nov). A New Linear Platoon Model Based on Reactive Multi-agent Systems. In *Proceedings of the 2011 IEEE 23rd International Conference on Tools with Artificial Intelligence* (pp. 898-899).

El-Zaher, M., Gechter, F., Hajjar, M., & Gruer, P. (2016, #oct#). An interaction model for a local approach to vehicle platoons. *International Journal of Vehicle Autonomous Systems*.

Fass, D., & Gechter, F. (2015). Towards a Theory for Bio Cyber Physical Systems Modelling. In V. G. Duffy (Ed.), *Digital Human Modeling. Applications in Health, Safety, Ergonomics and Risk Management: Human Modeling* (pp. 245-255). Springer International Publishing. 10.1007/978-3-319-21073-5_25

Fischer, K., Schillo, M., & Siekmann, J. (2003). Holonic multiagent systems: A foundation for the organisation of multiagent systems. In *Proceedings of the International Conference on Industrial Applications of Holonic and Multi-Agent Systems* (pp. 71-80). 10.1007/978-3-540-45185-3_7

Gechter, F., Beresford, A. R., & Rice, A. (2016). Reconstruction of Battery Level Curves Based on User Data Collected from a Smartphone. In *Proceedings of the International Conference on Artificial Intelligence: Methodology, Systems, and Applications* (pp. 289-298). 10.1007/978-3-319-44748-3_28

Gechter, F., Contet, J.-M., Gruer, P., & Koukam, A. (2010). Car-driving assistance using organization measurement of reactive multi-agent system. *Procedia Computer Science*, *1*(1), 317–325. doi:10.1016/j.procs.2010.04.035

Guillet, A., Lenain, R., Thuilot, B., & Martinet, P. (2014, March). Adaptable Robot Formation Control: Adaptive and Predictive Formation Control of Autonomous Vehicles. *IEEE Robotics & Automation Magazine*, *21*(1), 28–39. doi:10.1109/MRA.2013.2295946

Hagelbäck, J., & Johansson, S. J. (2008). Using multi-agent potential fields in real-time strategy games. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems* (Vol. 2, pp. 631-638).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hallé, S., Chaib-draa, B., & Laumonier, J. (2003). Car platoons simulated as a multiagent system. In *Proc. 4th Workshop on Agent-Based Simulation*, (pp. 57-63).

Hoolimath, P. K., Kiran, M., & Reddy, G. R. (2012, July). Optimized Termite: A bio-inspired routing algorithm for MANETś. In *Proceedings of the 2012 International Conference on Signal Processing and Communications (SPCOM)* (pp. 1-5). 10.1109/SPCOM.2012.6290230

Leitão, P., Barbosa, J., & Trentesaux, D. (2012). Bio-inspired multi-agent systems for reconfigurable manufacturing systems. *Engineering Applications of Artificial Intelligence*, *25*(5), 934–944. doi:10.1016/j.engappai.2011.09.025

Nolfi, S. (2002). Power and the limits of reactive agents. *Neurocomputing*, *42*(1-4), 119–145. doi:10.1016/S0925-2312(01)00598-7

Posadas, J. L., Poza, J. L., Simó, J. E., Benet, G., & Blanes, F. (2008). Agent-based distributed architecture for mobile robot control. *Engineering Applications of Artificial Intelligence*, *21*(6), 805–823. doi:10.1016/j.engappai.2007.07.008

Sampaio, P. A., Ramalho, G., & Tedesco, P. (2010, Oct). The Gravitational Strategy for the Timed Patrolling. In *Proceedings of the 2010 22nd IEEE International Conference on Tools with Artificial Intelligence* (Vol. 1, pp. 113-120).

Sanislav, T., & Miclea, L. (2012). Cyber-physical systems-concept, challenges and research areas. *Journal of Control Engineering and Applied Informatics*, *14*, 28–33.

Scheuer, A., Simonin, O., & Charpillet, F. (2009, May). Safe longitudinal platoons of vehicles without communication. In *Proceedings of the 2009 IEEE International Conference on Robotics and Automation* (pp. 70-75). 10.1109/ROBOT.2009.5152629

Serugendo Di Marzo, G., Gleizes, M.-P., & Karageorgos, A. (2005). Self-organization in multi-agent systems. *The Knowledge Engineering Review*, *20*(02), 165–189. doi:10.1017/S0269888905000494

Shladover, S. E. (2007). PATH at 20—History and major milestones. *IEEE Transactions on Intelligent Transportation Systems*, *8*(4), 584–592. doi:10.1109/TITS.2007.903052

Simonin, O., & Gechter, F. (2005). An environment-based methodology to design reactive multi-agent systems for problem solving. In *Proceedings of the International Workshop on Environments for Multi-Agent Systems* (pp. 32-49).

Teknomo, K. (2006). Application of microscopic pedestrian simulation model. *Transportation Research Part F: Traffic Psychology and Behaviour*, *9*(1), 15–27. doi:10.1016/j.trf.2005.08.006

Yazbeck, J., Scheuer, A., & Charpillet, F. (2014, May). Decentralized near-to-near approach for vehicle platooning based on memorization and heuristic search. In *Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 631-638). 10.1109/ICRA.2014.6906921

Zhai, Y., Xu, L., & Yang, Y. (2015, August). Ant Colony Algorithm Research Based on Pheromone Update Strategy. In *Proceedings of the 2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics* (Vol. 1, pp. 38-41).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ENDNOTES

[1] http://www.sartre-project.eu/en/Sidor/default.aspx
[2] http://www.multiagent.fr/MultiAgentWiki:SafePlatoon
[3] http://www.furbot.eu

This research was previously published in the International Journal of Monitoring and Surveillance Technologies Research (IJMSTR), 5(2); edited by Konstantina S. Nikita and Nikolaos Bourbakis; pages 15-28; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 13
A Geometric Dynamic Temporal Reasoning Method with Tags for Cognitive Systems

Rui Xu
Beijing Institute of Technology, China

Pingyuan Cui
Beijing Institute of Technology, China

Zhaoyu Li
Beijing Institute of Technology, China

Shengying Zhu
Beijing Institute of Technology, China

Ai Gao
Beijing Institute of Technology, China

ABSTRACT

Temporal reasoning is one of the cognitive capabilities humans involve in communicating with others and everything appears related because of temporal reference. Therefore, in this paper a geometric dynamic temporal reasoning algorithm is proposed to solve the temporal reasoning problem, especially in autonomous planning and scheduling. This method is based on the representation of actions in a two dimensional coordination system. The main advantage of this method over others is that it uses tags to mark new constraints added into the constraint network, which leads the algorithm to deal with pending constraints rather than all constraints. This characteristic makes the algorithm suitable for temporal reasoning, where variables and constraints are always added dynamically. This algorithm can be used not only in intelligent planning, but also computational intelligence, real-time systems, and etc. The results show the efficiency of our algorithm from four cases simulating the planning and scheduling process.

1. INTRODUCTION

The notion of time is obvious in any activity that requires intelligence (Pani & Bhattacharjee, 2001; Wang, 2014). The temporal reasoning is one of the cognitive capabilities humans involve in communicating with others (Stock, 1998). Everything appears related because of temporal reference (Pani & Bhattacharjee,

DOI: 10.4018/978-1-7998-0951-7.ch013

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2001). Therefore, cognitive models or intelligent systems have to be developed to mode this capability of representing time and reasoning about temporal relationship (Stracuzzi, Li, Cleveland, & Langley, 2009; Wang & Patel, 2009). Certainly, it is also important many possible application of computer-based systems, for instance planning actions in the real world which is one of high cognitive abilities, scheduling within a complex system (Cervantes, Rodríguez, López, Ramos, & Robles, 2013; Stock, 1998).

This paper takes the application of temporal representation and reasoning in autonomous planning and scheduling as an example to specify our work, because concepts about time are very important when people are talking about actions or events. In real world, actions are not instantaneous but have start times, end times and durations. An important challenge in current planning and scheduling research is time dimension management to allow for the simultaneous execution of non-instantaneous actions (M. C. Cooper, Maris, & Régnier, 2013). The sequential nature of classical plans is often too restrictive to cope with a temporal plan that consists of a set of instances of durative actions (M. Cooper, Maris, & Régnier, 2014). Therefore, time and temporal constraints must be considered in planning and scheduling methods.

Plan-space planning is well suited to temporal planning problems (Cushing, 2012). In plan-space planning, temporal reasoning is needed to check whether a simple temporal problem (STP) instance is consistent once a flaw and conflict is solved (Barták, Morris, & Venable, 2014). It means the temporal reasoning algorithm is called for many times. So improving the efficiency of temporal reasoning in dynamic environment can make the performance of the planning and scheduling method better. Therefore, the authors are interested in dynamic reasoning, namely whenever a new constraint or variable is asserted, this paper wants to maintain a solution of the STP rather than recomputing it using a "static" algorithm (Gerevini, 2005). The static algorithm, such as the Floyd-Warshall algorithm (Aini & Salehipour, 2012), leads to computation cost, because when the value of local variables changes or a variable/constraint is added to an STN, the whole network needs to be propagated again (Xu, 2004).

The most commonly used method is based on a simple temporal network (STN) and there is a list of algorithms. Xu et al. (2004) has proposed a dynamic algorithm which only computes the partial network influenced by the new introduced constraint rather than the whole network with the worst time complexity of the algorithm. Mouhoub and Yip (2002) presented a comparative study of dynamic arc-consistent algorithms in the case of temporal constraint problems and shown the efficiency of AC-3.1|DC algorithm. Planken et al. (2010) presented a brief survey of existing incremental algorithms and proposed a new algorithm that incrementally enforces partial path consistency. Gerevini (2005) proposed a collection of new polynomial algorithms to incrementally decide satisfiability, to maintain a solution, or to update the minimal representation of the CSP. This approach based on STN also has been used in several planning and scheduling systems in order to solve real-world problems, such as the Heuristic Scheduling Testbed System of Deep Space 1 (Muscettola, 1994), the EUROPA system (Barreiro et al., 2012; Bresina, 2015; Iwig et al., 2015), IxTeT (Lemai, 2004) and VHPOP (Younes & Simmons, 2003).

The authors aim to propose an efficient dynamic temporal reasoning algorithm to determine the consistency of the STP and to eliminate possible values to variables in planning and scheduling problems. In order to verify the efficiency of our algorithm, the authors compare it with the temporal reasoning algorithm in EUROPA. It provides incremental algorithms e. g. adaptive Bellman-Ford algorithm to detect a negative cycle which is sufficient for backtrack-free search (Chandrachoodan, Bhattacharyya, & Liu, 2001). On the one hand, the authors want to solve planning and scheduling problems in the future, so it is more suitable and convenient to select a reasoning algorithm having used in planning and scheduling system. On the other hand, the algorithm in EUROPA is actually efficient in temporal reasoning.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Durative actions are 2-dimensional (2-D) in nature, making their 1-dimensional (1-D) representation ambiguous (Rit, 1986). Therefore, in order to solve problems above and take advantage of the 2-D nature, this paper introduces a geometric method to represent durative actions based on our prior work (Li, Xu, Cui, Xu, & He, 2016) and proposes a dynamic algorithm using tags on constraints to propagate temporal constraints.

This paper is organized as follows. First, the authors introduce related work to help readers understand our method, including defining the planning and scheduling problems and the basic representation of actions in the 2-D space. Then, details of the geometry-based dynamic temporal constraint reasoning algorithm are described. In the fourth section, the authors present experiments to verify the performance of our method compared with that of the algorithm in the EUROPA system.

2. RELATED WORK

2.1. Problem Statement

Planning and scheduling techniques have been improved thanks to the application of constraint satisfaction models and tools (Barták, Salido, & Rossi, 2010). In order to apply the efficient constraint satisfaction techniques to planning and scheduling, our problems can be described as constraint satisfaction problems (CSPs) (Tsang, 2014).

Definition 1: A planning and scheduling problem $P=<P,O,C,s_0,g>$ consists of:

- A set of variables $V=\{v_1,v_2,\ldots,v_m\}$, and $v\in D_v$, where D_v is the variable range;
- A set of actions $O=\{o_1,o_2,\ldots,o_n\}$ where an action o has a start time s, an end time e, and a duration d; in other words, $o=\{d,s,e\}$;
- A set of constraints C, namely temporal constraints between two actions, such as $C(o_x,o_y)$ = "o_x *Cons* o_y";
- An initial state s_0;
- A goal state g.

A solution to a planning and scheduling problem P is the result of scheduling all actions to satisfy all constraints, in which each action is assigned an execution time. Therefore, the representation of actions and temporal constraints between them and then the propagation of the temporal constraints are essential tasks for solving a P.

2.2. Geometric Representation

The representation of actions and temporal constraints is the basis of temporal reasoning. Because actions are 2-dimensional (2-D) in nature, making their 1-dimensional (1-D) representation ambiguous (Rit, 1986), this paper takes advantage of geometric representation of durative actions in a 2-D coordinate system to represent actions and temporal constraints.

Rit (Rit, 1986) proposed a graphical representation of constrained set of possible occurrences first. He also described a constraint propagation algorithm which is a Waltz-type filtering algorithm. Pujari, Kumari and Sattar later discussed this type of representation (Rit, 1986), while other works have only

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

briefly mentioned it (Aigner & Miksch, 2006; Duftschmid, Miksch, & Gall, 2002). Recently, Ullberg used this method to represent and propagate several simple quantitative constraints to solve the context recognition problem (Ullberg & Pecora, 2012).

2.2.1. Representation of Actions

The most straightforward way of representing durative actions would be to interpret an action as a single point in a 2-D graph as in Figure 1. The x-axis defines the action's start-time and the y-axis defines the end-time. An action is not permitted to end before it has started, so no point is in the shadow region.

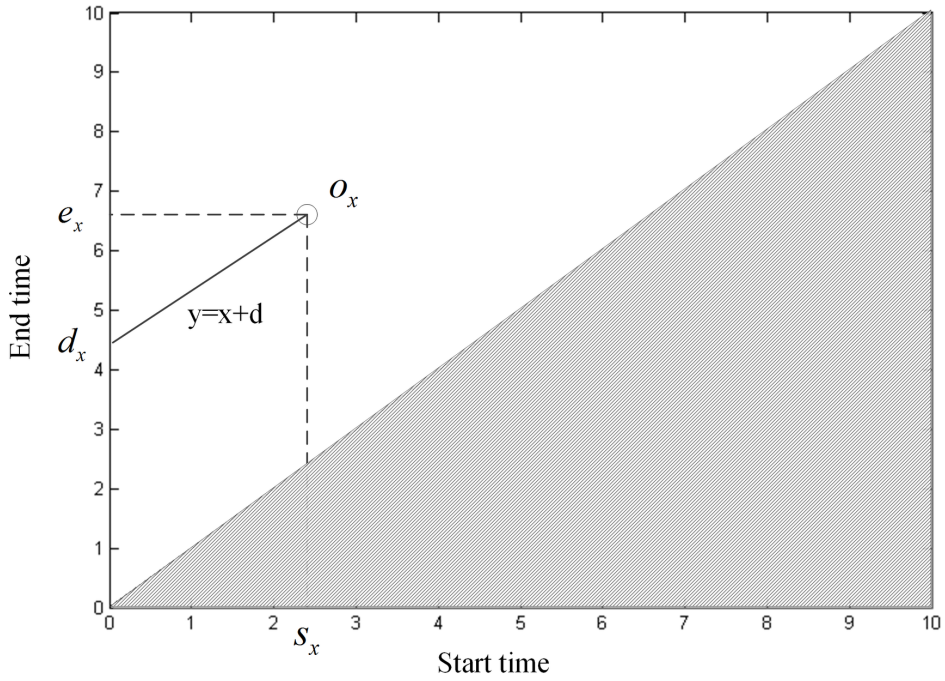
In the planning and scheduling process, actions usually have flexible range for their start time, end time and duration to deal with uncertainty in execution. Figure 2 shows the representation of an action $o_x = \{s_x, e_x, d_x\}$ with flexible domain $s_x \in [s1_x, s2_x]$, $e_x \in [e1_x, e2_x]$ and $d_x \in [d1_x, d2_x]$. If there is a relative zero point, then $s2_x \geq s1_x \geq 0$. Similarly, $e2_x \geq e1_x \geq 0$ and $d2_x \geq d1_x \geq 0$. The bounds of domain are represented using lines in the 2-D system.

Compared with the simple temporal network (STN) representation method, the geometric method takes an action as a whole while the STN method has to use two points to represent one action. This may reduce time to propagate constraints between two actions.

2.2.2. Representation of Constraints

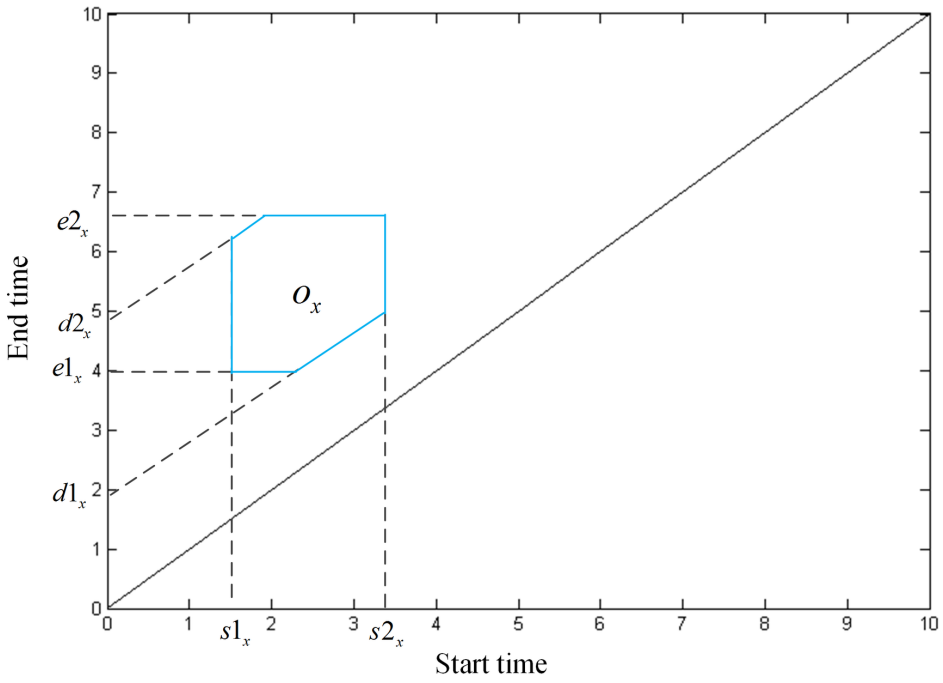
Constraints between actions consist of qualitative and quantitative constraints. This paper focuses on quantitative constraints. Assume there are two actions o_x and o_y which belong to the set of actions $O = \{o_1, o_2, \ldots, o_n\}$, and $o_x = \{d_x, s_x, e_x\}$, $o_y = \{d_y, s_y, e_y\}$. The constraint between o_x and o_y is represented by $C(o_x o_y)$

Figure 1. Representation of an action in the 2-D system



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Representation of an action o_x with flexible domain



= "o_x *Cons* o_y". *Cons* has four elements including [*ss*1,*ss*2], [*se*1,*se*2], [*es*1,*es*2] and [*ee*1,*ee*2]. [*ss*1,*ss*2] means that the action o_x starts from *ss*1 to *ss*2 earlier than o_y starts. [*se*1,*se*2] means that the action o_x starts from *se*1 to *se*2 earlier than o_y ends. [*es*1,*es*2] means that the action o_x ends from *es*1 to *es*2 earlier than o_y starts. [*ee*1,*ee*2] means that the action o_x ends from *ee*1 to *ee*2 earlier than o_y ends. According to the domain of action o_x and constraints, the domain of o_y can be got. As an example, Figure 3 shows the domain of o_y according to constraints [*se*1,*se*2] and [*es*1,*es*2], respectively.

3. GEOMETRIC DYNAMIC TEMPORAL REASONING ALGORITHM

During planning and scheduling, the deletion and addition of actions or constraints will lead to the dynamic changes of the constraint network. Recomputation every time may cause massive cost computation. Therefore, in this section, a geometry-based dynamic reasoning algorithm (GD) is proposed to check the consistency of the network and give the minimal range of the start time and end time of all actions as well, shown in Figure 4.

The inputs of the GD algorithm include the set of actions O, the set of action pairs R, the set of constraints C and the set of tags *Tags*. O and C are defined in the section 2. There is a one-to-one correspondence between R and C, i. e. R contains all action pairs ($o_x o_y$) that have constraints contained in C such as $C(o_x o_y)$ = "o_x *Cons* o_y". Tags are the key role in dynamic propagation. Every action pair has a tag contained in *Tags*. The domain of a tag is (0, 1). If *Tags*(i)=1, it indicates that the action pair $R(i)$ needs processing. If *Tags*(i)=0, it is unnecessary to process $R(i)$ because neither of actions in $R(i)$ changes. At the beginning, all values of *Tags* are initialized to 1.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Representation of constraints [se1,se2] and [es1,es2]

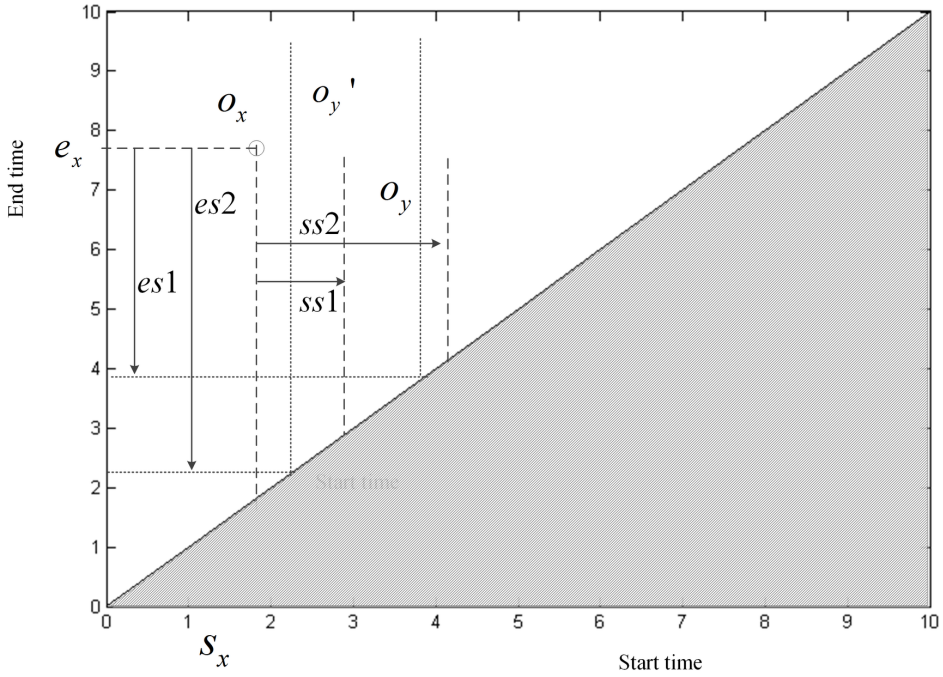


Figure 4. The geometric dynamic propagation algorithm

Algorithm GD

```
1:  function dynamic_propagation (O, C, R, Tags)
2:      for  ∀o_i ∈ O
3:          D(o_i) = EvaluateAction ( o_i );
4:  Tags ← 1
5:  while  ∃  Tags(i) = 1 do
6:      for any Tags(i) = 1 in Tags
7:          Tags(i) = 0;
8:          if Propagate( o_x , o_y , C(i) )
9:              If  ∃R[ j](o_y , o_z )
10:                 Tags(j) = 1;
11:     end for
12:   end while
```


EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.1. Evaluating Domain of Actions

At first, the algorithm checks weather all actions exist using function *Evaluate Action*(o_i), where $o_i \in O$. For the simplify of specification, this paper uses graphical method to describe steps in *Evaluate Action*(o_i). Here, assume there are two actions o_x and o_y, and $o_x=\{d_x,s_x,e_x\}$, $o_y=\{d_y,s_y,e_y\}$. In Figure 2, the final region in blue is equal to the initial domain of action o_x. If:

$$[y1_x - s2_x, y2_x - s1_x] \cap [d1_x, d2_x] = \varnothing$$

The action o_x does not exist and the constraint network is not consistent.

Shown in Figure 5(a), if the line $y=x+d2_x$ intersects with the line $y=e1_x$ between $x=s1_x$ and $x=s2_x$, the domain of s_x is changed into $s_x \in [s1_x^*, s2_x]$; if the line $y=x+d2_x$ intersects with the line $x=s2_x$ between $y=e1_x$ and $y=e2_x$, the domain of e_x is changed into $e_x \in [e1_x, e2_x^*]$. Similarly, the line $y=x+d1_x$ can also change the domain of s_x and e_x as shown in Figure 5(b).

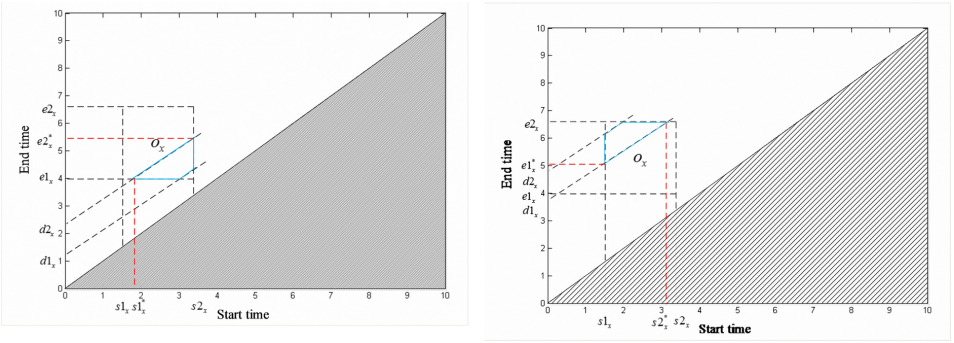
In a conclusion, points of intersection of the six lines $x=s1_x$, $x=s2_x$, $y=e1_x$, $y=e2_x$, $y=x+d2_x$ and $y=x+d1_x$ determine the final domain of action o_x.

3.2. Propagation of Constraints

This section is based on the authors' prior work (Duftschmid et al., 2002) and it gives details of propagation of constraints between actions using the function *Propagate*(o_x,o_y,$C(i)$). Here, this paper only uses the conclusion and gives the brief introduction.

Constraints $C(o_x,o_y)$ between two actions $o_x=\{s_x,e_x,d_x\}$ and $o_y=\{s_y,e_y,d_y\}$ are called a constraint pair. A constraint pair includes four elements, i. e. the constraint [$ss1$,$ss2$] placed by the start time of o_x on the start time of o_y, the constraint [$se1$,$se2$] placed by the start time of o_x on the end time of o_y, the constraint [$es1$,$es2$] placed by the end time of o_x on the start time of o_y and the constraint {$ee1$,$ee2$] placed by the end time of o_x on the end time of o_y. Here, the paper gives the method of using the domain of o_x and constraint pair $C(o_x,o_y)$ to calculate all possible values of action o_y directly. This paper will refer to this set as $D^*(o_y)$ and use it to geometrically intersect the previous domain of o_y so as to remove inconsistent values.

Figure 5. Diagram of the influence of domain of d_x on domain of s_x and e_x



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The possible domain of the start time of o_y is shown in (1):

$$D^*(s_y) = [s1_x + ss1,\ s2_x + ss2] \cap [e1_x + es1, e2_x + es2]$$

The possible domain of the end time of o_y is shown in (2):

$$D^*(e_y) = [s1_x + se1, s2_x + se2] \cap [e1_x + ee1, e2_x + ee2]$$

Calculation of the possible domain of the duration of o_y is a little complex as follows:

$$D^*(d_y) = d1 \cap d2 \cap d3 \cap d4$$

$d1$, $d2$, $d3$ and $d4$ can be obtained by:

$$\begin{cases} d1 = [se1 - ss2,\ se2 - ss1] \\ d2 = [ee1 - es2, ee2 - es1] \\ d3 = [d31, d32] \\ d4 = [d41,\ d42] \end{cases}$$

where:

$$d31 = \min(d1_x - ss1 + ee1, d1_x - ss1 + ee2,\ d1_x - ss2 + ee1, d1_x - ss2 + ee2)$$

$$d32 = \max(d2_x - ss1 + ee1, d2_x - ss1 + ee2,\ d2_x - ss2 + ee1, d2_x - ss2 + ee2)$$

$$d41 = \min(-d1_x - es1 + se1, -d1_x - es1 + se2,\ -d1_x - es2 + se1, -d1_x - es2 + se2)$$

$$d42 = \max(-d2_x - es1 + se1, -d2_x - es1 + se2,\ -d2_x - es2 + se1, -d2_x - es2 + se2)$$

The final step of propagating one constraint pair is removing redundant values of o_y using (5):

$$D^{new}(o_y) = \{D^*(s_y) \cap [s1_y, s2_y],\ D^*(e_y) \cap [e1_y, e2_y],\ D^*(d_y) \cap [d1_y, d2_y]\}$$

3.3. Analysis of Algorithm

In plan-space planning, the initial plan is incomplete and the aim of planning and scheduling is to fill the plan with actions or states. When an action is added in the plan, some temporal constraints may also be added into the constraint network besides points. The propagation algorithm will be called every time

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

when actions are added into the plan in order to check the consistency of the constraint network. Only when the network is consistent can the next step of planning continues. Therefore, the efficiency of the propagation algorithm has a significant impact on that of planning and scheduling.

The GD algorithm is dynamic because of the introduction of tags on action pairs. Unlike the traditional method that needs to process all constraints every time, the GD algorithm only needs to process the marked constraints. In other words, if new actions and their constraints are added to the constraint network, action pairs related to new constraints are marked with $Tags(i)=1$, and then the GD algorithm process these action pairs at the beginning. This brings the advantage of reducing the number of pending constraints.

Figure 6 is the flow diagram of the GD algorithm. Actions and constraints are inputs of the algorithm. First, the function *Evaluate Action* is used to check weather all actions exist. Then, all constraints are tagged with the value 1 and one constraint is selected to be propagated using $Propoagate(o_x,o_y,C(i))$. At the same time, the value 0 is assigned to the tag related to $C(i)$. If the domain of o_y is changed because of the constraint, tags of all constraints related to o_y will get the value 1. Perform the function *Propagate* until all tags have the value 0. At last, the consistency will be judged and minimal domain of start time and end time of all actions will be got.

Figure 7 is the diagram to show the dynamic characteristic of the algorithm. New constraint pairs are added in the list above, which are marked red. The values of their tags are 1. After new action pairs are processed, the domain of action o_2 may be changed. If it happens, actions pairs related to o_2 have to be checked with the tag from 0 to 1 marked red in the list below.

4. EXPERIMENTAL EVALUATION

In this section of the paper we compare the performance of the GD algorithm against the results that can be obtained using the temporal reasoning algorithm in the EUROPA system, which provides incremental algorithms e. g. adaptive Bellman-Ford algorithm to detect a negative cycle which is sufficient for backtrack-free search (Barták et al., 2010). Here, the algorithm in EUROPA is called EU algorithm for simplify. The data used for this evaluation comes from the reference (Planken, De Weerdt, & Van Der Krogt, 2012). For reflecting the planning and scheduling process, four cases are considered, which demonstrates the efficiency of our algorithm from multiple perspectives.

Both algorithms were implemented in C++. Experiments were conducted with Microsoft Visual Studio 2008 on Intel Core i7-2600 CPUs running 32-bit Windows. The run times were averaged over 10 runs for each point in the simulation diagrams. Details about data are stated in each case.

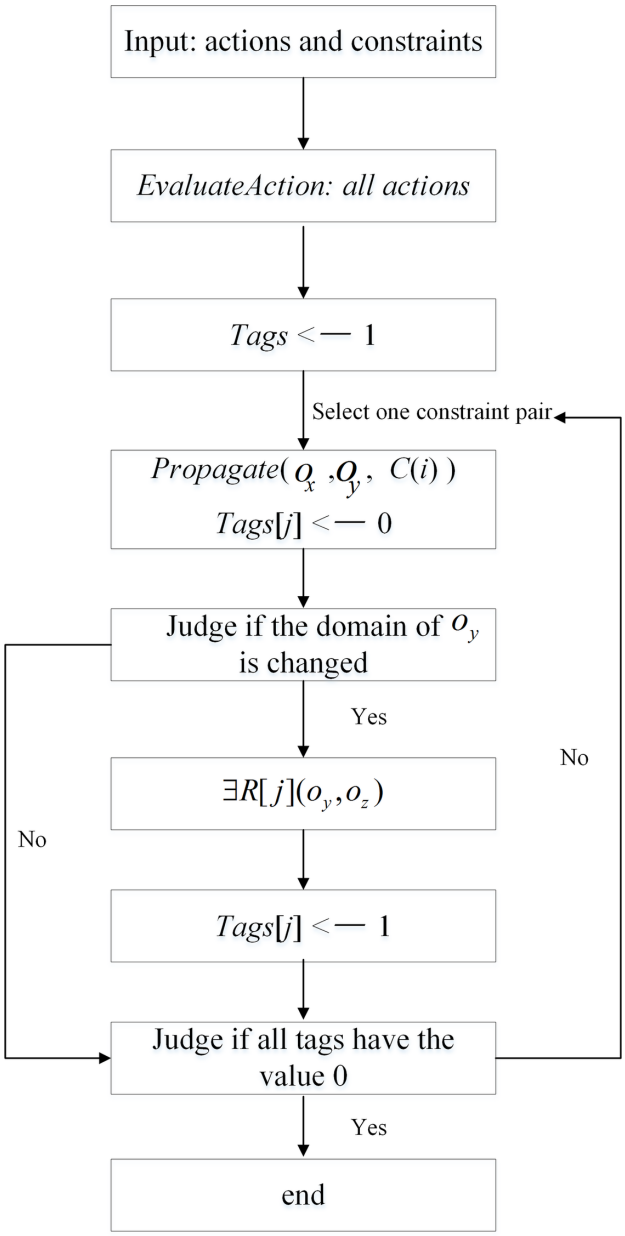
4.1. Case 1

In this case, we want to verify how the number of constraint pairs added can influence the performance of algorithms. The initial constraint network has 250 variable points and 3800 constraint edges. When new constraint edges are added every time, the number of variable points and constraint edges of the initial constraint network remain the same. The number of new constraint edges added is from 10 to 1000 with the interval 10.

Figure 8 shows the results, in which the value of x axis is the sum of the number of initial constraint edges, 3800, and that of the added constraint edges. For example, in the x-axis, 3900 means that 100

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. The flow diagram of the GD algorithm

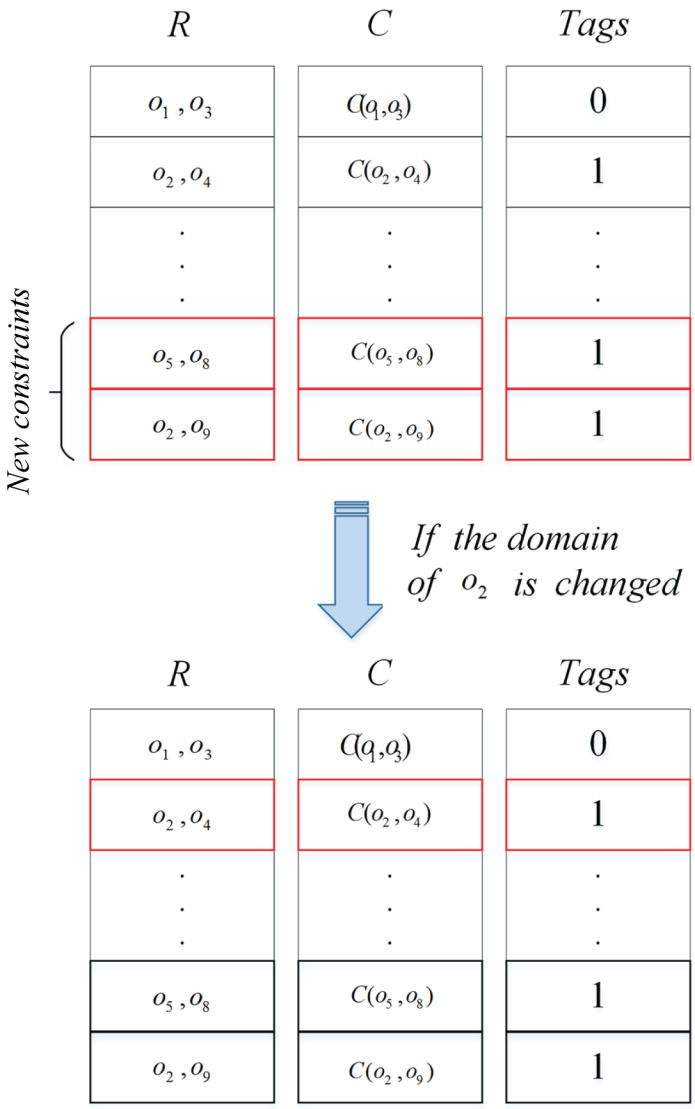


constraint edges are added into the constraint network with 3800 edges; in the same way, 4000 means that 200 constraint edges are added into the network with 3800 edges. Table 1 gives data of six points in Figure 8, which is more specific to show the run time.

From Figure 8 and Table 1, we can see that run time of both two algorithms increases with the increment of added constraint edges. The GD algorithm is more efficient, especially when the number of added constraints is small. For example, in Table 1, when the number of constraint edges is 3810, the run time of GD is faster by two orders of magnitude.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Diagram of the dynamic characteristic



However, the growth rate of GD is larger. Maybe when the number of new constraints added is very large such as 10000, the EU algorithm will be faster than GD. But it is unusual that a large number of constraints is added into the constraint network at the same time in planning and scheduling. Therefore, the GD algorithm has a certain advantage on this case.

4.2. Case 2

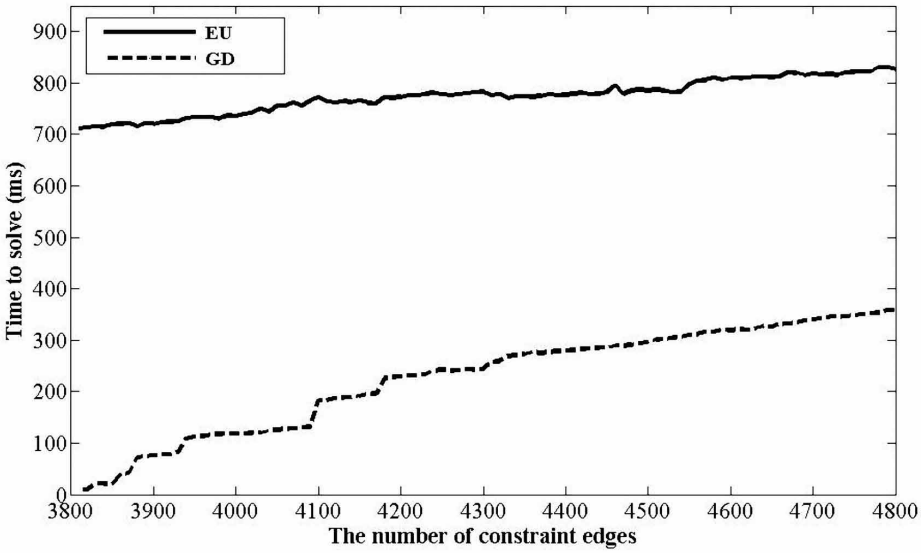
The case 1 has verified that the number of added constraints has influence on run time. In this case, we focus on the influence of the initial network on run time. The initial network is same with that of case 1 where it has 250 variable points and 3800 constraint edges. When new constraint edges are added into the network, the number of points stays same, but the number of constraint edges changes. For example,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Data of six points in case 1

Number of constraint edges		3810	4000	4200	4400	4600	4800
Run time / ms	*EU*	712	736	773	777	809	827
	GD	8	119	229	280	319	359

Figure 8. Comparison of the run times of the GD and EU for case 1



ten constraint edges are added at the first time into the network with 3800 constraints. At the second time, another ten constraint edges are added into the network with 3900 constraints, that is to say the ten constraint edges added last time are not removed. Therefore, the basic number of the constraint network changes every time when ten new constraints are added.

The results of this experiment can be found in Figure 9. Specific data is shown in Table 2. Here we should note that the x axis has the same value with Figure 9, but the way of changing the number of constraint edges is different. The reason has been stated above.

We observe that the efficiency of the GD algorithm is significantly better than that of the EU. The run time of the GD remains about the same while that of EU has a slight increase. From Table 2, we can see that all run time of GD is faster by two orders of magnitude than that of the EU.

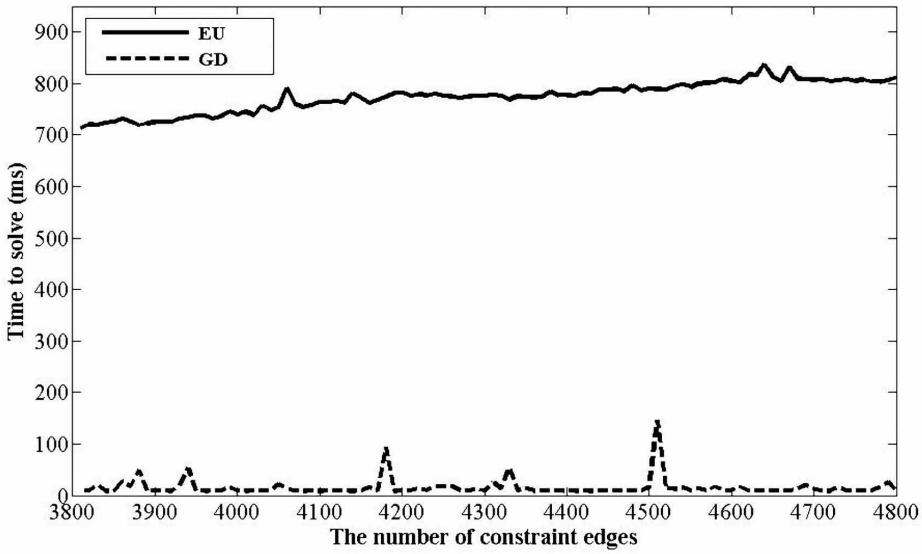
Combined with the result in case 1, we also conclude that the number of added constraints has lager effect on run time of the GD while the number of initial constraints of the network has little influence. However, for the EU algorithm, both the number of added constraints and initial constraints can influence its run time.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Data of six points in case 2

Number of constraint edges		3810	4000	4200	4400	4600	4800
Run time / ms	*EU*	714	740	781	777	805	811
	GD	9	9	10	9	10	10

Figure 9. Comparison of the run times of the GD and EU for case 2



4.3. Case 3

For simulating planning and scheduling process, we design the third experiment. In planning and scheduling, actions with constraints are added into the plan constantly. Therefore, not only new constraint edges are added into the constraint network, but also variable points. In this case, the initial network T_0 contains 45 variable points and 100 constraint edges. The initial network does not change and 10 - 1000 constraint edges with interval 10 are added every time. For example, we add ten constraints into the initial network at the first time and then at the second time we add twenty constraints into the same initial network. Three to five variable points are added into the network with the addition of ten constraint edges.

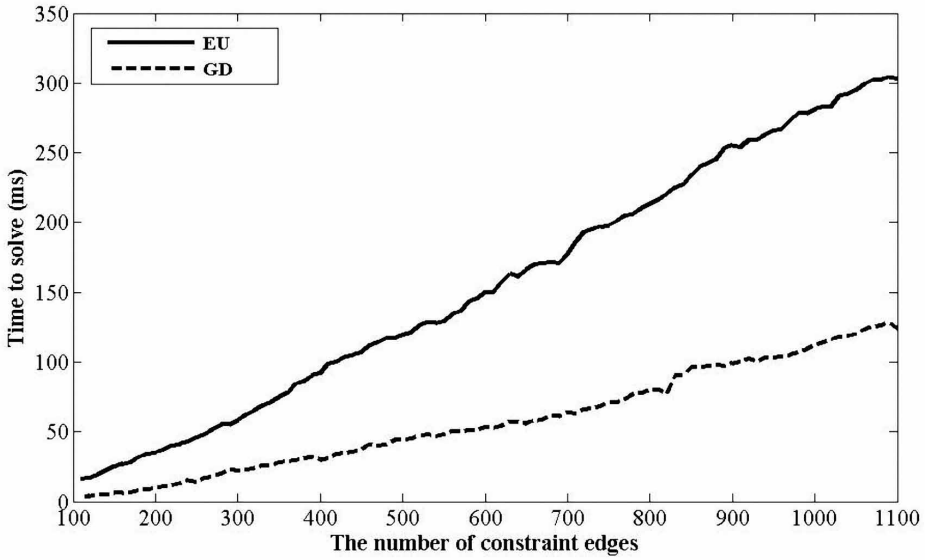
The results are shown in Figure 10 and Table 3 has the specific data. With the increase of the number of added constraints, run time of both algorithms increases and advantages of the GD are highlighted. From Table 3, we can get that run time of the GD algorithm is less than half of that of EU algorithm. Compared the results with that in case 1, we can conclude that the EU algorithm is influenced by both the number of variable points and constraint edges.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Data of six points in case 3

Number of constraint edges		110	300	500	700	900	1100
Run time / ms	*EU*	16	58	119	177	255	303
	GD	2	22	44	64	99	124

Figure 10. Comparison of the run times of the GD and EU for case 3



4.4. Case 4

The case is the one most closely associated with planning and scheduling. The initial network T_0 is the same with that in case 3. But when new constraints and variable points are added into the network, they will not be removed. For example, ten constraint edges and 3 points are added into the network with 100 constraints at the first time and then the new network T_1 has 110 constraints and 48 points. Next, another ten constraint edges and 3 points are added into the network T_1. It is a superposed process.

Figure 11 shows the results of this experiment as a function of the number of constraint edges of the constraint network. Table 4 gives the specific data. Like the case 2, the GD algorithm is not influenced by the added constraints. From this case, we can also get that our algorithm is not influenced by variable points. However, the efficiency of the EU algorithm is associated with the number of variable points and constraint edges, so it increases with the addition of new constraints and points.

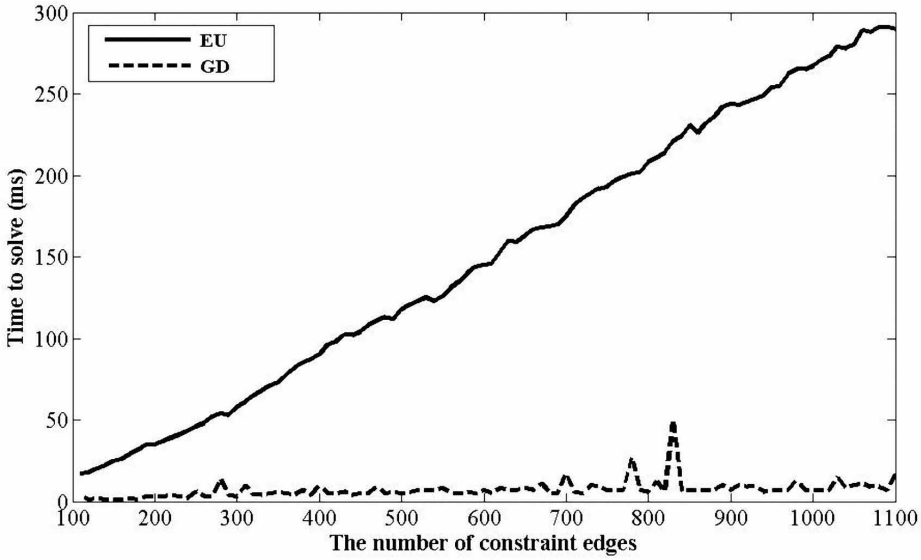
In a conclusion, in all cases above, our dynamic method has obvious advantages in the efficiency over the EU algorithm. It is more suitable for temporal constraint reasoning in planning and scheduling.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Data of six points in case 4

Number of constraint edges		110	300	500	700	900	1100
Run time / ms	*EU*	17	58	118	175	244	290
	GD	3	3	5	17	7	17

Figure 11. Comparison of the run times of the GD and EU for case 4



5. CONCLUSION

In this paper, we focus on temporal reasoning which is one of cognitive capabilities people involve in communicating with others and the reference of everything changing. Here, planning and scheduling is taken as an example to specify our topic and method.

we use our prior work to represent actions and filter redundant values of actions in a 2-D coordination system. For reducing the influence of addition of constraints or actions on the efficiency of temporal reasoning algorithm, a dynamic algorithm with tags based on geometric representation (the GD algorithm) is proposed. This approach was described in detail and also evaluated against the algorithm in EUROPA. In the evaluation, we found that our algorithm is more efficient when new constraint edges and variable points are added into the constraint network.

Future work will implement the GD algorithm on the planning and scheduling system to check the consistency of temporal constraint network and give the minimal domain of the start time and end time of all actions. Furthermore, we want to evaluate the performance of our algorithm in other computer-based systems. Also, the formal properties of the GD algorithm remain to be analyzed.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ACKNOWLEDGMENT

We thank to the support of Ministry of Science and Technology of the People's Republic of China and the National Natural Science Fund Committee.

REFERENCES

Aigner, W., & Miksch, S. (2006). CareVis: Integrated visualization of computerized protocols and temporal patient data. *Artificial Intelligence in Medicine*, *37*(3), 203–218. doi:10.1016/j.artmed.2006.04.002 PMID:16781851

Aini, A., & Salehipour, A. (2012). Speeding up the Floyd–Warshall algorithm for the cycled shortest path problem. *Applied Mathematics Letters*, *25*(1), 1–5. doi:10.1016/j.aml.2011.06.008

Barreiro, J., Boyce, M., Frank, J., Iatauro, M., Morris, P., Smith, T., & Do, M. (2012). EUROPA: A platform for timeline-based AI planning, scheduling, constraint programming, and optimization. *PSTL*, *2012*, 6.

Barták, R., Morris, R. A., & Venable, K. B. (2014). An introduction to constraint-based temporal reasoning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, *8*(1), 1–121. doi:10.2200/S00557ED1V01Y201312AIM026

Barták, R., Salido, M. A., & Rossi, F. (2010). Constraint satisfaction techniques in planning and scheduling. *Journal of Intelligent Manufacturing*, *21*(1), 5–15. doi:10.100710845-008-0203-4

Bresina, J. L. (2015). Activity Planning for a Lunar Orbital Mission. *Paper presented at the AAAI*.

Cervantes, J. A., Rodríguez, L. F., López, S., Ramos, F., & Robles, F. (2013). Cognitive process of moral decision-making for autonomous agents. *International Journal of Software Science and Computational Intelligence*, *5*(4), 61–76. doi:10.4018/ijssci.2013100105

Chandrachoodan, N., Bhattacharyya, S. S., & Liu, K. (2001). Adaptive negative cycle detection in dynamic graphs. *Proceedings of the International Symposium on Circuits and Systems*, Sydney, Australia (pp.163-166). 10.1109/ISCAS.2001.922010

Cooper, M., Maris, F., & Régnier, P. (2014). Monotone temporal planning: Tractability, extensions and applications. *Journal of Artificial Intelligence Research*, 50, 447–485.

Cooper, M. C., Maris, F., & Régnier, P. (2013). Managing temporal cycles in planning problems requiring concurrency. *Computational Intelligence*, *29*(1), 111–128. doi:10.1111/j.1467-8640.2012.00430.x

Cushing, W. A. (2012). *When is temporal planning really temporal? (Doctor)*. Arizona State University.

Duftschmid, G., Miksch, S., & Gall, W. (2002). Verification of temporal scheduling constraints in clinical practice guidelines. *Artificial Intelligence in Medicine*, *25*(2), 93–121. doi:10.1016/S0933-3657(02)00011-8 PMID:12031602

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Gerevini, A. (2005). Incremental qualitative temporal reasoning: Algorithms for the point algebra and the ORD-Horn class. *Artificial Intelligence*, *166*(1), 37–80. doi:10.1016/j.artint.2005.04.005

Iwig, C., Oglesby, J., Shimono, M., Stowers, K., Leyva, K., & Salas, E. (2015). Space flight task contexts for long distance and duration exploration missions: Application to measurement of human automation interaction. *Regional Science and Urban Economics*, *59*(1), 941–945.

Lemai, S. (2004). *IxTeT-eXeC: planning, plan repair and execution control with time and resource management*. Institut National Polytechnique de Toulouse-INPT.

Li, Z., Xu, R., Cui, P., Xu, L., & He, W. (2016). Geometry-based propagation of temporal constraints. *Information Systems Frontiers*.

Mouhoub, M., & Yip, J. (2002). Dynamic CSPs for interval-based temporal reasoning. In Developments in Applied Artificial Intelligence*, LNCS* (Vol. *2358*, pp. 575–585). doi:10.1007/3-540-48035-8_56

Muscettola, N. (1994). HSTS: Integrating planning and scheduling. *Mark Fox & Monte Zweben Intelligent Scheduling*, *15*(5), 2001.

Pani, A., & Bhattacharjee, G. (2001). Temporal representation and reasoning in artificial intelligence: A review. *Mathematical and Computer Modelling*, *34*(1), 55–80. doi:10.1016/S0895-7177(01)00049-8

Planken, L., De Weerdt, M., & Van Der Krogt, R. (2012). Computing all-pairs shortest paths by leveraging low treewidth. *Journal of Artificial Intelligence Research*, *43*, 353–388.

Planken, L., Weerdt, M. D., & Yorke-Smith, N. (2010). Incrementally solving STNs by enforcing partial path consistency. *Proceedings of the International Conference on Automated Planning and Scheduling '10,* Toronto, Ontario, Canada (pp.129-136).

Rit, J. F. (1986). Propagating temporal constraints for scheduling. *Proceedings of the National Conference on Artificial Intelligence*, Philadelphia, PA, US (pp. 383-388).

Stock, O. (1998). *Spatial and temporal reasoning*. Springer Science & Business Media.

Stracuzzi, D. J., Li, N., Cleveland, G., & Langley, P. (2009). Representing and reasoning over time in a symbolic cognitive architecture. *Proceedings of the 31st Annual Meeting of the Cognitive Science Society,* Amsterdam, Netherlands (pp. 2986-2991).

Tsang, E. (2014). *Foundations of constraint satisfaction: the classic text*. Helsinki: BoD–Books on Demand.

Ullberg, J., & Pecora, F. (2012). Propagating temporal constraints on sets of intervals. *Proceedings of ICAPS Workshop on Planning and Scheduling with Timelines (PSTL)*, Atibaia, Brazil.

Wang, Y. (2014). On a novel cognitive knowledge base (CKB) for cognitive robots and machine learning. *International Journal of Software Science and Computational Intelligence*, *6*, 42–64.

Wang, Y., & Patel, S. (2009). Exploring the cognitive foundations of software engineering. *International Journal of Software Science and Computational Intelligence*, *1*(2), 1–19. doi:10.4018/jssci.2009040101

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Xu, R. (2004). Dynamic planning and scheduling algorithm based on temporal constraint network. *Computer Integrated Manufacturing Systems*, *10*, 188–194.

Younes, H. L. S., & Simmons, R. G. (2003). VHPOP: Versatile heuristic partial order planner. *Journal of Artificial Intelligence Research*, *20*, 405–430.

This research was previously published in the International Journal of Software Science and Computational Intelligence (IJSSCI), 8(4); edited by Yingxu Wang; pages 43-59; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 14
Artificial Intelligence for the Identification of Endometrial Malignancies: Application of the Learning Vector Quantizer

Abraham Pouliakis
National and Kapodistrian University of Athens, Greece

Niki Margari
National and Kapodistrian University of Athens, Greece

Effrosyni Karakitsou
University of Barcelona, Spain

Evangelia Alamanou
Tzaneio Hospital, Greece

Nikolaos Koureas
St. Savas Hospital, Greece

George Chrelias
National and Kapodistrian University of Athens, Greece

Vasileios Sioulas
National and Kapodistrian University of Athens, Greece

Asimakis Pappas
MHTERA Maternity Hospital, Greece

Charalambos Chrelias
National and Kapodistrian University of Athens, Greece

Emmanouil G Terzakis
St. Savas Hospital, Greece

Vasileia Damaskou
National and Kapodistrian University of Athens, Greece

Ioannis G Panayiotides
National and Kapodistrian University of Athens, Greece

Petros Karakitsos
National and Kapodistrian University of Athens, Greece

DOI: 10.4018/978-1-7998-0951-7.ch014

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ABSTRACT

Aim of this article is to investigate the potential of Artificial Intelligence (AI) in the discrimination between benign and malignant endometrial nuclei and lesions. For this purpose, 416 histologically confirmed liquid-based cytological smears were collected and morphometric characteristics of cell nuclei were measured via image analysis. Then, 50% of the cases were used to train an AI system, specifically a learning vector quantization (LVQ) neural network. As a result, cell nuclei were classified as benign or malignant. Data from the remaining 50% of the cases were used to evaluate the AI system performance. By nucleic classification, an algorithm for the classification of individual patients was constructed, and performance indices on patient classification were calculated. The sensitivity for the classification of nuclei was 77.95%, and the specificity was 73.93%. For the classification of individual patients, the sensitivity was 90.70% and the specificity 82.79%. These results indicate that an AI system can have an important role in endometrial lesions classification.

INTRODUCTION

The endometrium is a human system with inherent complexity because it undergoes cyclic regeneration under estrogen effects during each menstrual cycle. These effects, as well as sex steroids, oncogene products, growth factors and various peptides (Murphy, Murphy, & Friesen, 1987; Shyamala & Ferenczy, 1981) may cause premalignant and malignant transformation of the endometrium. Early diagnosis is crucial because it is associated with patient management and therapy. However, and despite that in the last decades the incidence of endometrial cancer has increased (Jemal et al., 2006), there is still not a global and well-established screening method for the early detection of endometrial cancer. Moreover, there is no automated procedure.

Artificial Intelligence (AI) techniques are not new in medicine (Adams, Bello, & Dumancas, 2015; Darcy, Louie, & Roberts, 2016; Foran, Chen, & Yang, 2011; Grabe et al., 2010; Karakitsos et al., 1996; Luo et al., 2015; Seffens, Evans, Minority Health, & Taylor, 2015; Su, Xu, He, & Song, 2016). During the last decades, numerous applications have been reported; these involve the use of classical statistical models (Karakitsos et al., 2004), as well as more advanced techniques, such as neural networks. In the field of oncology related medical disciplines there are numerous efforts (Cruz & Wishart, 2006; Hassan, Ruusuvuori, Latonen, & Huttunen, 2015; Karakitsos et al., 1997; Kourou, Exarchos, Exarchos, Karamouzis, & Fotiadis, 2015; Pouliakis et al., 2016; Simões et al., 2014). However, the literature related to endometrial cytological material evaluation by AI and especially publications reporting the comparison of machine learning performance are rather poor.

In this article, a methodology aiming to classify endometrial lesions based on cell nuclei morphometry data will be presented. Initially, the morphometrical features and their biological relation will be described and then the construction of machine learning systems classifying individual nuclei will be outlined. Furthermore, according to the nuclei classification results, a second stage classifier aiming to categorize individual cases will be constructed. The validation of the results will be based on the histological result, which will be considered our gold standard. The applied Machine Learning Algorithm (MLA) was the Learning Vector Quantizer (LVQ) ANN and was selected because it has the capability to provide solutions by creating clusters. The applied classifier was compared with the performance of the cytological outcome.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

MATERIALS AND METHODS

Biological Sample, Material Collection and Diagnoses

For this study, cytological samples from the endometrial cavity were collected from 416 female patients. The biological material was taken using the EndoGyn® Sampler (Biogyn S.n.c., Mirandola, Italy). Subsequently, the sampling device was withdrawn, and the material was immersed into a vial containing CytoLyt® solution (Cytyc Corporation). The use of this methodology (i.e. Liquid Based Cytology - LBC) for material collection and preservation is essential, not only because it facilitates the cytological diagnosis through the microscope, but also provides a clear material for subsequent image analysis, as required by this study and therefore ensures a constant sample quality. The vials with the biological material were used to prepare a single cytological slide via an automated slide processor (ThinPrep® TP2000, Cytyc Corporation). TP2000 prepares a smear in a 2 cm diameter area containing cells in a single layer with almost no overlap in 90 seconds. Each slide was stained with Papanicolaou stain, using an automated staining machine (Varistain® 24-3 Thermo Electron Corporation [formerly Shandon], Runcorn, U.K.).

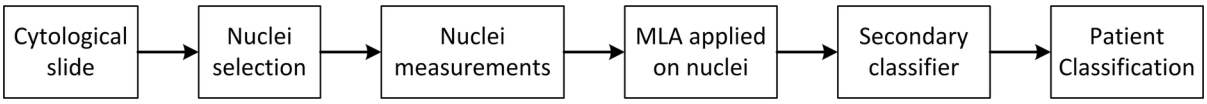
Both the cytological and histological examination results were available for every patient included in this study. The histological examination was performed on biological material harvested via endometrial curettage and/or surgically. Women without histological examination results were excluded from the study. Histological confirmation is imperative, as it is the golden standard and the lack of it makes it impossible to have an accurate final diagnosis and therefore evaluate the results and performance of the proposed methodology.

The cytological diagnosis was based on a new classification scheme introduced in the Department of Cytopathology of University General Hospital "Attikon" (Margari et al., 2016) and the histological diagnosis was in accordance with the International Society of Gynecological Pathology (Blaustein & Kurman, 2011) classification. We only considered cases for which the cytological result was in agreement with the histological conclusion; otherwise, the data were excluded from the study. That was due to the aim of this study to include cell nuclei that have a certain histological type exclusively.

Image Analysis System, Measurements, and Their Biological Relation

The process applied for the classification of individual patients is depicted in Figure 1. The workflow starts with the patient's cytological slide; this is analysed via the image analysis workstation. Following the identification of individual nuclei and their measurements, the MLA was applied; the MLA outcome is the classification of every identified nucleus in the slide as benign or malignant. The final stage is the application of a secondary classifier; this classifies a patient into the benign or malignant category on the basis of the classification results for the nuclei after their measurement and subsequent classification by the MLA.

Figure 1. Work flow for the classification of patients



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

From each slide we selected about 100 representative nuclei via a manual selection/segmentation process; this requires about 10 - 20 minutes, in practice, this is a process of definition of nuclei borderline. The characteristics measurement time following nuclei selection is negligible and was performed massively through software built in the programming language of ImagePro Plus. In total, were 38,326 nuclei from the 416 slides were measured. The features are summarized in Table 1 and details for the applied algorithms can be found in the literature (Baak, 1991; Russ, 1995; Sonka, Hlavac, & Boyle, 1993; Varlatzidou et al., 2011).

These features were selected because they are related to the characteristics that cytopathologists examine during the routine microscopic examination. Specifically, cytopathologists during slide screening examine the nuclear size, boundary abnormalities, elongation of nuclei, and other shape based characteristics of the nuclei. Moreover, nucleus density (i.e. darkness), the distribution of DNA in the nucleus centre or membrane as well as the texture caused by chromatin distribution (i.e. the existence of coarse or fine dots of chromatin in the nucleus) are important diagnostic characteristics (Margari et al., 2016).

Classification Targets

The primary target during diagnosis both in the cytological and histological material is to discriminate between benign and malignant cases. Thus, the data were categorized into these two broad categories: a) benign cases: i.e. histologically normal cases including atrophy, secretory and proliferative endometrium, disordered endometrium, polyps and hyperplasias without atypia, and b) the group of malignant cases composed of the hyperplasias with atypia and the carcinomas (endometrioid, serous and adenocarcinomas).

Training and Test Sets

In order to obtain a similar distribution of the data to the relevant histological categories, 50% of the patients/cases (along with their nuclei) were randomly selected and were used as a training set; the remaining nuclei (from the rest 50% of the cases) formed the test set. The distribution of the cases and the nuclei into the two sets is presented in Table 2. By this methodology, it was ensured that a) representative nuclei obtained via cytological material from all histological groups were included in both sets b) there was no mixing of patients' nuclei between the training set and test set and c) there was a balance between the training and test set.

Artificial Intelligence Systems

The Learning Vector Quantization (LVQ) classifier (Teuvo Kohonen, 1989) was selected as a tool for nuclei classification. It was implemented within the MATLAB v2011b environment and the available programming language (The MathWorks, Inc. Natick, Massachusetts, U.S.A.).

ANNs are complex mathematical models inspired by the human brain structure; they are capable of learning and pattern recognition (Duda, Hart, & Stork, 2001; Theodoridis & Koutroumbas, 2009). This capability is suitable for classification and prediction tasks in numerous practical situations. The core component of ANNs is the artificial neuron (Figure 2). Neurons are inspired and have similarities with brain dendrite cells, where each neuron cell is interconnected with other neurons through synapses. In ANNs neurons perform a relatively simple task: the inputs (i_1- i_n) are multiplied one by one by synaptic weight (w_1- w_n), the products are added, and the result is passed via a nonlinear function which represents the artificial neuron output.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Nuclear morphometry features involved in the study

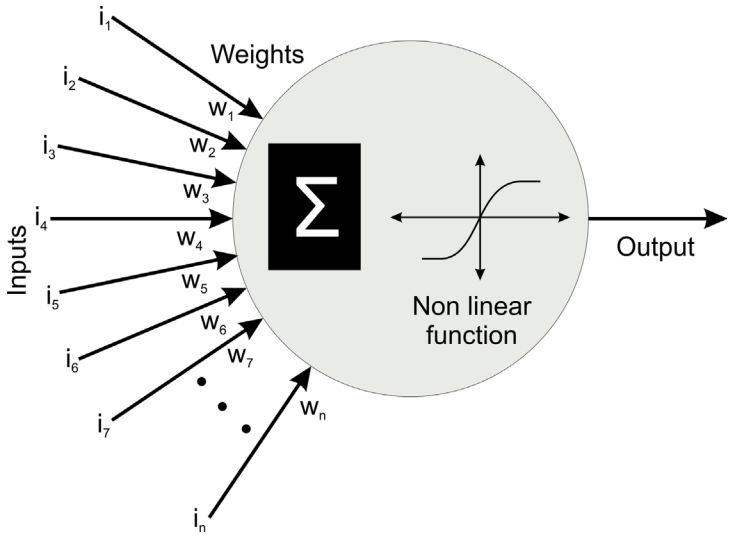
	Feature	Description
Characteristics of nucleus shape	Nucleus area	The area occupied by a nucleus expressed in square μm
	Nucleus major axis	The length of the major axis of an ellipse fitted o the nucleus (expressed in μm)
	Nucleus minor axis	The length of the minor axis of an ellipse fitted o the nucleus (expressed in μm)
	Aspect ratio	The ratio of the major and minor axis considering that an ellipse is fitted to the nucleus (expressed as number without units)
	Maximum caliper	Maximum distance between two perimeter points while the nucleus is between two parallel lines (expressed in μm)
	Minimum caliper	Minimum distance between two perimeter points while the nucleus is between two parallel lines (expressed in μm)
	Average value of calliper	The mean value of the distances measured between two perimeter points while the nucleus is between two parallel lines (expressed in μm)
	Maximum nucleus radius	The highest distance between the nucleus centroid and the nucleus outline (the nucleus centroid is the center of nucleus mass when considering that the nucleus is uniform) (expressed in μm)
	Minimum nucleus radius	The minimum distance between the nucleus centroid and the nucleus outline (expressed in μm)
	Radius ratio	The ratio between the maximum and minimum radius (expressed as a number without units)
	Nucleus perimeter	The length of the nucleus boundary (expressed in μm)
	Nucleus roundness	Defined by the formula (Perimeter*Perimeter)/(4π*Area), it is an indication of the nucleus circularity, an absolutely circular nucleus has roundness equal to 1 while anything not perfectly circular has higher roundness, for example, the roundness of a square is 4/π=1.27 (expressed as a number without units)
	Fractal dimension	The fractal dimension of the nucleus boundary is an indication of repetitive patterns in the nucleus outline (Mandelbrot, 1983)
Density characteristics	Integrated optical density	The sum of the pixel values of the nucleus (when color is not considered) if the nucleus intensity is light and the area is large then integrated optical density becomes higher. Can be calculated the formula (Area)*(Mean optical density)
	Mean value of nucleus red color	The average value of the nucleus pixels values of the red color plane, it is a number between 0 and 255
	Mean value of nucleus green color	The average value of the nucleus pixels values of the green color plane, it is a number between 0 and 255
	Mean value of nucleus blue color	The average value of the nucleus pixels values of the blue color plane, it is a number between 0 and 255
	Mean value of optical density	The average value of the nucleus pixels when color is not considered, it is a number between 0 and 255
	Maximum value of optical density	The maximum value that has pixels inside the nucleus, it is a number between 0 and 255
	Minimum value of optical density	The minimum value that has pixels inside the nucleus, it is a number between 0 and 255
	Standard deviation of optical density	The standard deviation of the nucleus pixels when no color is considered, it is an indication of the range of grey values that has the nucleus
	Margination	The percentage of nucleus pixels that deviate more than 10% from the mean value of the nucleus density
	Heterogeneity	An index of homogeneity between the center of the nucleus and the borders, a homogeneous nucleus (i.e. all nuclei pixels have the same value) has heterogeneity = 0.33 while nuclei with the brighter center have higher heterogeneity. Heterogeneity is indicative of the distribution of DNA in the center or the nucleus membrane

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Distribution of the cases and the related cell nuclei into the training and test sets

	Histological Category	Cases			Nuclei		
		Training set	Test set	Total	Training set	Test set	Total
Benign	Proliferative	36	36	72	3,256	3,593	6,849
	Secretory	22	22	44	2,194	2,396	4,590
	Atrophy	24	24	48	2,081	2,236	4,317
	Disorder	2	2	4	181	180	361
	Polyp	12	12	24	1,214	1,096	2,310
	Hyperplasia without atypia	26	26	52	2,577	2,520	5,097
Malignant	Hyperplasia with atypia	10	10	20	773	870	1,643
	Endometrioid carcinoma	62	62	124	5,503	5,091	10,594
	Serous carcinoma	5	5	10	367	457	824
	Adenocarcinoma	9	9	18	792	949	1,741
	Total	**208**	**208**	**416**	**18,938**	**19,388**	**38,326**

Figure 2. The standard model of an artificial neuron, inputs (i_1- i_n) are multiplied by weights (w_1- w_n), subsequently the products are summarized (block Σ) and finally the output is produced as the response of the nonlinear function



The capability of the ANNs to learn is due to the neurons' weights, and the capability to address complex problems is due to the nonlinear function and architectures that interconnect many neurons together in several layers and the use of training (learning) algorithms.

The LVQ is a supervised classifier; during the training phase LVQ creates partitions of the feature space, each partition is characterized by a vector in the centre of the partition being called codebook vector; the class of the codebook vector characterizes the class of the entire partition. During training the codebook vectors are modified to represent the full feature space; passing of all the training vectors

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

to the classifier several times and by the application of a training algorithm LVQ based MLAs store the knowledge in the network components (i.e. in the codebook vectors) (Teuvo Kohonen, 1989; T. Kohonen, Kangas, Laaksonen, & Torkkola, 1992).

During a test or operational phase, unknown vectors are presented to the LVQ, the class of each unknown vector is determined to be the same to the class of the partition that the unknown vector resides. This class is determined by identifying the codebook vector that is nearest to the unknown vector.

Method for the Classification of the Cases

We constructed case classifiers based on the result of nuclei classification by the LVQ using a method similar to a system proposed in the past for the classification of thyroid lesions (Varlatzidou et al., 2011), endometrial lesions (Makris et al., 2017) as well as for HPV methylation analysis (Kottaridi et al., 2017). This second stage classifier is built on the basis of the percentage of nuclei classified as malignant from the LVQ. During training of second stage classifier, the aim is to find a threshold that optimizes patients' classification based on the training data. For this purpose, we evaluated the case classification results for a broad range of thresholds; starting from 0% and increasing up to 100% (indicating the percentage of the patient's malignant nuclei as assigned by the MLA) the increment step was 0.1%. For each threshold value the number of cases that were true positive was calculated, true negative and false positive or false negative, thus it was possible to calculate the various performance indexes: sensitivity, specificity, Positive Predictive Value (PPV), Negative Predictive Value (NPV), False Positive Rate (FPR), False Negative Rate (FNR), Overall Accuracy (OA), odds ratio and Yuden's index. As an optimum threshold was selected the threshold that maximized the overall accuracy on the training cases. The software for the logic of the cases' classifiers and the performance indices was implemented within the MATLAB environment.

RESULTS AND DISCUSSION

Results are reported in two levels: a) the nuclei classification level where the LVQ was trained by the measurements obtained from the 18,938 nuclei of the training set (208 cases) and tested on the nuclei from the remaining 208 cases (19,388 measured nuclei) and b) the case classification level where the thresholds for the case classifiers were found on the basis of the training results obtained from the LVQ and subsequently tested on the rest of the cases.

Performance on Nuclei Classification

In terms of architectural characteristics of the LVQ ANN, the number of the nodes (codebook vectors) representing the number of clusters used to separate the data was 300, the results of the clusters' layer were fed to a second layer with 2 nodes in order to discriminate into the two target categories (benign and malignant); the number of iterations on the training data set was only 3. As training algorithm, it was used the LVQ.1 algorithm (T. Kohonen et al., 1992). The ANN was trained and evaluated on the training set and subsequently assessed on the test set, the performance is presented in Table 3.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Performance of the LVQ in nuclei classification for the training and test sets

	Training set	Test set
Sensitivity	74.05%	77.95%
Specificity	74.31%	73.93%
PPV	49.59%	50.50%
NPV	89.35%	90.77%
FPR	25.69%	26.06%
FNR	25.95%	22.05%
OA	74.24%	74.96%
Odds Ratio	8.25	10.03
Yuden Index	48.36%	51.89%

Case Classification Results Analysis and Performance Metrics

The case classifier based on the results of the LVQ outcome missed 14 malignant cases and 37 benign cases (Table 4) leading to a sensitivity and specificity of the test set equal to 90.70% and 82.79% respectively (Table 5).

CONCLUSION

Endometrial carcinoma is the most common malignancy of the female genital tract (Papaefthimiou, Symiakaki, Mentzelopoulou, Tsiveleka, et al., 2005; Parazzini, La Vecchia, Bocciolone, & Franceschi, 1991); it affects women at reproductive age and after menopause as well. The mean age of diagnosis in the USA is 60 years (American Cancer Society, 2005), in the UK is 61 years (Soliman et al., 2005) and in European countries is 63 (Colombo et al., 2013). The most frequent symptom of endometrial carcinoma is postmenopausal bleeding (ACOG, 2005) but, only 10% of the women with postmenopausal bleeding have endometrial carcinoma (Iatrakis et al., 1997; Taskin, Bozaci, Seval, & Unlu, 2006); moreover 5-10% of endometrial carcinoma cases are asymptomatic (Smith-Bindman et al., 1998). According to the American National Institute of Health (NIH), and specifically the Surveillance, Epidemiology and

Table 4. Cross tabulation of the assignment of the cases for the LVQ classifier according to the histological group. Results are presented separately for the training and test sets

	Case classifier based on the LVQ ANN			
	Histological result	Benign	Malignant	Grand Total
Training set	Benign	106	16	122
	Malignant	6	80	86
Test Set	Benign	101	21	122
	Malignant	8	78	86
	Grand total	221	195	416

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5 Performance indices of the cases classified on the basis of the LVQ outcomes (training and test sets)

	Training set	Test set
Sensitivity	93.02%	90.70%
Specificity	86.89%	82.79%
PPV	83.33%	78.79%
NPV	94.64%	92.66%
FPR	13.11%	17.21%
FNR	6.98%	9.30%
OA	89.42%	86.06%
Odds Ratio	88.33	46.89
Yuden Index	79.91%	73.48%

End Result (SEER) program, the number of new cases of endometrial cancer for 2016 is estimated to be 60,050 with 10,470 deaths. The age-adjusted rate, based on 2009-2013 data, was 25.4 endometrial cancers per 100,000 women per year. The number of deaths is 4.5 per 100,000 women per year. Moreover, the lifetime risk of developing endometrial cancer is approximately 2.8% for women who will be diagnosed with endometrial cancer at some point during their lifetime (based on 2011-2013 data). In 2013, it was estimated that 635,437 women were living with endometrial cancer in the United States. According to the World Health Organization, there are, worldwide, 188,000 new endometrial carcinomas every year. According to GLOBOCAN, the endometrial cancer age standardized rate for 2012 is 8.2 per 100,000 for the world and for Japan (a single country that is applied endometrial cancer screening) the incidence is significantly reduced to 3.9 per 100,000.

According to various studies (Karakitsos et al., 2002; Papaefthimiou, Symiakaki, Mentzelopoulou, Giahnaki, et al., 2005; Papaefthimiou, Symiakaki, Mentzelopoulou, Tsiveleka, et al., 2005) morphological characteristics of nuclei have an important role in endometrial cytology diagnosis irrelevant of the technique (conventional or LBC); in addition, characteristics of structures formed by the observed cells (contextual characteristics) are important as well.

The application of LBC is in line with the objectivity requirements and quality control and assurance, as advantages of LBC are related to common diagnostic approach, because there is a standardized procedure related to fixation, transfer and cyto-preparation of the biological material; additionally observed cellular overlap is linked to the real 3 dimensional structures and not due to the slide preparation methods; finally the cells are layered in non- overlapping sheets, a fact that facilitates their measurement. The proposed methodology introduces new objectivity in two levels: first cell nuclei are measured in a reproducible method (a procedure facilitated by LBC), and secondly, these characteristics are evaluated via computerized algorithms.

Obviously, the performance of the LVQ on nuclei classification is worse than the performance of the complete system for the classification of cases. For instance, the sensitivity of the test set for nuclei classification was 77.95%, and the PPV 50.50% (Table 3), while the sensitivity of the complete system of case classification was 90.70% and the PPV 78.79% (Table 5). This performance improvement is due to the second stage classification system and the application of the optimal threshold that maximizes the overall accuracy (on the training set).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The presented results are concordant to the wider context of encouraging findings related to the performance of cytology in the diagnosis of endometrial cancer (Norimatsu et al., 2006; Papaefthimiou, Symiakaki, Mentzelopoulou, Giahnaki, et al., 2005; Sams, Currens, & Raab, 2012; Yanoh et al., 2009). Nuclear morphological characteristics seem to play a central role in this context, a finding which has been observed in previous studies as well (Karakitsos et al., 2002; Makris et al., 2017; Papaefthimiou, Symiakaki, Mentzelopoulou, Giahnaki, et al., 2005; Papaefthimiou, Symiakaki, Mentzelopoulou, Tsiveleka, et al., 2005; Pouliakis et al., 2014; Zygouris et al., 2014). The rationale behind this is the fact that carcinogenesis destroys the cellular structure and the in particular nucleoskeleton, thus resulting in nuclear deformations. Shape characteristics reflect such changes, while the second group of density characteristics reflects nucleus DNA changes. Although nuclear morphological characteristics can be evaluated both by conventional and liquid- based cytology, the latter contributes towards standardization due to lysis of mucus and blood using a standardized procedure, standardized fixation, transfer and cytological preparation of the biological material and in this study by improving image analysis due to the reduction of cell overlapping.

Concluding, the proposed objective approach using ANN, gave promising results in the discrimination between malignant and benign endometrial cases; the MLA showed encouraging results; the integration of this technique in the laboratory routine might become a useful adjunct in the future leading towards a more objective diagnosis of endometrial malignancies.

REFERENCES

ACOG. (2005). ACOG practice bulletin, clinical management guidelines for obstetrician-gynecologists, number 65, August 2005: Management of endometrial cancer. *Obstetrics and Gynecology*, *106*(2), 413–425. PMID:16055605

Adams, L. J., Bello, G., & Dumancas, G. G. (2015). Development and Application of a Genetic Algorithm for Variable Optimization and Predictive Modeling of Five-Year Mortality Using Questionnaire Data. *Bioinformatics and Biology Insights*, *9*(Suppl. 3), 31–41. doi:10.4137/BBI.S29469 PMID:26604716

American Cancer Society. (2005). Detailed Guide: Endometrial Cancer: What are the Risk Factors for Endometrial Cancer? Retrieved 20 March 2015, 2015, from http://www.cancer.org/cancer/endometrialcancer/detailedguide/endometrial-uterine-cancer-risk-factors

Baak, J. P. A. (1991). *Manual of quantitative pathology in cancer diagnosis and prognosis*. Berlin, New York: Springer-Verlag.

Blaustein, A., & Kurman, R. J. (2011). *Blaustein's pathology of the female genital tract* (6th ed.). New York, NY: Springer.

Colombo, N., Preti, E., Landoni, F., Carinelli, S., Colombo, A., Marini, C., & Sessa, C. (2013). Endometrial cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology: Official Journal of the European Society for Medical Oncology*, *24*(Suppl. 6), vi33–vi38. doi:10.1093/annonc/mdt353 PMID:24078661

Cruz, J. A., & Wishart, D. S. (2006). Applications of machine learning in cancer prediction and prognosis. *Cancer Informatics*, *2*, 59–77. doi:10.1177/117693510600200030 PMID:19458758

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Darcy, A. M., Louie, A. K., & Roberts, L. W. (2016). Machine Learning and the Profession of Medicine. *Journal of the American Medical Association*, *315*(6), 551–552. doi:10.1001/jama.2015.18421 PMID:26864406

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern classification* (2nd ed.). New York: Wiley.

Foran, D. J., Chen, W., & Yang, L. (2011). Automated image interpretation computer-assisted diagnostics. *Analytical Cellular Pathology (Amsterdam)*, *34*(6), 279–300. doi:10.1155/2011/239761 PMID:22082571

Grabe, N., Lahrmann, B., Pommerencke, T., von Knebel Doeberitz, M., Reuschenbach, M., & Wentzensen, N. (2010). A virtual microscopy system to scan, evaluate and archive biomarker enhanced cervical cytology slides. *Cellular Oncology*, *32*(1-2), 109–119. doi:10.3233/CLO-2009-0508 PMID:20208139

Hassan, S. S., Ruusuvuori, P., Latonen, L., & Huttunen, H. (2015). Flow Cytometry-Based Classification in Cancer Research: A View on Feature Selection. *Cancer Informatics*, *14*(Suppl. 5), 75–85. doi:10.4137/CIN.S30795 PMID:27081305

Iatrakis, G., Diakakis, I., Kourounis, G., Sakellaropoulos, G., Rammos, G., Ladopoulos, J., ... Tzingounis, V. (1997). Postmenopausal uterine bleeding. *Clinical and Experimental Obstetrics & Gynecology*, *24*(3), 157. PMID:9478305

Jemal, A., Siegel, R., Ward, E., Murray, T., Xu, J., Smigal, C., & Thun, M. J. (2006). Cancer statistics, 2006. *CA: a Cancer Journal for Clinicians*, *56*(2), 106–130. doi:10.3322/canjclin.56.2.106 PMID:16514137

Karakitsos, P., Kyroudes, A., Pouliakis, A., Stergiou, E. B., Voulgaris, Z., & Kittas, C. (2002). Potential of the learning vector quantizer in the cell classification of endometrial lesions in postmenopausal women. *Analytical and Quantitative Cytology and Histology*, *24*(1), 30–38. PMID:11865947

Karakitsos, P., Megalopoulou, T. M., Pouliakis, A., Tzivras, M., Archimandritis, A., & Kyroudes, A. (2004). Application of discriminant analysis and quantitative cytologic examination to gastric lesions. *Analytical and Quantitative Cytology and Histology*, *26*(6), 314–322. PMID:15678613

Karakitsos, P., Stergiou, E. B., Pouliakis, A., Tzivras, M., Archimandritis, A., Liossi, A., & Kyrkou, K. (1997). Comparative study of artificial neural networks in the discrimination between benign from malignant gastric cells. *Analytical and Quantitative Cytology and Histology*, *19*(2), 145–152. PMID:9113307

Karakitsos, P., Stergiou, E. B., Pouliakis, A., Tzivras, M., Archimandritis, A., Liossi, A. I., & Kyrkou, K. (1996). Potential of the back propagation neural network in the discrimination of benign from malignant gastric cells. *Analytical and Quantitative Cytology and Histology*, *18*(3), 245–250. PMID:8790840

Kohonen, T. (1989). *Self-organization and associative memory* (3rd ed.). Berlin, New York: Springer-Verlag. doi:10.1007/978-3-642-88163-3

Kohonen, T., Kangas, J., Laaksonen, J., & Torkkola, K. (1992, June 7-11). LVQPAK: A software package for the correct application of Learning Vector Quantization algorithms. *Paper presented at International Joint Conference on Neural Networks IJCNN '92*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kottaridi, C., Kyrgiou, M., Pouliakis, A., Magkana, M., Aga, E., Spathis, A., ... Karakitsos, P. (2017). Quantitative Measurement of L1 Human Papillomavirus Type 16 Methylation for the Prediction of Preinvasive and Invasive Cervical Disease. *The Journal of Infectious Diseases*, *215*(5), 764–771. doi:10.1093/infdis/jiw645 PMID:28170039

Kourou, K., Exarchos, T. P., Exarchos, K. P., Karamouzis, M. V., & Fotiadis, D. I. (2015). Machine learning applications in cancer prognosis and prediction. *Computational and Structural Biotechnology Journal*, *13*, 8–17. doi:10.1016/j.csbj.2014.11.005 PMID:25750696

Luo, H., Ye, H., Ng, H. W., Shi, L., Tong, W., Mendrick, D. L., & Hong, H. (2015). Machine Learning Methods for Predicting HLA-Peptide Binding Activity. *Bioinformatics and Biology Insights*, *9*(Suppl. 3), 21–29. doi:10.4137/BBI.S29466 PMID:26512199

Makris, G. M., Pouliakis, A., Siristatidis, C., Margari, N., Terzakis, E., Koureas, N., ... Karakitsos, P. (2017). Image analysis and multi-layer perceptron artificial neural networks for the discrimination between benign and malignant endometrial lesions. *Diagnostic Cytopathology*, *45*(3), 202–211. doi:10.1002/dc.23649 PMID:28160459

Mandelbrot, B. B. (1983). The fractal geometry of nature (Updated and augm. ed.). New York: W.H. Freeman.

Margari, N., Pouliakis, A., Anoinos, D., Terzakis, E., Koureas, N., Chrelias, C., ... Karakitsos, P. (2016). A reporting system for endometrial cytology: Cytomorphologic criteria-Implied risk of malignancy. *Diagnostic Cytopathology*, *44*(11), 888–901. doi:10.1002/dc.23605 PMID:27653446

Murphy, L. J., Murphy, L. C., & Friesen, H. G. (1987). Estrogen induces insulin-like growth factor-I expression in the rat uterus. *Molecular Endocrinology (Baltimore, Md.)*, *1*(7), 445–450. doi:10.1210/mend-1-7-445 PMID:2484720

Norimatsu, Y., Shimizu, K., Kobayashi, T. K., Moriya, T., Tsukayama, C., Miyake, Y., & Ohno, E. (2006). Cellular features of endometrial hyperplasia and well differentiated adenocarcinoma using the Endocyte sampler: Diagnostic criteria based on the cytoarchitecture of tissue fragments. *Cancer*, *108*(2), 77–85. doi:10.1002/cncr.21734 PMID:16463402

Papaefthimiou, M., Symiakaki, H., Mentzelopoulou, P., Giahnaki, A. E., Voulgaris, Z., Diakomanolis, E., ... Karakitsos, P. (2005). The role of liquid-based cytology associated with curettage in the investigation of endometrial lesions from postmenopausal women. *Cytopathology*, *16*(1), 32–39. doi:10.1111/j.1365-2303.2004.00224.x PMID:15859313

Papaefthimiou, M., Symiakaki, H., Mentzelopoulou, P., Tsiveleka, A., Kyroudes, A., Voulgaris, Z., ... Karakitsos, P. (2005). Study on the morphology and reproducibility of the diagnosis of endometrial lesions utilizing liquid-based cytology. *Cancer*, *105*(2), 56–64. doi:10.1002/cncr.21025 PMID:15751019

Parazzini, F., La Vecchia, C., Bocciolone, L., & Franceschi, S. (1991). The epidemiology of endometrial cancer. *Gynecologic Oncology*, *41*(1), 1–16. doi:10.1016/0090-8258(91)90246-2 PMID:2026352

Pouliakis, A., Karakitsou, E., Margari, N., Bountris, P., Haritou, M., Panayiotides, J., Koutsouris, D., & Karakitsos, P. (2016). Artificial Neural Networks as Decision Support Tools in Cytopathology: Past, Present, and Future. *Biomedical Engineering and Computational Biology, 7*. doi:10.4137/BECB.S31601

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Pouliakis, A., Margari, C., Margari, N., Chrelias, C., Zygouris, D., Meristoudis, C., ... Karakitsos, P. (2014). Using classification and regression trees, liquid-based cytology and nuclear morphometry for the discrimination of endometrial lesions. *Diagnostic Cytopathology*, *42*(7), 582–591. doi:10.1002/dc.23077 PMID:24273089

Russ, J. C. (1995). *The image processing handbook* (2nd ed.). Boca Raton: CRC Press.

Sams, S. B., Currens, H. S., & Raab, S. S. (2012). Liquid-based Papanicolaou tests in endometrial carcinoma diagnosis. Performance, error root cause analysis, and quality improvement. *American Journal of Clinical Pathology*, *137*(2), 248–254. doi:10.1309/AJCPLFBK1A2XJDQI PMID:22261450

Seffens, W., Evans, C., Minority Health, G. N., & Taylor, H. (2015). Machine Learning Data Imputation and Classification in a Multicohort Hypertension Clinical Study. *Bioinformatics and Biology Insights*, *9*(Suppl. 3), 43–54. doi:10.4137/BBI.S29473 PMID:27199552

Shyamala, G., & Ferenczy, A. (1981). The effect of sodium molybdate on the cytoplasmic estrogen and progesterone receptors in human endometrial tissues. *Diagnostic Gynecology and Obstetrics*, *3*(4), 277–282. PMID:7200003

Simões, P. W., Izumi, N. B., Casagrande, R. S., Venson, R., Veronezi, C. D., Moretti, G. P., Rocha, E. L. d., Cechinel, C., Ceretta, L. B., Comunello, E., Martins, P. J., Casagrande, R. A., Snoeyer, M. L., & Manenti, S. A. (2014). Classification of Images Acquired with Colposcopy Using Artificial Neural Networks. *Cancer Inform, 13*, 119-124. doi:10.4137/CIN.S17948

Smith-Bindman, R., Kerlikowske, K., Feldstein, V. A., Subak, L., Scheidler, J., Segal, M., ... Grady, D. (1998). Endovaginal ultrasound to exclude endometrial cancer and other endometrial abnormalities. *Journal of the American Medical Association*, *280*(17), 1510–1517. doi:10.1001/jama.280.17.1510 PMID:9809732

Soliman, P. T., Oh, J. C., Schmeler, K. M., Sun, C. C., Slomovitz, B. M., Gershenson, D. M., ... Lu, K. H. (2005). Risk factors for young premenopausal women with endometrial cancer. *Obstetrics and Gynecology*, *105*(3), 575–580. doi:10.1097/01.AOG.0000154151.14516.f7 PMID:15738027

Sonka, M., Hlavac, V., & Boyle, R. (1993). *Image processing, analysis, and machine vision* (1st ed.). London, New York: Chapman & Hall Computing. doi:10.1007/978-1-4899-3216-7

Su, J., Xu, X., He, Y., & Song, J. (2016). Automatic Detection of Cervical Cancer Cells by a Two-Level Cascade Classification System. *Analytical Cellular Pathology (Amsterdam)*, *9535027*. doi:10.1155/2016/9535027 PMID:27298758

Taskin, S., Bozaci, E. A., Seval, M. M., & Unlu, C. (2006). Transvaginal sonographic assessment of endometrial thickness and endometrial morphology in postmenopausal bleeding. *International Journal of Gynaecology and Obstetrics: the Official Organ of the International Federation of Gynaecology and Obstetrics*, *92*(2), 155–156. doi:10.1016/j.ijgo.2005.10.023 PMID:16360156

Theodoridis, S., & Koutroumbas, K. (2009). *Pattern recognition* (4th ed.). London: Academic Press.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Varlatzidou, A., Pouliakis, A., Stamataki, M., Meristoudis, C., Margari, N., Peros, G., ... Karakitsos, P. (2011). Cascaded learning vector quantizer neural networks for the discrimination of thyroid lesions. *Analytical and Quantitative Cytology and Histology*, *33*(6), 323–334. PMID:22590810

Yanoh, K., Norimatsu, Y., Hirai, Y., Takeshima, N., Kamimori, A., Nakamura, Y., ... Shiraishi, T. (2009). New diagnostic reporting format for endometrial cytology based on cytoarchitectural criteria. *Cytopathology*, *20*(6), 388–394. doi:10.1111/j.1365-2303.2008.00581.x PMID:18657157

Zygouris, D., Pouliakis, A., Margari, N., Chrelias, C., Terzakis, E., Koureas, N., ... Karakitsos, P. (2014). Classification of endometrial lesions by nuclear morphometry features extracted from liquid-based cytology samples: A system based on logistic regression model. *Analytical and Quantitative Cytopathology and Histopathology*, *36*(4), 189–198. PMID:25291856

This research was previously published in the International Journal of Reliable and Quality E-Healthcare (IJRQEH), 7(2); edited by Anastasius Moumtzoglou; pages 37-50; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 15
Expanding Bloom's Two-Sigma Tutoring Theory Using Intelligent Agents:
Application to Management Education

Owen P. Hall Jr.
Pepperdine University, USA

ABSTRACT

This article describes how management education is engaged in significant programmatic reforms in response to the business community's call for web-savvy, problem-solving graduates. Web-based intelligent tutors provide a readily accessible vehicle for enhancing business students' learning performance as well as prepare them for the rigors of the global marketplace. A primary goal of these AI-based systems is to approach Bloom's two-sigma learning performance standard. Bloom found that average students tutored one-to-one with mastery learning techniques performed two standard deviations better than students who learned via conventional teaching methods. Intelligent tutors can also be used to identify students at risk, to formulate appropriate intervention plans, and to support team learning. The purpose of this article is to highlight the growing potential for using intelligent tutors to enhance student and team learning opportunities and outcomes and to outline strategies for implementing this revolutionary process throughout the management education community of practice.

INTRODUCTION

Business leaders, accrediting bodies, and management educators alike are calling for a paradigm shift in the content and delivery of management education because of globalization, unprecedented economic uncertainty, changing demographics, and new learning technologies (Buil, 2016; Hall, 2013; Rubin, 2013; Thomas, 2012). Business schools are beginning to recognize that digital technology can play an important role in improving both the learning process and learning outcomes (Bowen, 2014; Mohapatra, 2015; Spada, 2014). Specifically, the rapid growth in Web 2.0+ technologies has ushered in a new era

DOI: 10.4018/978-1-7998-0951-7.ch015

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

for management education including: virtual learning environments, intelligent tutors, web-based collaborative learning groups, and crowdsourcing. Additionally, the ubiquitous availability of mobile devices has encouraged educators to utilize this technology as a vehicle for accessing context-based materials, without the constraints or limitations of time, location and pace (Aljohani, 2012). The increased use of learning technologies can help narrow the gap between a graduates' skill set and the ever-changing requirements of business. For example, an MBA Roundtable survey found that many business schools do not emphasize the skills desired by most hiring managers (Hanover, 2013). These included evidence-based decision-making, ethics, and a focus on results and quality.

A number of the concerns such as the lack of integration and too little emphasis on problem finding, international dimensions, entrepreneurship, and practical experience have seemingly not been widely addressed. Nor have many of the other competencies considered important by employers, e.g., motivation and commitment to the firm, creativity, quality focus, customer focus, ethics–integrity, teamwork, flexibility, and interpersonal skills (Herrington, 2013).

With globalization under increased scrutiny and an unpredictable economy, the business community is increasingly focusing on adaptive and sustainable positioning strategies (Waddock, 2013). The business community is looking for web-savvy graduates that are both problem-solving and entrepreneurial oriented. This is the ongoing challenge faced by the management education community. Even the business accrediting organizations have gotten into the act of critiquing business education. A report from the Association to Advance Collegiate Schools of Business (AACSB, 2016) found:

- A larger share of degree-based education will be delivered in flexible formats (modular, part-time) across providers, with students having more control over their curriculum
- Business schools will struggle to align the pace of curriculum development with the pace of evolution in business practice
- A shift toward more experiential learning and business engagement
- Business schools will be increasingly called on to serve the common good

Digital technology provides a vehicle that can be used for addressing these trends. More specifically, providing customized learning opportunities can enhance both student learning outcomes and preparation for the transition to the business community. The mastery of management education material lies at the heart of this proposition. In a 1984 study, Benjamin Bloom found that the average student tutored one-to-one using mastery learning techniques performed two standard deviations better than students who learn via conventional instructional methods (Bloom, 1984). The task at hand is to find technological-based methods of instruction that are scalable, which is not the case with the traditional Bloom model, and that approach the effectiveness of one-to-one human tutoring. Intelligent tutors represent one promising possibility (Anastassakis, 2014; Niu, 2012; Yeh, 2013). These systems, which consist of a combination of hardware devices and software programs, are designed to provide learners with both customized content and instructional feedback at a performance-driven pace. Intelligent tutors can be used to replicate the role of a one-to-one tutor on a massive scale and thus move toward the goal of approaching Bloom's Two-Sigma standard. To be effective, an intelligent tutoring system requires a pedagogical framework that is based on the latest learning constructs and that can justify its choice of activities, presentation

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

formats, and decisions (Bagheri, 2015). In recent years, a range of educational technologies have emerged throughout the business school universe including: homework support agents, simulations, massive open online courses (MOOCs), and conditional content release systems (Bassi, 2014; Chin, 2010).

Learner's attitude toward affective computing, agent tutor's expressiveness, emotion recognition accuracy, number of emotions recognized by agent tutor, pedagogical action and ease of system use have significant influence on learner's satisfaction (Mao, 2010).

These same learning support technologies can also be applied to team learning. For example, sensing technologies are now able to assess each team member's behavior and level of engagement. This data along with individual and team outcome metrics provides the agent with the required inputs to enhance the overall learning process. Accurate individual and team statistics (e.g., exam performance) turn out to be key factors for the agent to identify optimal learning strategies for enhancing overall performance (Holden, 2011). This paper is organized as follows: 1) a review of intelligent tutoring systems (ITS) in management education and 2) an overview on ITS implementation strategies. This article's primary contribution to management education is to identify how intelligent agents can be used to enhance student learning opportunities and outcomes and thus better prepare students for the growing demands of the business universe.

Management education has always striven to inculcate a culture that is both practical and responsive to society's needs. Every tension the industry has experienced in its past continues to affect it today—and will reverberate into its future (Spender, 2016).

INTELLIGENT TUTORS IN MANAGEMENT EDUCATION

The changing demands placed on the management education universe by the business community continues at a torrid pace. Many students now entering a management degree program, especially at the graduate level, often lack the prerequisite skills and know-how to successfully compete and complete their course of studies, thus leading to less than satisfactory employment results (Pianin, 2014).

Nearly three-quarters of hiring managers complain that millennials – even those with college degrees – aren't prepared for the job market and lack an adequate work ethic (Bentley University, 2014).

These competency issues can be addressed using the intelligent tutoring paradigm outline in this paper. An intelligent tutoring systems (ITS) is defined as a learning construct that provides feedback and support during each phase of the learning process or exercise (e.g., hints) as well as an overall assessment at the end of the process. ITS represent a natural extension to the classical computer assisted learning systems in that they are step-phased oriented allowing the student to advance to the next learning level only after they demonstrate a certain level of proficiency. Additionally, these systems can further enhance the learning process by making it more personalized based on the learner's characteristics, background, desired learning pace, and business requirements (Dixit, 2013; Xu, 2014). ITS can acquire real-time information and knowledge via the web to support the learning process. The fundamental challenge is

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

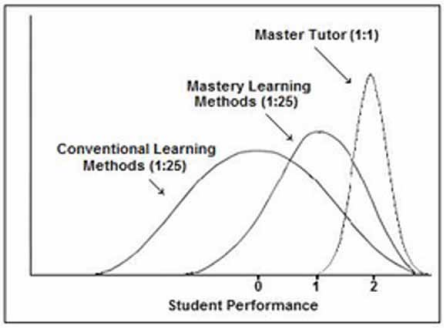
to identify the most effective ITS delivery mode for a given learning application. Again, the goal of incorporating tutors into the curriculum is to improve learning outcomes, which in turn, should lead to greater employment options (Zahid, 2013).

If learning goals are not properly aligned with managerial competencies, education and accreditation processes based on those goals are for naught. (Costigan, 2015).

Bloom Basics

Some thirty years ago, Benjamin Bloom observed that students receiving individualized tutoring in combination with corrective instruction tended to perform two standard deviations better than students receiving conventional classroom instruction. This suggests that the tutored students out-performed 98 percent, or about two-sigma, of the students engaged in a traditional learning environment (Guskey, 2007). Bloom also pointed out that master-based individual and team tutoring do not lend themselves to economies of scale, attesting that one-to-one tutoring would be too expensive for most schools to deploy on a wide scale basis. Accordingly, Bloom called for alternative learning strategies, which when used in combination might approach his two-sigma target. These included but were not limited to: 1) mastery learning, 2), proactive classroom participation, 3) time on task, and 4) cooperative engagement (i.e., team collaboration). Of these, mastery learning was found to yield the largest effect size (i.e., one-sigma). In mastery learning, the class size remains the same as in the traditional classroom setting, but learning pace for each student can vary. The goal of mastery learning is for each student to achieve general competency for a given topic (e.g., descriptive statistics) before moving on to the next, more advanced subject (e.g., inferential statistics). In mastery learning, the student is given the time needed to achieve the specified level of expertise, even though it may take longer than other students. Two challenges associated with the mastery learning paradigm are: 1) preserving consistency among multiple course sections and between different business disciplines and 2) maintaining the course pace for the entire class. Specifically, what happens to a student that is unable to complete the course requirements by the end of the term? Figure 1 highlights the basic construct behind Bloom's Two-Sigma paradigm. The numbers in parenthesis represent the ratio of instructor to class size. Notice that in both the one-sigma and two-sigma cases, the variability in performance is substantially reduced. Significantly increasing performance while reducing variability lies at the heart of the Bloom two-sigma paradigm and as such is viewed by many as the *Holy Grail* of modern education.

Figure 1. Bloom's two-sigma paradigm



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Intelligent Tutoring Systems

The proposal to utilize computer-based tutors to enhance the learning process is not new, but was first introduced in the late 1990s (Rickel, 1999). Today, the state of the art in intelligent tutoring systems (ITS) has improved dramatically since those early days (Rus, 2013). Specific advancements include:

- Inhabits the environment with the student, thus providing increased potential for "physical" collaboration (e.g., telepresence)
- Communicates nonverbally, for example through gestures and facial expressions
- Tracks learners in new ways, for example their visual attention and physical movements
- Maintains learning records over the course of the study program
- Establishes progressive learning goals
- Provides for team collaboration

Typically, ITS embodies a wide range of web-based learning formats including videos, simulations, interactive cases, homework, games, voiceover slide shows and streaming text. Two recent meta-studies examined the efficacy of ITS (Kulik, 2016; Ma, 2015). The former study found, based on 50 evaluations, that performance scores were raised to 0.66 standard deviations from conventional levels, i.e., 50th to the 75th percentile. This study did observe that performance increases varied based on locally developed or standardized testing methodologies. The latter study, based on a sample size of 107, found a 0.42 sigma increase in student performance across a broad range of subject domains. The first study utilized Glass's effect size estimator while the second study employed Hedges's g. Both these statistics are related to the more common Cohen's d, which translates 0.2 to a small effect, 0.5 to a moderate effect, and 0.8 to a large effect. This scale has been extended to 2.0, very large effect, to encompass Bloom's standard (Sawilowsky, 2009). Unfortunately, neither study broke out the results by business category (e.g., finance, marketing). MBA programs, by their very nature, not only embody a broad range of diverse disciplines (e.g. ethics to analytics) but also are highly interdisciplinary. Accordingly, finding the effect per discipline therefore, is particularly important given the continuing interdisciplinary trends associated with many management education programs (Jacob, 2015). The average of the two meta-studies yielded a 0.5 sigma improvement in student performance. It should be noted that these two meta-studies contain performance data that is based on earlier intelligent tutor technology and usage plans. Furthermore, both studies involved K-12 and post-secondary students.

A basic tenant of this paper is that intelligent tutors will continue to improve student performance as new learning technologies are introduced that tend to mimic more human behaviors and qualities.

The past fifteen years have seen considerable AI advances in education. Though quality education will always require active engagement by human teachers, AI promises to enhance education at all levels, especially by providing personalization at scale. Similar to healthcare, resolving how to best integrate human interaction and face-to-face learning with promising AI technologies remains a key challenge. Over the next fifteen years the use of intelligent tutors and other AI technologies to assist teachers is likely to expand significantly, as will learning based on virtual reality applications (Stone, 2016).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Recent research suggests that the new generation of management education students face serious concentration, engagement, and socialization issues (Karakas, 2015). This, of course, represents a major issue for business schools in terms of preparing students for the job market. To that end, a well-designed ITS should be able to modify both the lesson plan and content level according to specific student characteristics (Clemente, 2014; Diziol, 2013; Felijakowski, 2014). Faculty, administrators, designers and practitioners of management education need to develop coherent and reliable assessments of this ITS learning paradigm. Agents can support an adaptive, exploratory learning environment for learning using a multi-layered interface design. Each layer's mechanisms and supportive resources are tailored to facilitate learning for a given situation. Based on a student's background and performance, the agent would select the most appropriate interface layer for that context. For example, if a student is having difficulty mastering a particular subject (e.g., basic statistics) as detected by testing or self-assessment, then the tutor would prescribe specific additional learning content to the student. This content can take the form of videos, computing tutorials, or simulations.

Education-based intelligent tutoring systems should possess the following four basic characteristics: autonomy, pro-activity, adaptability, and sociability. An important feature of an effective tutor is to identify something positive in a student's response (i.e., positive reinforcement). Furthermore, intelligent tutors can be used to identify students at risk and to formulate an appropriate intervention plan for students that are challenged. Specifically, Purdue University's Signals Project has applied analytics to the problem of improving student success within a course and, hence, to improving retention and graduation rates (Pistilli, 2012). This project has led to the development of student success data mining algorithms with intervention messages sent to students based on performance via dashboards. These tools, which often include decision trees, rule induction, and neural networks, can be employed to 1) select student groups with similar characteristics and reactions to learning strategies; 2) detect student misuse and lurking; 3) identify students who, in certain types of test formats, are hint-driven; 4) locate students who exhibit low motivation and find alternate means of reaching them; and 5) predict probable student outcomes (MacFayden, 2010; Shahiri, 2015). The domain variables for data mining usually include previous course grades, class test grades, individual assignment performance, admission exam scores, and team assessments (Baradwaj, 2011).

An integrative education is about blending knowledge and skills from different disciplinary areas; putting knowledge in to practice, considering multiple perspectives to advance collaborative problem solving, adapting the skills learned in one situation to problems encountered in another, reflecting upon connections made over time among academic, co-curricular and pre-professional experiences and across-the-curriculum integration of skills with learning in disciplinary or interdisciplinary settings (Khan, 2015).

The emergence of big data and cloud analysis technologies can be used to develop enhanced ITS, which would assess why students are failing or struggling in a particular course. Typically, students fail or struggle for a variety of reasons including a lack of active participation and inadequate background in the subject material. Intelligent tutors should assess whether the lack of performance is an isolated incident or a reoccurring phenomenon. Big data-driven ITS can be transferred to new student populations, which in turn can be used to provide even more customized learning interventions and educational innovations. In this regard, ITS can adapt to individual differences in both student know-how and level of experience (Baker, 2013, Suhadi, 2016).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Team Tutoring

Student team work has had a long history throughout management education. The empirical evidence suggests that students engaged in team activities (e.g., homework, presentations) exhibit improved learning outcomes (Hsiung, 2012). Many students view their team as a personnel trainer. In this team oriented learning universe, faculty become designers, coordinators, moderators, mediators, and mentors, rather than pure instructors, whereas students assume a proactive role including supporting team activities (Lee, 2015). Students with working group experience tend to be better prepared for placement in the business community (Erenli, 2011).

Teamwork competencies are becoming a vital part of a student's preparation for the global workplace. Identifying specific challenges and describing how students engage offer useful insights. This in turn provides added perspectives into how teaching might further support student learning (Gonzalez, 2014).

The United States military has seen an increase in the use of intelligent tutors, particularly as related to team training (Spain, 2012). The U.S. military's training requirements continue to expand in scope and complexity due to the ongoing introduction of new technology and the changing nature of warfare. In response to these challenges, the military is increasingly relying on an adaptive training paradigm. For example, the United States' Tactical Language and Culture Training Systems assists students to acquire functional skills in foreign languages and cultures in a flexible and timely manner (Johnson, W., 2009). Presented below are some specific military-based team-oriented organizational principles for enhancing the learning process, which can be applied to management education (Soetanto, 2012):

- Obliged to rely on one another to achieve specific goals
- Tasked with doing their share of the group assignment
- Charged with providing one another with feedback and support
- Encouraged to develop and practice trust-building leadership skills
- Motivated to facilitate changes in the team process as conditions warrant

The overall aim of these principles is to enable group members to monitor their progress and that of their peers in performing specific learning exercises. In addition, practitioners of this strategy can utilize intelligent tutors with information about each trainee's problem-solving behavior and group engagement for assessment and instructional purposes (Caballé, 2010). Currently, most military ITS focus on well-defined domains where procedures have explicit steps, ordering, measures, and standards (Sottilare, 2010). Ill-defined domains remain a serious issue in both the military space as well as the business universe. For example, hostage negotiations can be considered "ill-defined" since there could be several successful outcomes based on various strategies. One potential approach to overcome these obstacles is to combine team-based simulations with intelligent agents. Team-based business simulations, which are used extensively throughout the management education community, tend to provide limited external guidance while intelligent tutors usually require a structured knowledge domain. However, combining these two pedological platforms where agents provide advice, based on game conditions, to each team over the course of the simulation could prove to be extremely useful from a team learning perspective, particularly in business simulations involving multi-targets (e.g., market share, stock price).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In some instances, multi-agents can be used where each agent learns from other agents, which will help improve the overall consulting tasks (Jumadinova, 2014). This combination approach is illustrative of the experiential learning process, which is receiving increased attention throughout schools of business.

Experiential learning is an interdisciplinary approach based on management, education and psychology. It is particularly powerful in connection with management education as it is perceived to be effective in support of training and education in fields as diverse as talent management, leadership performance, competence development, change management, community involvement, volunteering, cross-cultural training and entrepreneurship (Bevan, 2012).

MANAGEMENT EDUCATION ITS EXAMPLES

The previously cited two meta-studies contained a paucity of business management courses. Another meta-study on ITS, which was limited to secondary students not enrolled in a professional program, reported an effect size, using Hedges's g, of a modest 0.26 for three business related subjects (Steenbergen-Hu, 2014). To be clear, many of the reports included in the meta-study were over ten years old, which obviously does not reflect the current state-of-the art in ITS technology. Some additional insights into the efficacy of ITS as pertains to specific business management courses is presented in Table 1. Unlike other educational sectors, the number of studies that have assessed the impact of ITS within the management education universe remains somewhat sparse. The first example was extracted from the first cited meta-study (Phillips, 2011). In this case, both groups received support beyond the traditional textbook and lecture modalities. The resultant effect size of 0.32 suggests a small to moderate impact. However, had the control group only received the traditional instructional support the effect mostly likely would have been somewhat larger. The next study investigated how enforcing the prerequisites impacted course outcomes in intermediate accounting. The prerequisites were enforced using a proficiency test administered online in a system that also offered tutoring to remedy any learning shortfalls, this study reported an effect size of 0.30 (Sargent, 2013). The application of an ITS in an information technology course revealed an effective size of 0.50 (Wang, 2014). The results indicated even larger performance improvement for those students with low-levels of prior subject knowledge. The next example involved an operations management course where the experimental student group was exposed to a series of tutorial videos (Winch, 2015). The resultant effect size was a moderate 0.66. In the supply chain management (SCM) course, the authors noted that the simulation-based, intelligent tutoring system, yielded an effect size of 0.78 (Bhuiyan, 2015). In this ITS application, a survey of both groups was conducted after completing the course. The survey results indicated that students participating in the simulation reported statistically higher levels of confidence and satisfaction than those that did not use the ITS. The use of an ITS in an introductory management course revealed a modest effect size of 0.37. Perhaps more significant was that those students exhibiting a minimal understanding of the material covered in the class at the pre-test phase showed the most improvement after utilizing the ITS (Stonebraker, 2016). The next example exhibited an ITS that offered an adaptive learning strategy in a database design course that selected the type of subsequent task based on the level of assistance the student received in previous problems (Najar, 2016). The results indicated that students using the ITS learned significantly more than their peers who were exposed with a fixed sequence of problem sets ($d = 0.58$). The last example

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Selected ITS business course examples

Discipline	Effect Size[1]
Financial Accounting (Phillips,2011)	0.32
Intermediate Accounting (Sargent, 2013)	0.30
Information Technology (Wang, 2014)	0.53
Operations Management (Winch, 2015)	0.66
Supply Chain Management (Bhuiyan, 2015)	0.78
Introductory Management (Stonebraker, 2016)	0.63[2]
Database Design (Najar, 2016)	0.58
Project Management (González-Marcos, 2016)	0.67

[1]Cohen's d was used in these examples, because Hedges's g was not reported in all cases.
[2]Effect size from the one significant experimental group.

featured a course in project management experimental group where the experimental student group had access to an ITS-based Project Management Information System (González-Marcos, 2016). The resultant Cohen's d of 0.67 was based on final grade scores.

The overall performance improvements reported in Table 1 are consistent with the outcomes delineated in the previously reference meta-studies. The general upward trend in effect size over time does suggest that the one-sigma goal of mastery learning is within reach. However, the lack of standardized assessment methods and small to modest sample sizes makes assessment comparisons somewhat challenging. Furthermore, some uses of ITS within a program of management education have not result in improved learning performance results (Hahn, 2013). This case was very similar in design and content to the first example given in Table 1. The lack of consistent assessment methodologies is one area that is ripe for future research. In contrast, a completely different set of performance outcomes was reported in a fourth ITS meta-study (VanLehn, 2011). The analysis revealed an effect size of 0.76 for ITS and a corresponding effect size of a mere 0.79 for human tutoring (Cohen's d). Clearly, if these observations stand the implications on both further research and system implementation are significant. To this end, one of the significant problems in evaluating ITS performance is establishing and using a consistent definition of content mastery and standardized assessments, e.g., 90 percent.

Students must understand and use learning targets, set their own learning goals, select effective learning strategies, and assess their own learning progress. And as students develop into more confident and competent learners, they become motivated (energized) to learn, increasingly able to persist during demanding tasks and to regulate their own effort and actions when they tackle new learning challenges (Hills, 2015).

IMPLEMENTATION STRATEGIES

The rate at which technology is adopted in organizations varies considerably. The Rogers' diffusion of technological innovation model, which was first used for predicting the introduction rate of hybrid corn seed, provides a helpful paradigm for better understanding the adoption process (Rogers, 2003). Apply-

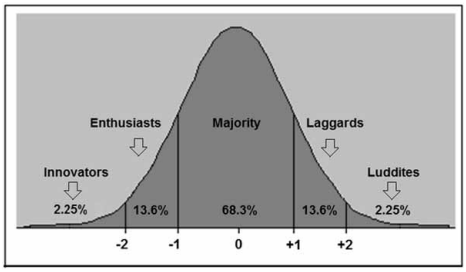
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ing Rogers' model to academia suggests that a significant proportion of faculty tend to react slowly in embracing technological innovation (Fathema, 2015; Sahin, 2006). Figure 2 depicts a slight variation to the basic Rogers' adoption distribution model. This format incorporates a fifth category – Luddites. In a commercial setting the ultimate incorporation of new technologies throughout the organization, for the most part, centers on economics. In academia, this is not always the case. There are some faculty members who most likely will not adopt new learning technologies no matter the nature of the incentives and disincentives and, in fact, may erect road blocks just as their namesakes did some 200 plus years ago. This reality is the primary reason for adding the Luddite category. The Rogers' model is based on the Empirical Rule gleaned from statistics. The Empirical Rules states that approximately 68 percent of the population resides within ± one standard deviation from the mean, while 95.5 percent are contained within ± two standard deviations. Figure 2 shows less than 16 percent of the faculty can be regarded as innovators or enthusiasts during the early phase of new technology adoption. Innovators and early adopters tend to embrace new learning modalities based on their own initiative and technical background without the need for significant institutional support. The emergence of ITS may provide sufficient motivation for these two groups to embrace this technology. Clearly, this percentage does not represent a critical mass. According to Rogers, the majority, which constitutes a little over two-thirds of the population, will begin to embrace new technology only after some period of consideration. This can present a problem with the speed at which new educational technology is being introduced (Rajasingham, 2011).

In general, adoption will significantly increase if the majority are recognized as a distinct group within the community and are made a part of the planning and policy making process. A technological innovation, such as ITS, must be demonstrated as both an effective and easily applied solution to a specific learning need. The motivation for the majority to adopt new learning technologies may not be the same across disciplines. Innovation by laggards is more likely to occur through ongoing majority involvement and continued administrative support in the form of incentives. Laggards will adopt only if they are certain that it will not fail and that there will be no negative consequences to themselves (e.g., negative student evaluations). The luddite segment, as the name implies, can be the source of potentially significant disruptions in the overall adoption process.

User based agents significantly help reduce the administration duties of the course and focus on responding to users' questions or preparing training materials. The agents have been used in many areas of learning. Yet, there remains a myriad of contexts where agents can be incorporated to make learning more efficient and fundamentally change the way education is being delivered (Jawahar, 2015).

Figure 2. Modified Rogers' adoption distribution model



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

One way to address these adoption concerns in an educational setting is to employ systems that are easy to use, that minimize potential problems, and that yield early success. In this regard, two promising ITS implementing vehicles are learning management system (LMS) based conditional release technologies and standalone authoring design tools. Typically, both of these approaches utilize an inner and outer loop format. The function of the inner loop is to execute the steps associated with the task at hand (e.g., taking a quiz). The inner loop can provide feedback and hints at each step. The inner loop also assesses the student's competence in terms of task performance, which is used by the outer loop to select the next task that is appropriate for the student, e.g., return to review the current material or advance to new more challenging material. Additional loops can be added (Niraula, 2015). However, both system complexity and implementation problems will dramatically increase.

Conditional Release Technologies (CRT)

Conditional release technologies (CRT) incorporated within an LMS provide a specific vehicle for offering students a wide variety of learning options based on progressively delivered, customized content (Li, 2015; Gardner, 2011). CRT can be defined as data-driven algorithms that release content based on student performance and characteristics. It is generally recognized that students entering a program of management education possess a wide range of experiences and capabilities, CRT allows customization of content based upon such differences. Three of the most significant advantages of integrating the ITS based CRT inside the LMS are: 1) ease of implementation, 2) maintenance of student performance records over the course and beyond, and 3) flexibility in providing customized content to meet the needs of a wide portfolio of management education programs. CRT offer the capability to assist students in overcoming specific learning challenges by assessing their performance and providing customized content.

The same technology that supports learning activities can also gather data that can be used for assessment. As students work, the system can capture their inputs and collect evidence of their problem-solving sequences (Bienkowski, 2012).

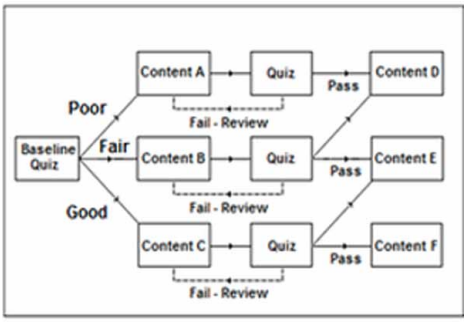
In this regard, students tend to participate more in learning systems that are content rich and that feature a wide variety of delivery modes, which is a hallmark of conditional release technologies. Recent research on the use of CRTs revealed that students (Fisher 2014):

- Recognized that videos provided the highest degree of choice since they can select which segments to watch at their convenience
- Selected tutorials that gave them the highest level of bandwidth and engagement without technological interference or latency
- Noted that on-line quizzes were the most challenging activity
- Expressed more interested in tutors that were based on their level of technological readiness

Figure 3 depicts a simplified CRT structure. In this application setting the student takes a baseline quiz to determine the appropriate content level based on test performance. Based on the initial baseline quiz results specific content is then presented. After reviewing the web-based content the student is presented with a quiz. If the student passes the quiz a new learning module is displayed, otherwise, the student is directed back to the previous material for further review or given additional supplemental

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. CRT design schematic



materials based on the quiz results. In some applications, additional review material can be provided via a variety of delivery modes, e.g., videos, interactive games, slideshow. This process can be further refined by including the student's prior background with the subject material and specific characteristics. While quizzes represent an important and easily implemented performance rubric they can also be augmented by discipline based simulations as conditions warrant. Figure 3 also illustrates incorporation of the two loops in the ITS. In this example, the quiz represents the inner loop while returning the student to the recently review content or providing the student with new content highlights the outer loop. One approach for incorporating additional information, beyond mere quiz scores, into the ITS is through the use of data mining models (Raju, 2016). For example, Bayesian networks can provide enhanced accuracy in the diagnostic of student' knowledge possession (Ramírez-Noriega, 2017). The Bayesian network is continuously updated as additional information about the student's interaction with the ITS is gleaned. These revised probabilities offer further insights into student learning behavior, which in turn allows for the identification of more effective learning content.

Authoring Tools

Authoring tools represent another approach for developing and deploying ITS. These systems have been shown to decrease the amount of time and resources needed for the development of intelligent tutoring systems (Adetokunbo, 2014). Typically, the tutor designer has two classic options: acquire the authoring systems or obtain an already existing tutoring system that can then be modified to meet the specific needs of the learning application. An authoring system, which are usually designed for non-programmers, is a software package that provides the capability to create courseware, multimedia applications, and the corresponding navigational appliances. Authoring tools provide the course designer with more flexibility and thus more creativity than the traditional CRT approach. This capability allows the designer to incorporate more complex and dynamic learning alternatives. Presented in the following are some specific attributes associated with an effective authoring system (Brawner, 2012; Koedinger, 2012).

- Supports the development of cognitive-based, example-tracing, and collaboration tutors
- Accommodates a wide range of application and technical domains
- Incorporates a dynamic and varied learning environment
- Utilizes the latest empirical evidence and learning techniques
- Skeptical of the student's level of understanding

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Rule-based cognitive tutors are designed to capture how students perceive, think, acquire knowledge, and solve problems and are being increasingly used throughout higher education as reported earlier. Example-tracing tutors assess learning performance by comparing student responses against examples of correct and incorrect responses. Typically, this class of tutors provide step-by step guidance on complex problems and situations, such as business cases, and tend to be less costly to develop than rule based systems (Aleven, 2016). Collaborative authoring tools represent a natural and important extension to the classical ITS in the sense that it provides for multiple user assessments (Olsen, 2014). To author a collaborative tutor each of the steps to create an individual tutor are followed for each member of the collaboration. Depending on the type of collaboration activities and roles depends on if different tutor interfaces and behavior graphs need to be made for each student in the group or if the same files can be used.

While the goal of this article is not to evaluate authoring tools, Table 2 lists of some of the better-known development systems. This list is by no means comprehensive. Furthermore, as mentioned earlier, it may be more cost-effective to acquire an existing ITS application than develop one using one or more of these tools.

Illustrative Business Course ITS Example

The Accounting Education Change Commission (AECC), among other bodies, has been active in calling for the incorporation of both innovative teaching as well as active learning paradigms into accounting courses (Larsen, 2014). Accounting principles courses, in particular, are well suited to an ITS learning environment, because they tend to be highly rule-based. To illustrate the use of an authoring tool based ITS, consider the following learning objectives for an entry-level accounting principles course:

- Prepare and analyze an income statement, balance sheet, and statement of cash flow.
- Evaluate the annual report of a public company.
- Assess emerging issues such as revenue recognition and restructuring costs.
- Appreciate the ethical, global and regulatory issues of accounting.
- Understand and use the information technology of accounting.

The ITS can be used to identify and deliver the most effective web-based resources to as illustrated in Table 3.

Table 2. Selected ITS authoring tools

System	Source
Situated Pedagogical Authoring (SPA)	University of Southern California
Cognitive Tutor Authoring Tools (CTAT)	Carnegie-Mellon University
Extensible Problem Specific Tutor (PST)	Iowa State University
Authoring Software Platform (ASPIRE)	University of Canterbury (NZ)
AutoTutor Script Authoring Tools (ASAT)	University of Memphis
Generalized Intelligent Framework for Tutoring (GIFT)	US Army Research Lab.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Example delivery modes

Topic	Content Delivery Mode
Income Statement	Computing Template
Annual Report	PowerPoint Presentation
Net Present Value (NPV)	Consultation
Ethics	Simulation
Information Technology	Streaming Video

Figure 4. Illustrative authoring tool format (NPV Example)

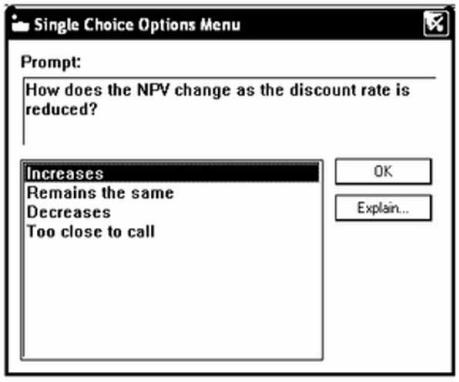


Figure 4 shows a sample consultation for a net present value learning application, which was developed using an authoring tool. In this example application, the student is guided through a series of prompts regarding the case and explanations are provided for each prompt. The consultation can be taken more than once since, among other things, many of the prompts are randomized. The use of this expert system is not limited to entering students requiring preparatory work but can also be used as a refresher by continuing students. For example, a student that is struggling with the notion of net present value (NPV) would be directed to a consultative system that outlines the basic NPV process and that provides more specific content.

Implementation Guidelines

Presented in the following are some guiding principles for implementing ITS in support of management education (Blas, 2014; Brew, 2014; Starke-Meyerring, 2008):

- **Early success**: Learning innovations, like intelligent tutors, are most likely to be accepted and used by the majority of management educators if success is experienced early on. Early on success extends to the peer network, both within and outside the institution, thereby magnifying the impact on adoption and diffusion.
- **On-going peer support**: Live and virtual peer support not only serves as a repository of encouragement; but also contributes diffusion throughout an educational community.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- **Real learning activities**: Initiatives designed to introduce and use ITS should address real activities and requirements. Furthermore, they should support adaptive learning strategies based on students' responses and backgrounds.
- Using intelligent tutors to access to information and resources can address many of today's management educational challenges and requirements.
- **Design philosophy**: Intelligent tutors are fit for purpose. The purpose determines how the tutors are structured and used.
- **Multi-user collaborations**: Intelligent tutors should support collaborative learning environment which improves student and team opportunities and outcomes.

To reap the full potential of integrating ITS throughout management education universe, the design must be more than simply "attaching" a series of standalone tutors to the curriculum. Generally, the program/course structure must be redesigned to provide a seamless transition between face-to-face learning and the ITS. Some specific administrative implementation tasks include: 1) training faculty for successful system deployment and usage, 2) providing high quality and consistent system access, 3) setting specific performance goals and metrics, 4) preparing students for entry and ongoing use, 5) sustaining system operation, and 6) establishing and maintaining the overall culture (Denton, 2013; Duff, 2014.). An internalized only approach may not take advantage of the ongoing developments in technology. Developing institutional partners represents one possible strategy that helps overcome some of these issues. The current state-of-the-art in ITS technology will continue to improve and learning applications, some of which are at the present time unknown, will continue to unfold. For example, the recent partnership between IBM-Watson and Pearson Publishing, which is designed to apply artificial intelligence to enhancing the education learning process (Luckin, 2016).

The ultimate goal of the field of artificial intelligence in education (AIED) is not to promote artificial intelligence, but to promote education. The leading systems in AIED seem to represent a different paradigm than the classic paradigm of intelligent tutoring systems (Baker, 2016).

CONCLUSION

The management education community is engaged in significant programmatic reforms in light of the business community's call for web-savvy, problem solving graduates. Today the demands from a globalized economy are causing a sea change in the way business education is being delivered. The traditional one-size-fits-all educational approach of the past is being replaced with a customized and flexible learning paradigm that focuses on student outcomes and performance. Web-based intelligent tutoring systems provide a readily accessible vehicle for enhancing business students' learning performance as well as preparing them for the rigors of the global marketplace. To support these goals, intelligent tutors should be designed to perform three basic functions: 1) perception of the dynamic conditions in the learning environment, 2) formulation of problem solving actions, and 3) implementation of the identified actions. Furthermore, they need to exhibit both a user-friendly and adaptive format as a vehicle for enhancing student learning outcomes.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The empirical evidence to date suggests that ITS performance, in the management education universe, is approaching the one-sigma standard (i.e., mastery learning). However, Bloom's two-sigma goal still remains elusive. Moreover, the actual use of ITS on a wide scale basis throughout the community of practice is still in an incipient stage. This should not be surprising since the rate at which technology is adopted in organizations varies considerably. The Rogers' diffusion of technological innovation model provides a helpful paradigm for better understanding the adoption process in business management. Applying Rogers' model to academia suggests that a significant proportion of faculty tend to react slowly in embracing technological innovation. Rogers' model suggests that less than 20 percent of the faculty can be regarded as innovators or enthusiasts during the early phase of any new technology adoption, such as intelligent tutoring systems. Some specific research areas to help accelerate the ITS adoption process include improving the effectiveness of authoring tools to reduce development costs and time, enhancing ITS assessment methodologies to provide improved consistency, reducing students and student teams cognitive overload, and enriching adaptation methodologies with a focus on student learning behaviors.

REFERENCES

AACSB. (2016). *A collective vision for business education*. AACSB International.

Adetokunbo, A., A., & Patel, A. M. (2014). A metamodel for designing an intelligent tutoring systems authoring tool. *Computer and Information Science*, *7*(2), 82–98.

Aleven, V., McLaren, B. M., Sewall, J., van Velsen, M., Popescu, O., Demi, S., ... Koedinger, K. R. (2016). Example-tracing tutors: Intelligent tutor development for non-programmers. *International Journal of Artificial Intelligence in Education*, *26*(1), 224–269. doi:10.100740593-015-0088-2

Aljohani, N., & Davis, H. (2012, October). Learning analytics in mobile and ubiquitous learning environments. In *Proceedings of the 11th World Conference on Mobile and Contextual Learning*, Helsinki, Finland.

Anastassakis, G., & Panayiotopoulos, T. (2014). A transparent and decentralized model of perception and action for intelligent virtual agents. *International Journal of Artificial Intelligence Tools*, *23*(4). doi:10.1142/S0218213014600203

Bagheri, M. (2015). Intelligent and Adaptive Tutoring Systems: How to Integrate Learners. *International Journal of Education*, *7*(2), 1–16. doi:10.5296/ije.v7i2.7079

Baker, R. (2016). Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education*, *26*(2), 600–614. doi:10.100740593-016-0105-0

Baker, R., Corbett, A. T., & Gowda, S. M. (2013). Generalizing automated detection of the robustness of student learning in an intelligent tutor for genetics. *Journal of Educational Psychology*, *105*(4), 946–956. doi:10.1037/a0033216

Baradwaj, B., & Pal, S. (2011). Mining educational data to analyze students' performance. *International Journal of Advanced Computer Science and Applications*, *2*(6), 62–69.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bassi, R., & Sula, A. (2014). Software agents in large scale open e-learning: A critical component for the future of massive online courses (MOOCs). In *Proceedings of the International Conference on Intelligent Networking and Collaborative Systems* (pp. 184–188). Salerno, September. 10.1109/INCoS.2014.15

Bentley University. (2014). PreparedU Survey. Retrieved from http://www.bentley.edu/newsroom/latest-headlines/mind-of-millennial

Bevan, D., & Kipka, C. (2012). Experiential learning and management education. *Journal of Management Development*, *31*(3), 193–197. doi:10.1108/02621711211208943

Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing teaching and learning through educational data mining & learning analytics: An issue brief. U.S. Dept. of Education, Center for Technology in Learning.

Blas, M., & ... . (2014). A portal of educational resources: Providing evidence for matching pedagogy with technology. *Research in Learning Technology*, *22*(1), 1–27.

Bloom, B. (1984). The 2-Sigma problem: The search for methods of group instruction as effective as one-on-one tutoring. *Educational Researcher*, *13*(6), 4-16. doi:10.3102/0013189X013006004

Bowen, W. (2014). *Higher education in the digital age*. Princeton University Press.

Brawner, K., Holden, H., Goldberg, B., & Sottilare, R.(2012, December). Recommendations for modern tools to author tutoring systems. In *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference*, Orlando, FL.

Brew, A., & Cahir, J. (2014). Achieving sustainability in learning and teaching initiatives. *The International Journal for Academic Development*, *19*(4), 341–352. doi:10.1080/1360144X.2013.848360

Buil, M., Aznar, J., Galiana, J., & Rocafort-Marco, A. (2016). An explanatory study of MBA students with regards to sustainability and ethics commitment. *Sustainability*, *8*(3), 280–293. doi:10.3390u8030280

Zahid, M. M., Naeem, A., Ahmad, F., & Rehman, A. (2013). Business education curriculum and management skills: A viewpoint of students and managers. *Social Sciences and Humanities*, *4*(4), 431–441.

Caballé, S., Daradoumis, T., Xhafa, F., & Conesa, J. (2010). Enhancing knowledge management in online collaborative learning. *International Journal of Software Engineering and Knowledge Engineering*, *20*(4), 485–497. doi:10.1142/S0218194010004839

Chin, D., Dohmen, I. M., Cheng, B. H., Oppezzo, M. A., Chase, C. C., & Schwartz, D. L. (2010). Preparing students for future learning with teachable agents. *Educational Technology Research and Development*, *58*(6), 649–681. doi:10.100711423-010-9154-5

Clemente, J., Ramírez, J., & de Antonio, A. (2014). Applying a student modeling with non-monotonic diagnosis to intelligent virtual environment for training/instruction. *Expert Systems with Applications*, *41*(2), 508–520. doi:10.1016/j.eswa.2013.07.077

Fisher, L. S., Gardner, J. G., Brinthaupt, T. M., & Raffo, D. M. (2014). Conditional release of course materials: Assessing best practice recommendations. *Journal of Online Learning and Teaching / MERLOT*, *10*(2), 228–239.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Costigan, R., & Brink, K. (2015). Another perspective on MBA program alignment: An investigation of learning goals. *Academy of Management Learning & Education*, *14*(2), 260–276. doi:10.5465/amle.2013.0315

Denton, D., et al. (2013, July). Closing the two-sigma gap: Eight strategies to replicate one-to-one tutoring in blended learning. In *Proceedings of the 10th Annual Sloan Consortium: Blended Learning Conference & Workshop*, Milwaukee, WI.

Dixit, A., Dass, M., Pendleton, G. C., & Lundstrom, W. J. (2013). Transforming marketing education of the future: The role of intelligent agent technologies in enhancing student learning. *Atlantic Marketing Journal*, *2*(3), 1–13.

Diziol, D., Walker, E., Rummel, N., & Koedinger, K. R. (2013). Using intelligent tutor technology to implement adaptive support for student collaboration. *Educational Psychology Review*, *22*(1), 89–102. doi:10.100710648-009-9116-9

Duff, S., Thangarajah, J., & Harland, J. (2014). Maintenance goals in intelligence agents. *Computational Intelligence*, *30*(1), 71–114. doi:10.1111/coin.12000

Erenli, K., & Ortner, G. (2011). Collaborative and social learning using virtual worlds: Preparing students for virtually anything. *International Journal of Advance Corporate Learning, 4*(3).

Fathema, N., Shannon, D., & Ross, M. (2015). Expanding the technology acceptance model (TAM) to examine faculty use of learning management systems (LMSs) in higher education institutions. *MERLOT Journal of Online Learning and Teaching*, *11*(2), 210–232.

Felijakowski, K., & Kosinski, R. (2014). Opinion formation and self-organization in a social network in an intelligent agent system. *Acta Physica Polonica B*, *45*(11), 2123–2134. doi:10.5506/APhysPolB.45.2123

Gardner, J., G., Fisher, L. S., & Raffo, D. M. (2011). Best practices for using conditional release in online classes. *International Journal of Instructional Technology and Distance Learning*, *8*(10), 3–14.

Soetanto, R., Childs, M., Poh, P., Austin, S., & Hao, J. (2012). Global multidisciplinary learning in construction education: Lessons from virtual collaboration of building design teams. *Civil Engineering Dimension*, *14*(3), 173–181.

González-Marcos, A., Alba-Elías, F., Navaridas-Nalda, F., & Ordieres-Meré, J. (2016). Student evaluation of a virtual experience for project management learning. *Computers & Education*, *102*, 172–187. doi:10.1016/j.compedu.2016.08.005

Guskey, T. (2007). Closing achievement gaps: Revisiting Benjamin S. Bloom's earning for mastery. *Journal of Advanced Academics*, *19*(1), 8–31. doi:10.4219/jaa-2007-704

Hahn, W., Fairchild, C., & Dowis, W. B. (2013). Online homework managers and intelligent tutoring systems: A study of their impact on student learning in the introductory financial accounting classroom. *Issues in Accounting Education*, *28*(3), 513–535. doi:10.2308/iace-50441

Hall, R., Agarwal, R., & Green, R. (2013). The future of management education in Australia: Challenges and innovations. *Education + Training*, *55*(4/5), 348–369. doi:10.1108/00400911311326009

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hanover Research. (2013). MBA course trends. Market Insight Center.

Herrington, J., & Arnold, D. (2013). Undergraduate business education: It's time to think outside the box. *Journal of Education for Business*, *88*(2), 202–209. doi:10.1080/08832323.2012.672934

Hills, T. (2015). Crowdsourcing content creation in the classroom. *Journal of Computing in Higher Education*, *27*(1), 47–67. doi:10.100712528-015-9089-2

Hsiung, C. (2012). The effectiveness of cooperative learning. *Journal of Engineering Education*, *101*(1), 119–137. doi:10.1002/j.2168-9830.2012.tb00044.x

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education. London: Pearson.

Jacob, W. (2015). Interdisciplinary trends in higher education. *Palgrave Communications*, *1*(15001), 1–5.

Jawahar, J. (2015). A study on role of intelligent agent in education. *International Journal of Applied Engineering Research*, *10*(76), 497–501.

Johnson, W., & Valente, A. (2009). Tactical language and culture training systems: Using AI to teach foreign languages and cultures. *AI Magazine*, *30*(2), 72–83. doi:10.1609/aimag.v30i2.2240

Jumadinova, J., Dasgupta, P., & Soh, L.-K. (2014). Strategic capability learning for improving multi-agent collaboration in ad hoc environments. *IEEE Transactions on Systems, Man, and Cybernetics. Systems*, *44*(8), 1003–1014. doi:10.1109/TSMC.2013.2285527

Karakas, F., Manisaligil, A., & Sarigollu, E. (2015). Management learning at the speed of life: Designing reflective, creative, and collaborative spaces for millennials. *International Journal of Management Education*, *13*(3), 237–248. doi:10.1016/j.ijme.2015.07.001

Koedinger, K., Corbett, A., & Perfetti, C. (2012). The knowledge-learning-instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science*, *36*(5), 757–798. doi:10.1111/j.1551-6709.2012.01245.x PMID:22486653

Kulik, J., & Fletcher, J. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*, *86*(1), 42–78. doi:10.3102/0034654315581420

Lee, H., & Mehta, M. (2015). E-Mentoring through a network of practice on Facebook. *International Journal of Knowledge-Based Organizations*, *5*(1), 34–45. doi:10.4018/ijkbo.2015010103

Li, N., Matsuda, N., Cohen, W. W., & Koedinger, K. R. (2015). Integrating representation learning and skill learning in a human-like intelligent agent. *Artificial Intelligence*, *219*, 67–91. doi:10.1016/j.artint.2014.11.002

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, *106*(4), 901–918. doi:10.1037/a0037123

MacFayden, L., & Dawson, S. (2010). Mining LMS data to develop an early warning system for educators: A proof of concept. *Computers & Education*, *54*(2), 588–599. doi:10.1016/j.compedu.2009.09.008

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Mao, X., & Li, Z. (2010). Agent based affective tutoring systems: A pilot study. *Computers & Education*, *55*(1), 202–208. doi:10.1016/j.compedu.2010.01.005

Bhuiyan, T., Peng, W. W., & Mahmud, I. (2015). Measuring learning motivation of students in supply chain management games setting: A case study of Innov8.0 game. *Problems and Perspectives in Management*, *13*(4), 81–90.

Mohapatra, S. (2015). Business school education and technology: A case study. *Education and Information Technologies*, *20*(2), 335–346. doi:10.100710639-013-9287-3

Najar, A., Mitrovic, J., & McLaren, B. (2016). Learning with intelligent tutors and worked examples: Selecting learning activities adaptively leads to better learning outcomes than a fixed curriculum. *User Modeling and User-Adapted Interaction*, *26*(5), 459–491. doi:10.100711257-016-9181-y

Niu, F., Zhang, C., Ré, C., & Shavlik, J. (2012). Elementary: Large-Scale knowledge-based construction via machine learning and statistical inference. *International Journal on Semantic Web and Information Systems*, *8*(3), 24–41. doi:10.4018/jswis.2012070103

Olsen, J., Belenky, D. M., Aleven, V., Rummel, N., Sewall, J., & Ringenberg, M. (2014). Authoring tools for collaborative intelligent tutoring system environments. In Intelligent Tutoring Systems, *LNCS* (Vol. *847*, pp. 523–528). doi:10.1007/978-3-319-07221-0_66

Phillips, F., & Johnson, B. (2011). Online homework versus intelligent tutoring systems: Pedagogical support for transaction analysis and recording. *Issues in Accounting Education*, *26*(1), 87–97. doi:10.2308/iace.2011.26.1.87

Pianin, E. (2014). The surprising reason college grads can't get a job. *The Fiscal Times*, January, 19. http://www.thefiscaltimes.com/Articles/2014/01/29/Surprising-Reason-College-Grads-Can-t-Get-Job

Pistilli, M., & Arnold, K. (2012, April). Course signals at Purdue: Using learning analytics to increase student success. In Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, Vancouver, Canada.

Rajasingham, L. (2011). Will mobile learning bring a paradigm shift in higher education? *Education Research International*.

Raju, D., & Schumacker, R. (2016). Comparing data mining models in academic analytics. *International Journal of Knowledge-Based Organizations*, *6*(2), 38–53. doi:10.4018/IJKBO.2016040103

Ramírez-Noriega, A., Juárez-Ramírez, R., & Martínez-Ramírez, Y. (2017). Evaluation module based on Bayesian networks to Intelligent Tutoring Systems. *International Journal of Information Management*, *37*(1), 1488–1498. doi:10.1016/j.ijinfomgt.2016.05.007

Rickel, J., & Johnson, W. L. (1999). Animated agents for procedural training in virtual reality: Perception, cognition, and motor control. *Applied Artificial Intelligence*, *13*(4-5), 343–382. doi:10.1080/088395199117315

Rogers, E. (2003). *Diffusion of innovations* (5th ed.). New York: Free Press.

Rubin, R., & Dierdorff, E. (2013). Building a better MBA: From a decade of critique toward a decennium of creation. *Academy of Management Learning & Education*, *12*(1), 125–141. doi:10.5465/amle.2012.0217

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Rus, V., D'Mello, S., Hu, X., & Graesser, A. (2013). Recent advances in conversational tutoring systems. *AI Magazine*, *34*(3), 42–54. doi:10.1609/aimag.v34i3.2485

Rus, V., Niraula, N. B., & Banjade, R. (2015, January). DeepTutor: An effective, online intelligent tutoring system that promotes deep learning. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, Austin.

Sahin, I., & Thompson, A. (2006). Using Rogers' theory to interpret instructional computer use by COE faculty. *Journal of Research on Technology in Education*, *39*(1), 81–104. doi:10.1080/15391523.2006.10782474

Sargent, C. (2013). Find it, fix it, and thrive: The impact of insisting on proficiency in prerequisite knowledge in intermediate accounting. *Issues in Accounting Education*, *28*(3), 581–597. doi:10.2308/iace-50456

Sawilowsky, S. (2009). New effect size rules of thumb. *Journal of Modern Applied Statistical Methods; JMASM*, *8*(2), 467–474. doi:10.22237/jmasm/1257035100

Shahiri, A., Husain, W., & Rashid, N. A. (2015). A review on predicting student's performance using data mining techniques. *Procedia Computer Science*, *72*, 414–422. doi:10.1016/j.procs.2015.12.157

Sottilare, R. (2010, May). Challenges in the development of intelligent tutors for adaptive military training systems. In *Proceedings of the International Training and Education Conference (ITEC)*, London.

Sottilare, R. A., Holden, H. K., Brawner, K. W., & Goldberg, B. S. (2011). Challenges and emerging concepts in the development of adaptive, computer-based tutoring systems for team training. In *Proceedings of the Interservice/Industry Training, Simulation, and Education Conference*, Orlando, December.

Spada, K. (2014). Higher education in the digital age [Review]. *IEEE Transactions on Professional Communication*, *57*(2), 150–153. doi:10.1109/TPC.2014.2311873

Spain, R., Priest, H., & Murphy, J. (2012). Current trends in adaptive training with military applications: An introduction. *Military Psychology*, *24*(2), 87–95. doi:10.1080/08995605.2012.676984

Spender, J. (2016). How management education's past shapes its present. BizEd, March/April. Retrieved from http://www.bizedmagazine.com/archives/2016/2

Starke-Meyerring, D., & Wilson, M. (2008). *Designing globally networked learning environments: visionary partnerships, policies, and pedagogies*. Sense Publishers.

Steenbergen-Hu, S., & Cooper, H. (2014). A Meta-Analysis of the effectiveness of intelligent tutoring systems on college students' Academic Learning. *Journal of Educational Psychology*, *106*(2), 331–347. doi:10.1037/a0034752

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., ... & Leyton-Brown, K. (2016). Artificial intelligence and life in 2030. Stanford University. Retrieved from https://ai100.stanford.edu/2016-report

Stonebraker, I., Robertshaw, M. B., & Moss, J. D. (2016). Student see versus student do: A comparative study of two online tutorials. *TechTrends*, *60*(2), 176–182. doi:10.100711528-016-0026-7

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Suhadi, S., Mohamed, H., Abdullah, Z., Zaid, N. M., Aris, B., & Sanmugam, M. (2016). Enhancing student's higher order thinking skills (HOTS) through the Socratic Method approach with technology. *International Journal of Knowledge-Based Organizations*, *6*(4), 14–28. doi:10.4018/IJKBO.2016100102

Thomas, H., & Cornuel, E. (2012). Business schools in transition? Issues of impact, legitimacy, capabilities and re-invention. *Journal of Management Development*, *31*(4), 329–335. doi:10.1108/02621711211219095

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, *46*(4), 197–221. doi:10.1080/00461520.2011.611369

Gonzalez-Perez, M. A., Velez-Calle, A., Cathro, V., Caprar, D. V., & Taras, V. (2014). Virtual teams and international business teaching and learning: The case of the global enterprise experience. *Journal of Teaching in International Business*, *25*(3), 200–213. doi:10.1080/08975930.2014.925738

Waddock, S., & Lozano, J. (2013). Developing more holistic management education: Lessons learned from two programs. *Academy of Management Learning & Education*, *12*(2), 265–284. doi:10.5465/amle.2012.0002

Wang, D., Han, H., Zhan, Z., Xu, J., Liu, Q., & Ren, G. (2014). A problem solving oriented intelligent tutoring system to improve students' acquisition of basic computer skills. *Computers & Education*, *81*, 102–112. doi:10.1016/j.compedu.2014.10.003

Winch, J., & Cahn, E. (2015). Improving student performance in a management science course with supplemental tutorial videos. *Journal of Education for Business*, *90*(7), 402–409. doi:10.1080/08832323.2015.1081865

Xu, D., Huang, W. W., Wang, H., & Heales, J. (2014). Enhancing e-learning effectiveness using an intelligent agent-supported personalized virtual learning environment: An empirical investigation. *Information & Management*, *51*(4), 430–440. doi:10.1016/j.im.2014.02.009

Yeh, S., Hwang, W.-Y., Wang, J.-L., & Zhan, S.-Y. (2013). Study of co-located and distant collaboration with symbolic support via a haptics-enhanced virtual reality task. *Interactive Learning Environments*, *21*(2), 184–198. doi:10.1080/10494820.2012.705854

This research was previously published in the International Journal of Knowledge-Based Organizations (IJKBO), 8(3); edited by John Wang; pages 28-46; *copyright year 2018 by IGI Publishing (an imprint of IGI Global).*

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 16
A User-Aware and Semantic Approach for Enterprise Search

Giacomo Cabri
University of Modena and Reggio Emilia, Italy

Riccardo Martoglia
University of Modena and Reggio Emilia, Italy

ABSTRACT

This article describes how in addition to general purposes search engines, specialized search engines have appeared and have gained their part of the market. An enterprise search engine enables the search inside the enterprise information, mainly web pages but also other kinds of documents; the search is performed by people inside the enterprise or by customers. This article proposes an enterprise search engine called AMBIT[1]-SE that relies on two enhancements: first, it is user-aware in the sense that it takes into consideration the profile of the users that perform the query; second, it exploits semantic techniques to consider not only exact matches but also synonyms and related terms. It performs two main activities: (1) information processing to analyse the documents and build the user profile and (2) search and retrieval to search for information that matches user's query and profile. An experimental evaluation of the proposed approach is performed on different real websites, showing its benefits over other well-established approaches.

INTRODUCTION

Enterprises produce and rely on a large amount of information. A small part of information is available by public web sites, while the most part is exploited by employees of the enterprise itself by means of intranet, and by customers of the enterprise who have access to some information for business purposes.

In this scenario, the capability of searching for needed information plays a fundamental role. On the one hand, enabling internal employees to find the needed information in a short time is not only useful to speed up their work, but also to avoid or decrease the frustration of long and unsuccessful searches. On the other hand, precise and relevant answers to customers that exploit the company web sites for both searching and interacting can grant a high degree of customer satisfaction.

DOI: 10.4018/978-1-7998-0951-7.ch016

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In this scenario, the authors point out two aspects that can improve the use of search engines in an enterprise context by providing more relevant search results: user-awareness and semantics. Existing studies and surveys in general information management contexts have highlighted the benefits that can be brought to search results by the former (Xiang et al., 2010) and latter (Mangold, 2007). User-awareness means to exploit the knowledge of the user in terms of profile and context to effectively tailor the search on the base of the available information. Semantics can be useful to overcome the limitations of a syntactic approach, which is often exploited but does not consider similar pieces of information expressed in different ways. As far as the authors know, there are no enterprise search engines that exploit both aspects in a single approach.

Starting from the above considerations, the general semantic foundations introduced in (Martoglia, 2015) are exploited to proceed towards the goal of achieving a user-aware semantic enterprise search engine. The search engine is called AMBIT-SE (AMBIT Search Engine). It is not a generic search engine, but a search engine dedicated to searches in an enterprise scenario. The AMBIT-SE approach improves search results by:

- Taking advantage of textual information, including user information. Indeed, text is the primary component of the documents that should be presented / suggested to users, and also one of the main information characterizing user profiles. Consider, for instance, the contents of user browsing history, the description of users' interests, and so on;
- Exploiting semantic techniques: instead of a pure syntactic matching between the query keywords and the words in the available documents, it relies on their meaning and takes into account synonyms and related terms.

The approach presented in this paper brings the following novel contributions with regard to the initial idea of an enterprise search engine sketched in (Cabri, 2016) and to the state of the art:

- Differently from the state of the art on available enterprise search engines, it is able to combine semantics and user-awareness without requiring any manual work (e.g. for annotating documents, describing user profiles, etc.). Novel semantic and user-aware techniques allow the engine to go beyond standard syntactic search in a completely automatic way;
- The semantic text analysis techniques are adapted and refined from previous authors' studies on the effectiveness of semantic text management in specific subject areas such as software engineering (Bergamaschi et al., 2015; Martoglia, 2011), agricultural (Beneventano et al., 2016) and user-centric cultural enhancement data (Martoglia, 2015). In the presented approach, the techniques are generalized to work in a non-specialized enterprise search setting with general purpose ontologies and new weighting schemes, allowing them to be directly compared to the new class similarity contribution;
- The strength of the semantic text analysis contribution is added to the new contribution given by semantic categorization. Categorization classifies documents on the basis of a well-known taxonomy (defined by IPTC[2]) in order to provide improvements in the retrieval effectiveness. To this end, a novel class similarity metric is introduced, with a weighting scheme exploiting the novel concept of inverse (document) class frequency;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- A novel ranking selection/fusion technique is employed, producing a final document ranking which: (1) reflects both the query and user profile in a flexible and customizable proportion; (b2 fuses both class and text similarity contributions in the case they are judged as significant; (3) otherwise, it is able to automatically exclude one of the two contributions from the final result.

The combined text analysis, user-aware and semantic retrieval techniques ultimately provide enhanced searching effectiveness over standard search techniques, as also shown in the experimental tests. Moreover, the approach is devised for IT Small and Medium-sized Enterprises (SMEs), providing them with easy-to-apply methods that allow them to query for the information they need in the way they are used to.

This paper is organized as follows. First, a related work analysis is presented (Section "Related Work"). Then, Section "Search Engine Architecture" describes the architecture of the proposed search engine. Section "Information Processing" explains how the presented approach processes the documents that can be "searchable" by the users. Section "Search and Retrieval" illustrates how the system defines a ranking of retrieved documents to satisfy the user's query. Finally, the results of the experiments carried out on the system are reported, before the conclusions (Section "Conclusions"), in Section "Experimental Evaluation".

RELATED WORK

This section presents a review of the existing approaches related to user-aware semantic enterprise search engines. The authors have considered both academic and commercial approaches, and have classified them on the base of two aspects: the user-awareness and the semantics; the resulting taxonomy is reported in Figure 1.

Most of the approaches do not consider together these aspects and/or cannot be classified as enterprise search engines. For these reasons, related work will be presented in three separate subsections: semantic approaches, user-aware approaches and enterprise search engines.

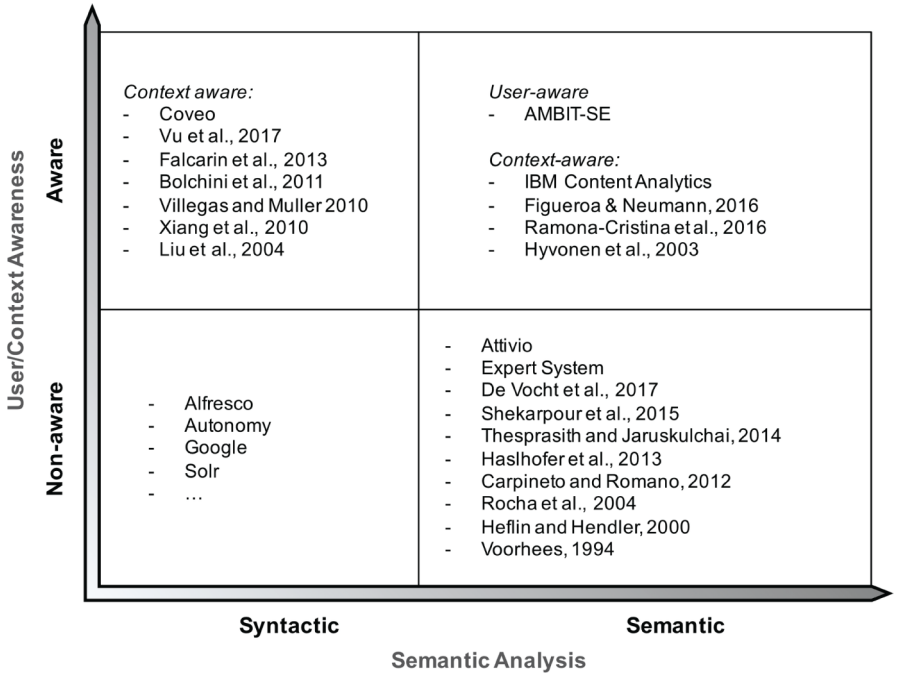
Semantic Approaches

The Semantic Web is likely to be the field where most of the approaches for semantic document retrieval has been proposed, as reported in (Mangold, 2007). It is a survey which covers approaches that exploit domain knowledge to process search requests. The authors discuss a large variety of domain knowledge utilization that include automatic query expansion and ontology-driven document retrieval.

The possible ineffectiveness of information retrieval systems is mainly due to the inaccuracy with which a query formed by a few keywords models the actual user information need. One well known method to solve this problem is automatic query expansion, whereby the user's original query is augmented by new features with a similar meaning (Carpineto & Romano, 2012). Differently from the approach presented in this paper, complex query expansion techniques, such as the ones discussed, usually require different parameters to be specified (as also stated in (Abdou & Savoy, 2008)). For instance, the method proposed in (Voorhees, 1994) involves a set of parameters specifying for each run and for each relation type included in the ontology the maximum length of a chain of that type of link that may be followed. Generally, there is no single theory capable of finding the most appropriate values (Abdou & Savoy, 2008) and therefore a long process of manual tuning is needed.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Taxonomy of user-aware semantic approaches



More and more document retrieval systems make use of ontologies to help users better specify their information needs and produce semantic representations of documents. (Haslhofer et al., 2013) proposes a Simple Knowledge Organization System (SKOS) based term expansion and scoring technique that leverages labels and semantic relationships of SKOS concept definitions. The Hybrid spreading activation approach (Rocha et al., 2004) requires tight coupling between the document base and the ontology, which is a graph where concepts and properties are nodes and edges, respectively. The set of nodes that match the given query terms is used as the start nodes of a spreading activation algorithm: documents with highest activation are ranked highest in the result set. Differently from the presented approach, the mentioned systems have no notion of user context.

Given the need for manual intervention, typical semantic retrieval techniques obtain a good degree of effectiveness only on manually annotated collections and/or with explicit user intervention. (Savoy, 2005) compare the retrieval effectiveness of different search models in a bibliographic database context. These models are founded on automatic syntactic text-word indexing or on manually assigned controlled descriptors. (Thesprasith & Jaruskulchai, 2014) proposes a query expansion technique working on MEDLINE documents that have been manually assigned to controlled MeSH (Medical Subject Headings) vocabularies. The indexing and retrieval approach proposed by AMBIT-SE, instead, exploits the semantics of the text while remaining completely automatic.

(De Vocht et al., 2017) presents a semantic search engine focusing in particular on integration of different sources of data in the science research field. To increase the precision of the results, the authors annotated and interlinked structured research data with ontologies from various repositories exploiting a semantic model. That approach does not consider user-awareness and requires annotation.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

(Figueroa & Neumann, 2016) focuses on the search in the context of Community question answering (cQA) platforms. It induces the semantic classes of question-like search queries by means of the contextual information. Context is set up or represented by inferred views of their respective search sessions, namely views modelling previous queries entered by the same user. The idea of introducing semantic classes and of combining them with contextual information is very interesting, but in that work is limited to a specific field (cQA).

SINA (Shekarpour et al., 2015) is a scalable keyword search system that can answer user queries by transforming user-supplied keywords or natural-languages queries into conjunctive SPARQL queries over a set of interlinked data sources. In this case, differently from the scenario considered in this paper, data are expressed in graph format. The system exploits semantics to improve search over different data sources, but does not take into consideration user information to refine queries.

User-Aware Approaches

The advantages of taking into consideration the context has been point out several times in the literature (Bolchini et al., 2011; Cabri et al., 2003; Falcarin et al., 2013; Xiang et al., 2010; Vu et al., 2017). In particular, a few works concern context modelling, representation, and effective handling. For instance, (Bolchini et al., 2011) proposes to design a context management system which is not application-dependent, (Falcarin et al., 2013) proposes an architectural framework for context data management, while (Villegas & Müller 2010) reports the result of a study on various context modelling and management approaches. (Xiang et al., 2010) addresses the problem of integrating context information into a ranking model. (Liu et al., 2004) proposes a method to derive a user profile based on the search history and on pre-determined category hierarchies. (Vu et al., 2017) proposes a personalised query suggestion framework for Intranet search, relying on two temporal user profiles.

Most of these approaches primarily focus on specific aspects such as external conditions or location, they do not consider the semantics of the context and/or they rely on manual work in order to classify and categorize users and documents. General purpose search engines, such as Google, typically provide only very simple localization of search results on the base of the IP-address.

Enterprise Search Engines

Many enterprise search engines have been proposed in the literature or are sold by companies. Most of the proposals do not provide an ontology-based semantic analysis, relying instead on syntactic and hand-coded rules. Some examples are Alfresco[3], Autonomy[4], Solr[5]. Google[6] is also an example of a syntactic search engine that can be exploited in a enterprise search context.

There are, of course, exceptions to this rule. The SHOE project (Heflin & Hendler, 2000), which requires a domain-ontology where document types correspond to ontology concepts. Attivio[7], a product which manages information in the RDF format, providing search results and alerts. Expert System's Cogito[8], which provides automated disambiguation, classification, entity extraction, and metadata. Nevertheless, these systems provide no notion of user context. Instead, Coveo[9] is a tool specifically oriented to exploit contextual knowledge for dealing with information related to customers and agents, but it does not exploit semantic information.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Considering the semantic and context information, there are few systems that exploit it, even if in a sometimes-limited way. The Ontogator system (Hyvonen et al., 2003), part of an image management and retrieval system, provides an interactive recommendation system that allows the user to browse images based on ontological properties. To exploit user contexts, it introduces views to the ontology that rely on different concept hierarchies, called "facets". Each view represents a specific information-need. PrEmISES, proposed in (Ramona-Cristina et al., 2016), is a framework that aims at addressing information management needs of SMEs relying on ontologies to add semantically enabled information integration. The framework definition is still in a very preliminary phase; therefore, semantic and context-sensitive features are currently only sketched. Another example is IBM's Content Analytics with Enterprise Search[10], which exploits a framework called Unstructured Information Management Architecture (UIMA), in order to build analytic applications and to find meanings, relationships and relevant facts hidden in unstructured text. Context information is provided by means of manual annotations. These approaches require manual intervention on the documents and/or adopt a still limited notion of context, i.e. they do not exploit all of the data potentially available related to the user, such as the contents of any web page visited, attachment downloaded, and similar documents.

DISCUSSION

The authors have reported several researches that are connected to the approach presented in this paper. They point out that there is no existing approach that addresses all the following aspects at the same time:

- **Semantics:** Exploitation of semantic techniques to improve the results of the query;
- **User-awareness:** Exploitation of user information to customize the results of the queries;
- **Generality:** Capability of being applied to general sources of information, not to a specific field;
- **Automation:** No need for manual annotation of the information.

Starting from the above analysis, the AMBIT-SE approach aims at addressing all the above-mentioned aspects. In the next section, the architecture of AMBIT-SE is presented.

SEARCH ENGINE ARCHITECTURE

Figure 2 shows the architecture of the proposed search engine, the main activities and the related modules. In (Cabri et al., 2016) the authors tackled in detail many text pre-processing issues such crawling techniques, paragraph identification and text tagging. In this paper, the focus is on defining the semantic text analysis and the novel semantic categorization techniques (described in Section "Information Processing") The general architecture is also significantly extended in order to manage the output of both kinds of techniques and to exploit it in the actual search (Section "Search and Retrieval"). The architecture highlights two main activities:

1. **Information processing (dashed line in figure):** This activity is applied to both the documents to be retrieved (e.g. web pages for a given site) and the documents useful to determine the user's profile (such as e-mails, web pages viewed, profile information, past search queries, etc.). The avail-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

able information is extracted (crawled) and indexed by means of ad-hoc techniques, also exploiting external knowledge sources. The data structures containing all the information processing results will be referred to as "semantic glossaries", one for the website(s) and one for the user profile. Both glossaries are then compared ("Semantic glossaries computation and comparison" in the figure) with ad-hoc document similarity algorithms, detailed in Section "Search and Retrieval". The generated "Profile rankings" symbolize how relevant the retrievable documents are in relation to the user's profile;

2. **Search and retrieval (solid line in figure):** This activity provides useful answers to the user by retrieving the most relevant documents with regard to the user query, also taking into account its pre-computed profile information. This is achieved ("Semantic query processing") by determining the "Query rankings" of the query with regard to the available documents (semantic similarity discussed in Section "Search and Retrieval"). Query rankings are finally fused with profile rankings in order to provide the final ranking.

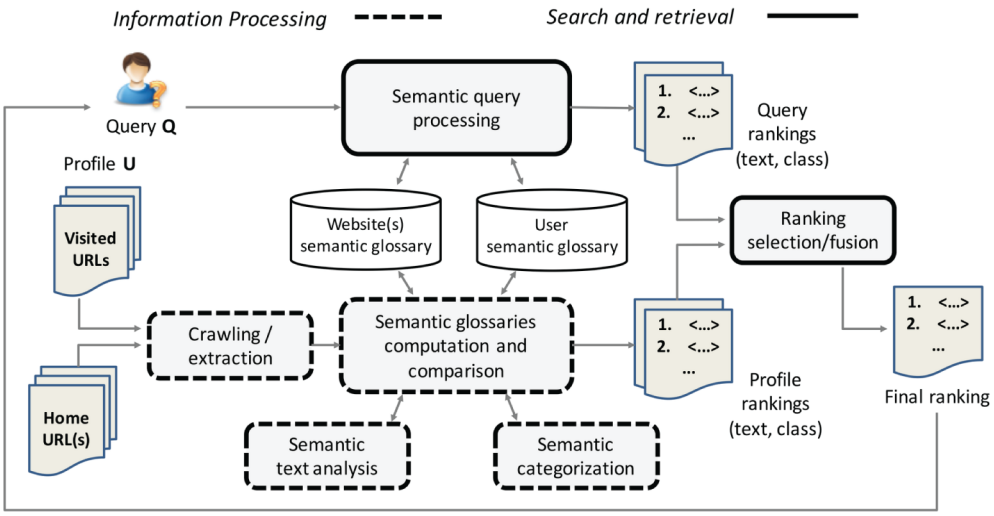
These two main activities are detailed in the next two sections.

INFORMATION PROCESSING

Information processing is common to both the documents of the website(s) to be indexed and the user profile documents. Each of them will contribute to a semantic glossary structure containing the results of the analysis.

The first step of the information processing is crawling: web/file crawling is performed in order to retrieve the raw data (e.g., Title, Content, URL, File Name, Meta Description, Meta Keywords) of the documents that will be further analysed by the subsequent steps. All of the crawling operations are handled by the open-source enterprise-class search engine software, OpenSearchServer[11].

Figure 2. An overview of the AMBIT-SE architecture



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Semantic Text Analysis

As single existing tools do not allow sufficient configuration and extension options, the authors designed a customized Semantic text analysis module, which exploits several open source libraries to perform specific actions. The analyser allows us to extract the contents (and meanings) of the processed information, by means of:

- Text extraction, an operation that can vary greatly depending on the format of each document; this is taken care of by components of the open-source software GATE[12];
- Tokenization, the terms are identified and punctuation is removed;
- Stemming and Part of Speech (POS) Tagging, the tokens are "normalized" and "stemmed", i.e., they are reduced to their base form (managing plurals and inflections) and "tagged" with POS tags (i.e., nouns, verbs, etc.); this is taken care of by the TreeTagger[13] library;
- Composite term identification, possible composite terms (such as "production area" or "wine tasting") are identified by means of a simple state machine and of POS tags information;
- Word Sense Disambiguation (WSD), the meaning of the keywords resulting from the previous steps is made explicit with regard to a reference thesaurus (typically, WordNet[14], (Miller, 1995)); in particular, the relevant synsets are extracted and associated;
- Weight computation, keywords are finally enriched with weight information (see below).

The extracted information enables the retrieval of the most relevant information for the user in the search and retrieval phase. In particular, the proposed approach exploites the semantic information from the thesaurus and the similarity functions described in Section "Search and Retrieval". Thanks to them, AMBIT-SE will be able to retrieve documents on the basis of synonyms and related keywords; for instance, documents about "dogs" are considered as relevant to a query about a "terrier". Moreover, by means of WSD, documents about a "pet" in the sense of "domesticated animal" will not be mixed up with those where "pet" means "a special loved one".

For the weight information, text analysis exploits a variant of the standard "tf-idf" weighting scheme. This is introduced to convey the information of the classic scheme while also keeping the weights normalized in the range of [0,1]: this enables an effective ranking comparison and fusion (more on this in Section "Search and Retrieval"). Given a document D^x, each keyword $k_i^x \in D^x$ is assigned a *keyword weight* kw_i^x defined as:

$$kw_i^x = \overline{kf}_i^x \cdot \overline{idf}_i \tag{1}$$

where:

$$\overline{kf}_i^x = \frac{f_i^x}{\max_l f_l^x} \tag{2}$$

$$\overline{idf}_i = log\left(\frac{N}{n_i}\right) / max_l\left(log\left(\frac{N}{n_l}\right)\right) \tag{3}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$\overline{kf}$ and $\overline{idf}$ are normalized keyword frequency and inverse document frequency: f_i^x is the raw frequency of keyword k_i in document D^x, N is the total number of indexed documents and n_i the number of documents where keyword k_i appears. Note that, by definition, $0 \leq \overline{kf}_i^x \leq 1$ and $0 \leq \overline{idf}_i \leq 1$.

Semantic Categorization

Besides semantic text analysis, a semantic categorization process is also applied to the documents in order to tag each document with appropriate subject classes. The Media Topic NewsCodes taxonomies and vocabularies provided by IPTC are adopted. These are well-known taxonomies offering a very good level of detail and coverage of a wide range of topics. In this case, the starting point is the output of ad-hoc categorization tools from Expert System[15], where each class tag is given a score $s(c_i)$, $0 \leq s(c_i) \leq 1$; the higher the weight the more relevant the class is for the document. For instance, a document about the typical "Terrier food products" will presumably have "Pet product and service" among its highest scoring associated tags. However, the authors deem that if most of the documents of the collection are tagged with the same class, this class will not be particularly distinctive and useful in the retrieval phase. Therefore, this fact is captured by going beyond the plain score and by defining a new weighting scheme for the classes inspired by the text analysis scheme (Equation 1): given a document D^x, each class tag $c_i^x \in D^x$ is assigned a class weight cw_i^x defined as:

$$cw_i^x = s\left(c_i^x\right) \cdot \overline{icf}_i \tag{4}$$

where:

$$\overline{icf}_i = log\left(\frac{N}{t_i}\right) \Big/ max_l \left(\log\left(\frac{N}{t_l}\right)\right) \tag{5}$$

$\overline{icf}$ stands for *inverse (document) class frequency*; N is the total number of documents and t_i is the number of documents tagged with class c_i.

Website(s) and User Semantic Glossaries

By applying batch information processing to the document collection, the *website(s) semantic glossary* is automatically generated. Conceptually, it consists of a *global view* (all keywords/classes together with their occurrences and additional extracted data), and a *per-document view* (keywords/classes occurrences in each document with their statistics). A simplified sample of the latter view is shown in Table 1-left and Table 1-right). In particular, "Document" is the list of the documents IDs in which each keyword/class occurs, "Synsets(s)" are the synsets selected as the output of WSD, while "KF", "KWeight" and "CWeight" are the weights illustrated in Equations (1) and (2).

Each time a user logs in AMBIT-SE, a batch analysis is also scheduled on the data of her profile U in order to generate/update its *user semantic glossary* (which shares the same structure of the website semantic glossary discussed above). In particular, the profile contains *action history data* including a list

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Sample portions of the extracted Document Semantic Glossary: per-doc view for keywords (left) and classes (right)

Document	Keyword	Synset(s)	KF	KWeight		Document	Class	CWeight
P02001	Terrier	00919240-n	0.445	0.277		P02001	Pet product and service	0.645
P02005	Veal	00414222-n	0.210	0.131		P02005	Food industry	0.442
…	…	…	…	…		…	…	…

of past accessed documents and past searches performed on the website. The idea is that such data can be analysed in the same way as the website documents, therefore exploiting all the power of the document text analysis and categorization in order to associate meaningful keywords and classes to users; these will provide more for relevant results to their queries in the search phase. Due to their complexity and so the need for computational power, all the analyses are performed when no users are logged in; they will be available in order to process future requests from the same user in a more accurate way.

SEARCH AND RETRIEVAL

When a user U submits a query Q, AMBIT-SE has to answer it as effectively as it can. The technique illustrated in this section aims to produce the most effective ranking $\hat{\tau}$ of the indexed documents $D \in D$ with regard to both U and Q. This is done by taking into account all the information available in the semantic glossaries produced by the analysis process (text analysis and classification) and by means of the following ad-hoc similarity metrics between two generic documents (i.e. keyword sets) D^x and D^y:

- The similarity *TextSim*(D^x,D^y), considering *keyword* information;
- The similarity *ClassSim*(D^x,D^y), considering *class* information.

In particular, *TextSim* and *ClassSim* will be applied to both:

- The user profile U (i.e.,) with regard to each indexed document (as performed on past navigated data): this produces the profile rankings τ^U_{text} and τ^U_{class};
- The query Q (i.e.,) with regard to each document; this produces the query rankings τ^Q_{text} and τ^Q_{class}.

In the final part of this section, the authors will show how this ranking information is exploited in *ranking selection/fusion* in order to produce the final ranking $\hat{\tau}$. Next, the *TextSim* and *ClassSim* similarity metrics will be defined.

Text and Class Similarity Metrics

Building on previous research on text retrieval for specific subject areas as software engineering (Bergamaschi et al., 2015; Martoglia, 2011), agricultural (Beneventano et al., 2016) and user-centric cultural enhancement data (Martoglia, 2015), the text similarity *TextSim*(D^x,D^y) formula between documents D^x and D^y is defined as:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$TextSim\left(D^x, D^y\right) = \frac{\sum_{k_i^x \in D^x} max_{k_j^y \in D^y} \left(KSim\left(k_i^x, k_j^y\right)\right) \cdot kw_i^x \cdot kw_j^y}{\left|D^x\right|} \quad (6)$$

where the weighting scheme illustrated in Equation (1) is exploited and *KSim* is a term similarity formula (see later) taking into account the semantic information extracted from the semantic glossary. Simply put, the similarity $TextSim(D^x,D^y)$ between two documents D^x and D^y is determined by summing the maximum keyword similarity scores $KSim(k_i,k_j)$ between each pair of keywords k_i and k_j belonging to different documents, multiplied by the weights of both. As to $KSim(k_i,k_j)$, in AMBIT-SE semantic framework, k_i and k_j can be:

1. *Equal or synonyms*: $KSim(,)= 1$;
2. *Related*, i.e. the thesaurus hypernymy path similarity between the keywords'synsets exceeds a given threshold *Th* (Bergamaschi et al., 2015): $KSim(,)= r$, where r is an arbitrary value between 0 and 1, default 0.7;
3. *Unrelated* otherwise: $KSim(,)= 0$.

Please note that, since both $0 \leq KSim(,) \leq 1$ and $0 \leq kw \leq 1$, then $0 \leq TextSim(,) \leq 1$. This will be useful in the ranking selection/fusion phase (Section "Ranking Selection / Fusion").

In addition to document keywords, the classes associated by the semantic classifier can also significantly help in retrieving useful documents. This is obviously true if both documents are strongly characterized by a common IPTC class (e.g. "process industry"); however, also documents about cola factories tagged with a similar class "food industry" would be of interest. This is achieved through $ClassSim(D^x,D^y)$, which quantifies the similarity of D^x and D^y on the basis of their associated IPTC classes $c_i^x \in D^x$ and $c_j^y \in D^y$:

$$ClassSim\left(D^x, D^y\right) = \frac{\sum_{c_i^x \in D^x} max_{c_j^y \in D^y} \left(CSim\left(c_i^x, c_j^y\right)\right) \cdot cw_i^x \cdot cw_j^y}{\left|\left\{c^x \in D^x\right\}\right|} \quad (7)$$

As for Equation (6), $CSim(c_i,c_j)$ between classes c_i and c_j ranges between 1 (equal classes), r (similar, i.e. "near" on the IPTC taxonomy, classes) and 0 (otherwise). Again, since $0 \leq CSim(,) \leq 1$, then $0 \leq ClassSim(,) \leq 1$.

Ranking Selection/Fusion

As previously seen, given a profile U, Equations (3) and (4) induce rankings τ_{text}^U and τ_{class}^U on each document $D \in D$ with regard to U, respectively. Similarly, when the two equations are computed with regard to a query Q, rankings τ_{text}^Q and τ_{class}^Q are induced on the same documents. The aim of the *ranking selection/fusion* process presented in this paper is to exploit the information of those rankings in order to obtain the final ranking $\hat{\tau}$. The aims are the following:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. The final ranking $\hat{\tau}$ should reflect both Q and U, and it should be possible to define the relevant importance of the query Q with regard to the profile U (for instance, to privilege the information contained in Q);
2. The final ranking $\hat{\tau}$ should be flexibly defined so to reflect both text and class information (Equations. 3 and 4, respectively), possibly in a customizable proportion. Moreover, the system should be able to automatically decide if one of the two kinds of rankings (text, class) is not significant, for instance due to very low similarities, which would only negatively affect the final ranking, therefore excluding it from the fusion;
3. The *score* of each document D, and not only its position in the rankings, should be taken into account in order to directly reflect its relevance in the final ranking.

Let us see how AMBIT-SE achieves these points. By means of a linear combination score fusion method, the scores inducing the fused rankings τ_{text} and τ_{class}, which take into account both U and Q, are defined as:

$$s^{\tau_{text}}\left(D\right) = \alpha_Q \cdot s^{\tau_{text}^{Q}}\left(D\right) + \left(1-\alpha_Q\right) \cdot s^{\tau_{text}^{U}}\left(D\right) \tag{8}$$

$$s^{\tau_{class}}\left(D\right) = \alpha_Q \cdot s^{\tau_{class}^{Q}}\left(D\right) + \left(1-\alpha_Q\right) \cdot s^{\tau_{class}^{U}}\left(D\right)$$

where $0 \leq \alpha_Q \leq 1$ is a preference weight determining the relevant importance of Q with regard to U. The default value is $\alpha_Q = 0.7$, meaning that the information in Q is typically more significant than that in U; this can be varied, for instance, by the system administrator depending on the actual usage scenarios.

The above defined τ_{text} and τ_{class} rankings are eventually fused in a final ranking $\hat{\tau}$, which takes into account both text and class contributions as of Equation (8):

$$s^{\hat{\tau}}\left(D\right) = \beta_{text} \cdot s^{\tau_{text}}\left(D\right) + \left(1-\beta_{text}\right) \cdot s^{\tau_{class}}\left(D\right) \tag{9}$$

where $0 \leq \beta_{text} \leq 1$ is a preference weight determining the relative importance of τ_{text} with regard to τ_{class}. The default value is 0.5. Note that ranking selection is implemented in the following way: if $\sum_D s^{\tau_{text}}\left(D\right) \gg \sum_D s^{\tau_{class}}\left(D\right)$ then β_{text} is automatically set to 1 in order to exclude the contribution of τ_{class} (and vice-versa), therefore avoiding possibly detrimental noise. Only documents that are part of all rankings will appear in the final ranking.

EXPERIMENTAL EVALUATION

This section presents the results of several tests performed on different kinds of websites. This paper is focused on effectiveness evaluation; for readers interested in time performances, the current prototype has a response time of 40 ms on average on a standard single-node configuration.

Five websites were selected for evaluation purposes; for each one of them, appropriate information needs were established by examining common searches performed in the past.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The choice of the website was driven by the aims of the proposed approach. First, the sites are all representative of real and typical business-relevant sites. Moreover, the sites were selected so as to cover a number of different topics (e.g. health, cars, recycling, etc.) and thus different terminologies and text content. Finally, they have different features in terms of structure. This is in line with the main goal of this section, i.e. to evaluate the effectiveness and the flexibility of an enterprise search engine.

As to possible experimental comparisons, as anticipated in the related, there are currently no approaches that can be directly compared with the proposed one. In particular, among the approaches whose implementation was publicly available, those offering semantic features require unavailable manual annotations and/or were strictly designed to work on specific ontologies (Haslhofer et al., 2013; Thesprasith & Jaruskulchai, 2014; De Vocht et al., 2017). This is also true for semantic enterprise search engines (Cogito, Attivio, Content Analytics). The considered user-aware search engines do not exploit the semantics (Bolchini et al., 2011; Cabri et al., 2003; Falcarin et al., 2013; Xiang et al., 2010; Vu et al., 2017). Also, the publicly available approaches working on context information are restricted to aspects such as time and location (Falcarin et al., 2013; Xiang et al., 2010; Villegas & Müller 2010) or very specific scenarios (e.g. agent management for Coveo, image search for Ontogator, cQA for (Figueroa & Neumann, 2016)) which makes them not applicable to the considered case. Summarizing, the reason that prevents direct comparisons are: first, only a few number of approaches exploit both semantics and user awareness; second, most of the approaches focus on a specific application field and cannot be applied to general websites; third, several approaches are not freely available to be tested. Anyway, for reference, the final part of this section is devoted to a direct comparison with Google search engine; even if it does not exhibit all the features of AMBIT-SE, the authors considered it an important benchmark in the field.

Table 2 reports, for each website, the address, a brief description of its contents and an example of a considered information need. In the context of the information needs of each website, 50 plausible queries were submitted to the system. Moreover, two options for user profiles are considered: a simpler setting, where the profile contains only documents relevant to the information need ("homogeneous profile"), and a more complex one, where the profile contains also an equal number of irrelevant documents ("heterogeneous profile").

Table 2. Experimental setting: details of the five business-relevant websites selected for evaluation purposes

Website	Description	Examples of Information Needs
http://www.cobat.it/	A relatively small website that provides information and services for disposing and recycling four problematic waste categories: batteries and accumulators, tires, electric and electronic devices, and photovoltaic panels.	Retrieve documents pertaining to the disposal of batteries and accumulators.
http://evergreensmallbusiness.com/	A Blog that publishes different kinds of information and advice for small businesses, all classified in categories such as business taxes, management, personal finance, etc.	Retrieve articles pertaining to bookkeeping.
http://www.bagnolottanta.it/	A journalistic website with several thematic columns such as history, culture, theatre, etc.	Retrieve articles in specific columns such as books.
http://truegoods.com/	An Indie online shop that specializes on healthy and natural products.	Retrieve information on products belonging to the pet-care category.
http://www.gruppozatti.it/	An authorized car dealer which sells several brands of both new and used cars.	Retrieve different information about cars belonging to the used category.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

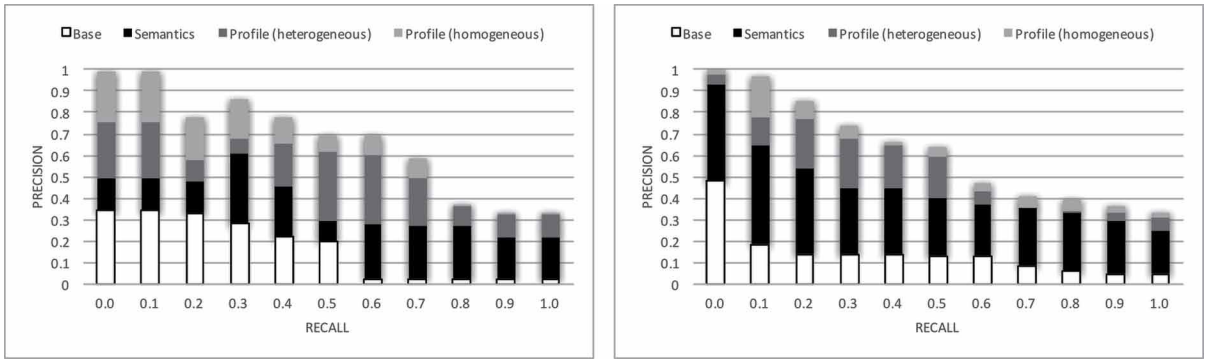
In each situation, the output of AMBIT-SE is compared with a "gold standard", i.e. relevant answers manually selected from the websites, and precision and recall are assessed; in particular, the tests compute interpolated precision at 11 recall levels of 0.0, 0.1, 0.2, ..., 1.0 (0, 10, 20, ..., 100 percent) (Baeza-Yates & Ribeiro-Neto, 1999), thus taking into account not only the relevance of the results but also their position in the returned ranking. Moreover, a "stable" situation is assumed, where users and documents have been already automatically processed and their relevant keywords and classes stored in the Semantic Glossaries. All the parameters are set at default.

Figures 3 through 5 depict the results (averaged on all the queries) obtained by AMBIT-SE, compared with the baseline of a syntactic retrieval method ignoring synonyms, related terms and class information. In particular, this baseline is representative of the document retrieval techniques commonly exploited by most commercial systems and standard enterprise search engines (e.g. Alfresco, Autonomy, Solr). Further comparisons with available systems (e.g. Google) will be discussed in the final part of the experimental analysis. In order to better visualize the difference between AMBIT-SE's results and the baseline, the figures detail the the precision achieved by the different features of AMBIT-SE. In particular, the white bars (at the bottom) represent the precision of the syntactic method (baseline); the contributions on top are the improvements achieved by means of the semantic (black) and user-aware features (gray).

As readers can see, the improvements in precision offered by the semantic and user-aware features of AMBIT-SE are significant in each scenario, ranging from 20% to even 80%. Let us start by analysing the http://www.cobat.it/ results (Figure 3 (left)). Here, the semantic features offer an advantage at all recall levels; the results also benefited from the user-aware features, because of the large number of keywords contained within the documents associated to the user profiles. For instance, the use of semantics enables the matching of different but very related terms like "battery" and "accumulator". Those contributions go unnoticed in standard syntactic search.

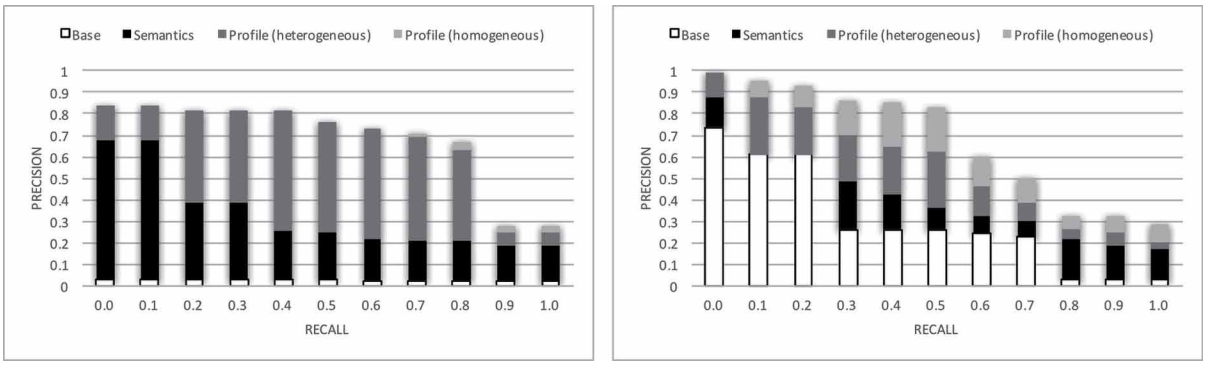
Considering the second website (Figure 3 (right)), the use of semantics provides even greater benefits: for instance, the use of synonyms and related terms are able to exploit a number of terms correlations such as between "money" and "bookkeeping." This was especially evident in the longer and less direct queries submitted, which are harder to satisfy by a search engine without additional information in the form of synonyms and related terms. Summing up the contribution of the profile analysis, AMBIT-SE reaches in some cases an improvement of nearly 80% in precision.

Figure 3. Test results for http://www.cobat.it/ (left) and http://evergreensmallbusiness.com/ (right)



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

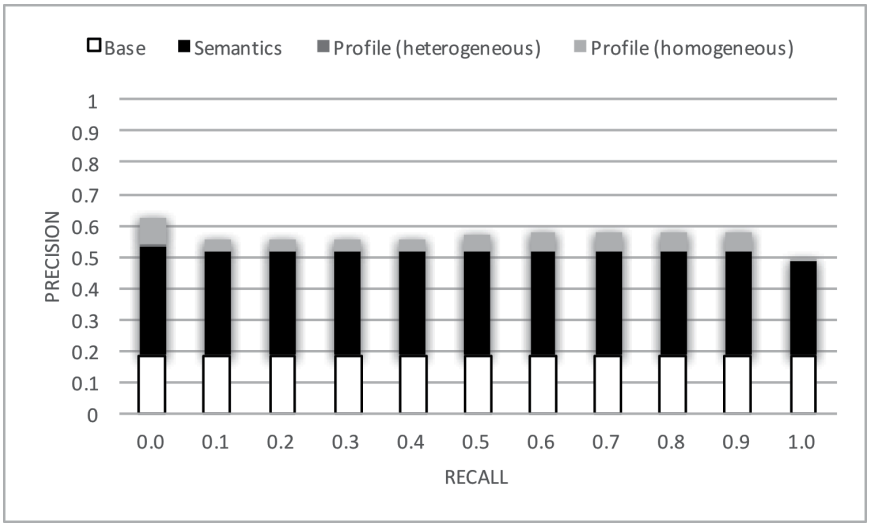
Figure 4. Test results for http://www.bagnolottanta.it/ (left) and http://truegoods.com/ (right)



In the test of Figure 4 (left), the authors found that in some cases the queries did not directly benefit from synonyms and related terms management, maybe due to very specific terminology that is used in the pages. Anyway, the use of semantics and the profile analysis provide relevant advantages even in this scenario. Moving to Figure 4 (right), the use of word stems, synonyms, related terms and profile can turn even difficult queries into manageable ones. For instance, in some cases, the syntactic baseline could not retrieve any records since the queries don't include terms that match exactly the ones found in the relevant documents: among the exploited terms correlations, the very frequent one between "pet" and "animal".

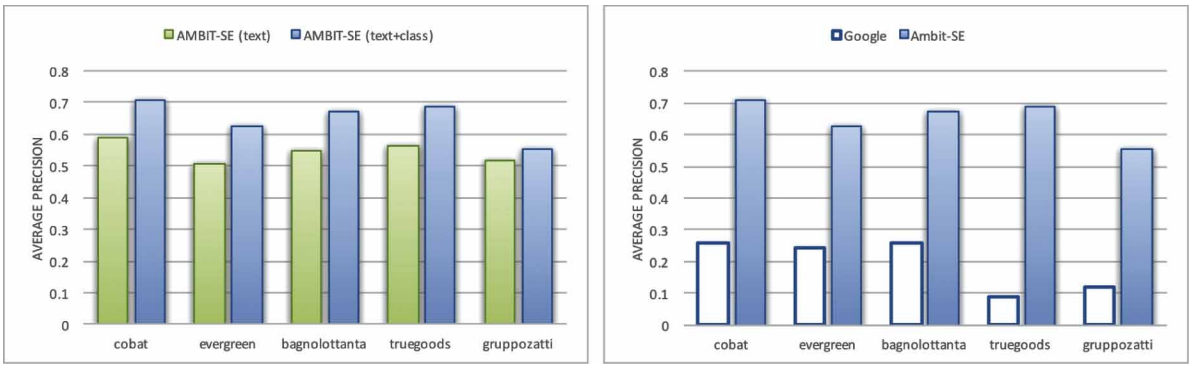
The final website considered (Figure 5) is characterized by pages containing very little text, thus providing a different task with regard to the others. Also in this case, a lot of complex queries did not yield satisfactory results for the syntactic baseline, thus providing very consistent advantages from the semantic features (precision improvement of 30% or greater). Due to the peculiar nature of the website, some queries were apparently too difficult even with the additional input provided by the semantics and profiles. Anyway, on average, the effect of the advanced AMBIT-SE features is evident even in this situation.

Figure 5. Test results for http://www.gruppozatti.it/



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Class improvement analysis (left) and comparison with Google (right)



A further test is presented which is aimed at evaluating the impact of the classification features of the presented engine, including the novel ranking selection and fusion capabilities. Figure 6 (left) shows the precision figures (averaged over all the queries and all recall levels) that are achieved by exploiting only the text ranking and similarities (left bars) with regard to the ones achieved by the complete setup presented in this paper. The benefits are usually in the order of more than 10% of improvement. Looking at specific cases, the fused ranking was able to take the best from the class and text rankings, together capturing the user interests more completely.

In order to complete the evaluation, the authors performed a quantitative comparison with a state-of-the-art search engine and a discussion about results is reported in the following. The chosen search engine was Google, as mentioned before.

Figure 6 (right) shows the results of the system compared with those obtainable through the Google search engine. Indeed, Google is one of the most widely used search engines also in the enterprise search area. Similarly to AMBIT-SE, Google is completely automatic and requires no manual work (e.g. annotation) on the documents to be processed. In this case, the considered document set has been restricted to the specific considered website pages. Differently from AMBIT-SE, Google does not perform semantic analysis of the texts, however it employs text processing techniques such as stemming. Multiple users were simulated in Google through different search sessions. The results show that the average precision achieved by Google (in the range of 10-30%) is quite lower than the one achieved by AMBIT-SE (always above 55%). In particular, Google offers a level of performance that is only marginally better than the syntactic baseline discussed in the previous tests.

CONCLUSION

Search engines represent a means essential for a lot of activities. This is especially true in an enterprise context, where the success of the enterprise is often strictly dependent on the ability of its employees and customers to find the needed information.

In this context the authors have proposed AMBIT-SE, an enterprise search engine approach that relies on two aspects, user-awareness and semantics, jointly exploited. First, it builds a profile of the user, which is exploited to search for the information that best matches with her needs. Second, semantic techniques enable the retrieval of interesting information that could not be considered exploiting a

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

standard syntactic search. The experimental evaluation has shown that the presented approach performs better than traditional search methods.

In the future, further ways of exploiting semantics and user information in the search will be considered, for instance by adapting some of the ideas coming from past works in different contexts (e.g., semantic search in heterogeneous and dynamic graph data (Catania et al., 2013) and multimedia data (Grana et al., 2013)).

ACKNOWLEDGMENT

This work was supported by the project "Algorithms and Models for Building context-dependent Information delivery Tools" (AMBIT) co-funded by Fondazione Cassa di Risparmio di Modena (SIME 2013.0660), by the "SocialGQ" FAR 2016 FIM Department Project, University of Modena and Reggio Emilia, and by the EU H2020 FIRST project (GA number: 734599 — H2020-MSCA-RISE-2016).

REFERENCES

Abdou, S., & Savoy, J. (2008). Searching in medline: Query expansion and manual indexing evaluation. *Information Processing & Management*, *44*(2), 781–789. doi:10.1016/j.ipm.2007.03.013

Baeza-Yates, R. A., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.

Beneventano, D., Bergamaschi, S., & Martoglia, R. (2016). Exploiting semantics for searching agricultural bibliographic data. *Journal of Information Science*, *42*(6), 748–762. doi:10.1177/0165551515606579

Bergamaschi, S., Martoglia, R., & Sorrentino, S. (2015). Exploiting semantics for filtering and searching knowledge in a software development context. *Knowledge and Information Systems*, *45*(2), 295–318. doi:10.100710115-014-0796-1

Bolchini, C., Orsi, G., Quintarelli, E., Schreiber, F. A., & Tanca, L. (2011). Context modeling and context awareness: Steps forward in the context-addict project. *A Quarterly Bulletin of the Computer Society of the IEEE Technical Committee on Data Engineering*, *34*, 47–54.

Cabri, G., Gaddi, S., & Martoglia, R. (2016). AMBIT-SE: Towards a User-aware Semantic Enterprise Search Engine. In *Proceedings of the 12th International Conference on Web Information Systems and Technologies (WEBIST 2016)* (Vol. 2, pp. 98–108). Springer. 10.5220/0005788800980108

Cabri, G., Leonardi, L., Mamei, M., & Zambonelli, F. (2003). Location-dependent Services for Mobile Users. *IEEE Transactions on Systems, Man, and Cybernetics. Part A, Systems and Humans*, *33*(6), 667–681. doi:10.1109/TSMCA.2003.819496

Carpineto, C. and Romano, G. (2012). A survey of automatic query expansion in information retrieval. *ACM Comput. Surv.*, 44(1):1:1–1:50.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Catania, B., Guerrini, G., Belussi, A., Mandreoli, F., Martoglia, R., & Penzo, W. (2013). Wearable Queries: Adapting Common Retrieval Needs to Data and Users (Vision Paper). In *Proceedings of the 7th International Workshop on Ranking in Databases (DBRank)*. 10.1145/2524828.2524835

De Vocht, L., Softic, S., Verborgh, R., Mannens, E., & Ebner, M. (2017). Social Semantic Search: A Case Study on Web 2.0 for Science. *International Journal on Semantic Web and Information Systems*, *13*(4), 155–180. doi:10.4018/IJSWIS.2017100108

Falcarin, P., Valla, M., Yu, J., Licciardi, C. A., Frà, C., & Lamorte, L. (2013). Context data management: An architectural framework for context-aware services. *Service Oriented Computing and Applications*, *7*(2), 151–168. doi:10.100711761-012-0115-1

Figueroa, A., & Neumann, G. (2016). Context-aware semantic classification of search queries for browsing community question–answering archives. *Knowledge-Based Systems*, *96*, 1–13. doi:10.1016/j.knosys.2016.01.008

Grana, C., Serra, G., Manfredi, M., Cucchiara, R., Martoglia, R., & Mandreoli, F. (2013). UNIMORE at ImageCLEF 2013: Scalable Concept Image Annotation. In *Proceedings of the Image Retrieval in Conference and Labs of the Evaluation Forum (ImageClef)*.

Haslhofer, B., Martins, F., & Magalhães, J. a. (2013). Using skos vocabularies for improving web search. In *Proceedings of the 22nd International Conference on World Wide Web* (pp. 1253-1258).

Heflin, J., & Hendler, J. (2000). Searching the web with shoe. In *Artificial Intelligence for Web Search. Papers from the AAAI Workshop*.

Hyvonen, E., Saarela, S., & Viljanen, K. (2003). Ontogator: combining view- and ontology-based search with semantic browsing. In *Proceedings of XML Finland*.

Liu, F., Yu, C., & Meng, W. (2004). Personalized web search for improving retrieval effectiveness. *IEEE Transactions on Knowledge and Data Engineering*, *16*(1), 28–40. doi:10.1109/TKDE.2004.1264820

Mangold, C. (2007). A survey and classification of semantic search approaches. In Semantics and Ontology. doi:10.1504/IJMSO.2007.015073

Martoglia, R. (2011). Facilitate IT-Providing SMEs in Software Development: a Semantic Helper for Filtering and Searching Knowledge. In *Proceedings of the 23rd International Conference on Software Engineering and Knowledge Engineering* (pp. 130-136).

Martoglia, R. (2015). Ambit: Semantic engine foundations for knowledge management in context-dependent applications. In *Proceedings of the 27th International Conference on Software Engineering and Knowledge Engineering* (pp. 146–151). 10.18293/SEKE2015-027

Miller, G. A. (1995). WordNet: A Lexical Database for English. *Communications of the ACM*, *38*(11), 39–41. doi:10.1145/219717.219748

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ramona-Cristina, P., Vasilateanu, A., & Goga, N. (2016). Ontology based multi-system for SME knowledge workers. In *Proceedings of the 2016 IEEE International Symposium on Systems Engineering (ISSE)*, Edinburgh, UK, October 3-5. 10.1109/SysEng.2016.7753132

Rocha, C., Schwabe, D., & de Aragao, M. P. (2004). A hybrid approach for searching in the semantic web. In *WWW '04: Proceedings of the Thirteenth International Conference on World Wide Web*. 10.1145/988672.988723

Savoy, J. (2005). Bibliographic database access using free-text and controlled vocabulary: An evaluation. *Information Processing & Management*, *41*(4), 873–890. doi:10.1016/j.ipm.2004.01.004

Shekarpour, S., Marx, E., Ngonga Ngomo, A.-C., & Aue, S. (2015). SINA: Semantic interpretation of user queries for question answering on interlinked data. *Journal of Web Semantics*, *30*, 39–51. doi:10.1016/j.websem.2014.06.002

Thesprasith, O., & Jaruskulchai, C. (2014). Query expansion using medical subject headings terms in the biomedical documents. In *Intelligent Information and Database Systems - 6th Asian Conference, ACIIDS 2014*, Bangkok, Thailand, April 7-9 (pp. 93-102).

Villegas, N. M., & Müller, H. A. (2010). Managing dynamic context to optimize smart interactions and services. In M. Chignell, J. Cordy, J. Ng et al. (Eds.), The Smart Internet, LNCS (Vol. 6400, pp. 289-318). Springer Berlin Heidelberg. doi:10.1007/978-3-642-16599-3_18

Voorhees, E. M. (1994). Query expansion using lexical-semantic relations. In *Proceedings of the 17th Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval. Dublin, Ireland, 3-6 July* (pp. 61-69).

Vu, T., Willis, A., Kruschwitz, U., & Song, D. (2017). Personalised Query Suggestion for Intranet Search with Temporal User Profiling. In *Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval* (pp. 265-268). 10.1145/3020165.3022129

Xiang, B., Jiang, D., Pei, J., Sun, X., Chen, E., & Li, H. (2010). Context-aware ranking in web search. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval SIGIR '10* (pp. 451-458).

ENDNOTES

1 AMBIT stands for "Algorithms and Models for Building context-dependent Information delivery Tools"
2 IPTC stands for "International Press Telecommunications Council", http://www.iptc.org/site/Home/
3 http://www.alfresco.com/
4 http://www.autonomy.com/
5 http://lucene.apache.org/solr/
6 http://www.google.com/
7 http://www.attivio.com/
8 http://www.expertsystem.com/it/cogito/

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

[9] http://www.coveo.com/
[10] https://www.ibm.com/
[11] http://www.opensearchserver.com/
[12] https://gate.ac.uk/
[13] http://www.cis.uni-muenchen.de/.17ex~schmid/tools/TreeTagger/
[14] http://wordnet.princeton.edu/ or, in case of specific contexts, other specialized resources
[15] http://www.expertsystem.com/

This research was previously published in the International Journal on Semantic Web and Information Systems (IJSWIS), 14(4); edited by Miltiadis D. Lytras; pages 129-146; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 17
Use of Chaotic Randomness Numbers: Metaheuristic and Artificial Intelligence Algorithms

Alper Ozpinar
Istanbul Commerce University, Turkey

Emel Seyma Kucukasci
Istanbul Commerce University, Turkey

ABSTRACT

The timeless search for optimizing the demand and supply of any resource is one of the main issues for humanity nearly from the beginning of time. The relevant cost of adding an extra resource reacts by means of more energy requirement, more emissions, interaction with policies and market status makes is even more complicated. Optimization of demand and supply is the key to successfully solve the problem. There are various optimization algorithms in the literature and most of them uses various algorithms of iteration and some degree of randomness to find the optimum solution. Most of the metaheuristic and artificial intelligence algorithms require the randomness where to make a new decision to go forward. So this chapter is about the possible use of chaotic random numbers in the metaheuristic and artificial intelligence algorithms that requires random numbers. The authors only provide the necessary information about the algorithms instead of providing full detailed explanation of the subjects assuming the readers already have theoretical basic information.

INTRODUCTION

Most recent optimization problems of recent years is not only limited to balancing the classical supply and demand problems as well searching for the common optimization problems related with the real-world challenges. Those challenges can be listed as sustainable energy management, logistics, transportation, production, manufacturing, consumption, healthcare, education, financial, telecommunication, cloud

DOI: 10.4018/978-1-7998-0951-7.ch017

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

computing, smart grids, internet of things and even genetic research. Most of these studies can be solved by using combinatorial optimization problems (COPs) and most of them are classified as NP-hard type problems. These real-life COPs are frequently characterized by their large-scale sizes and the need for obtaining high-quality solutions in short computing times, thus they require the use of metaheuristic and also artificial intelligence algorithms (Juan, Faulin, Grasman, Rabe, & Figueira, 2015).

In practice initializations of the stochastic and combinatorial optimization algorithms are based on the random number sequences and feeds. The modifications on the initialization directly affect the later phases and the success of the applied metaheuristics algorithm. Moreover, some of the algorithms do selections in regard to random numbers to search the solution space during the iterations. The performances of the evolutionary algorithms are asserted to be improved by using qualified random number generators (Bastos-Filho, Oliveira, Nascimento, & Ramos, 2010; Caponetto, Fortuna, Fazzino, & Xibilia, 2003).

In order to solve the complex and big domain complex problems it's a preferred way to solve them in distributed and parallel within high performance computing systems and supercomputers. The main problem steps distributed among those computing machines which are the members of the distributed system. The term computing machine refers either a single computer with multiple core CPU's, a high performance computing cluster, a supercomputer or may be a distributed grid structure formed by many computing nodes. A good distributed computing structure requires the higher level of transparency. In general transparency refers to operation which does not requires a common operation set, synchronized time settings between nodes, share of memory and variables. Many multi-core processors are homogeneous hardware architectures by means of shared-memory and direct memory access properties, meaning that all cores are identical however heterogeneous configurations of multi-core systems also exist in various computing centers (DAngelo & Marzolla, 2014). The required transparency within the high performance computing and cloud computing applications, requires the ideal case that is all the computing inventory acting like a single system even it does not.

A common problem here is increasing the number of devices results in the increasing the coincidences related with the random computing. Due to the logic and circuit system foundations the outcome of the computers is always acting in the same way unless affected by some an external feedback or source. They can create a random number by using an algorithm and mostly by implementing a deterministic function. Without a dynamic input or bias this double team never makes a surprise to its user. For example in most of the programming courses a lecture is generally includes some practical daily examples about generating random numbers with computers for lottery or card games. A simple example for shooting n random numbers with the simplest code in the computer laboratory, all the students in the lab will get the same lucky numbers in the same order. Some programming languages like c, a seeding number can be used to change the process of random number generation however this time the students using the same number for seeding sequence also gets the same numbers.

Using current time as a continuous changing and non-repeating type of seeds also does not solve the uniqueness problem of random numbers. This is the sensitive and crucial point about solving the problems in parallel and distributed. Considering the top list of super computers with more than million cores the probability of coincidence for obtaining same random numbers increase with the decrease of the quality of the random numbers. As a result either computational performance is decreased or limited search domain obtained for solution progress, which is not desired. In conclusion a good quality random number generation progress is required. The following sections explain this idea in detail with definitions and applications.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

RANDOM NUMBER GENERATORS

To generate independently identically distributed random variables with good statistical properties there are algorithms which provide random numbers that mimic most features of the truly random numbers. The purpose of these algorithms is to produce sequences of numbers or objects whose behavior is very hard to distinguish from that of their "truly random" counterparts, at least for the application of interest (Gentle, 2004).

There are two main types of random number generators available for use in computational operations. The first one is Pseudo-Random Number Generators (PRNGs), based on algorithms that use mathematical methods like linear congruential method or preset data to produce sequences of numbers. Therefore standard microcontrollers and computers without a special modification or upgrade can be classified as PRNGs type. PRNGs can operate with high performance and speed when required to generate ordinary random numbers which are sufficient enough to use in daily operations that does not require high security, uniqueness and complexity.

A pseudo-random number generator (PRNG) can be defined as an algorithm enabling to generate sequences of numbers with some properties of randomness.(Francois, Grosges, Barchiesi, & Erra, 2014) The TRNG derives its output from a physical noise source whereas a PRNG expands a relatively short key (possibly from a TRNG) into a long sequence of seemingly random bits based on a deterministic algorithm (Tsoi, Leung, & Leong, 2003). This determinism brings about the name pseudo-random instead of random. Additionally, the initial inputs of the equations are the "seeds" which are also random numbers. Let S be a set and $s \in S$ be the randomly selected seed number. Also let f be an invertible function from S to S. The pseudo-random number sequence is such that $\{s, f(s), f^2(s), f^3(s), \ldots\}$.

Another option is to use True Random Number Generators (TRNGs) as a hardware or as a software service instead of classical random number function of the computers. The need for a true random number generation is not only limited to metaheuristics and artificial intelligence. Another essential are is to secure any type of communication between electronic devices and in cryptography.

TRNGs can be performed based on different sources like electrical noise, jitter, chaos, entropy change, some irrelevant information noise, weather conditions, traffic congestion or even from elliptic curve operations and biometric sources. (Szczepanski, Wajnryb, Amiga, Sanchez-Vives, & Slater, 2004; Lee & Wong, 2004; Cicek, Pusane, & Dundar, 2014) Those sources mostly based on a natural physical phenomenon and the properties of independence as well as the unpredictability of the generated values are guaranteed by physical laws (Marton, Suciu, & Ignat, 2010). Since true randomness and true random generators do not practically exist, the name physical random number generator is also preferred in the literature. Although TRNGs provide non-predictable random sequences which cannot be repeated, the generation process is too slow and it depends on the specifications of hardware.

CHAOTIC RANDOM NUMBER GENERATION

Random numbers are useful for a variety of purposes, such as generating data encryption keys, simulating and modeling complex phenomena and selecting random samples from larger data sets (Kumar, Shukla, Prakash, Mishra, & Kumar, 2011). Additionally, qualified random numbers are also needed in Monte Carlo techniques (Gentle,), system dynamics (Eberhart & Shi, 2001; Uchida, 2012) and metaheuristics (Alatas, 2010; Bastos-Filho et al., 2010; Jahan & Akbarzadeh, 2012; Mukhopadhyay & Banerjee, 2012;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Reese, 2009). Random number generation is to find an independent identically distributed sequence with property of non-repeatability where at each repeat, a different sequence must be obtained.

Recently, different random number generators are proposed in the literature based on quantum devices, elliptic curves, cellular automata and different map functions (Akhshani, Akhavan, Mobaraki, Lim, & Hassan, 2014; dos Santos Coelho, 2009; Karimi, Hosseini, & Jahan, 2013; Pivoluska & Plesch, 2015; Reyad & Kotulski, 2015; Wang, Yu, Ding, & Leng, 2008). Random number generators are selected to be used for these applications based on desired application-specified statistical quality. There are physical generators and computational algorithms that imitate the physical generators. Additionally, random numbers can also be generated exploiting chaos theory, where the chaotic behavior of the output provides necessary statistical and computational properties.

The defined function f can also be a chaotic function and the obtained pseudo-random sequence is a chaotic pseudo-random sequence if and only if f has the following properties (Keller & Wiese, 2007):

- f reacts sensibly to changes in s, i.e. even small changes to the value of s result in large changes of the sequence,
- f is topologically transitive, i.e. almost every element of S can be connected to almost every other element of S by a finite sequence,
- f is topologically dense, i.e. even small intervals of S contain periodic points of f.

Chaotic pseudo-random number generators (CPRNG) are preferred since after a certain period length, a pseudo-random sequence repeats itself. The idea of applying discrete chaotic dynamical systems, intrinsically, exploits the property of extreme sensitivity of trajectories to small changes of initial conditions (Lozi & Taralova, 2014).

Aim of this chapter is to explain he use of Chaotic pseudo-random number generators (CPRNGs) in suitable metaheuristic and artificial intelligence algorithms on the other hand search for new generations, new weights, new distances or new parameters based on a previous step and also a randomized touch to scatter, change, improve, mutate the results for new frontiers to the solution and problem set. (Gilli, Maringer, & Schumann, 2011)

CHAOTIC MAPS

Chaotic dynamic system behavior can be described by using several map functions. The chaotic time series sequences can be obtained by using polynomial nonlinearities or piecewise linear maps (Afrabandpey, Ghaffari, Mirzaei, & Safayani, 2014; Caponetto et al., 2003; Jin & Ryu, 2012; Kim, Bok, & Ryu, 2013; May, 1976). Simple linear equations, the initial random value and the knowledge of the parameters are sufficient to generate the chaotic pseudo-random sequences. The following map functions are the mostly used goal-oriented chaotic sequences.

1. **Logistic Map:** A nonlinear function that describes the dynamic behavior of the biological populations is proposed (May, 1976). The logistic difference equation is

$$x^{t+1} = a\, x^t (1 - x^t)$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

where t is the iteration number and $a > 0$ is the biotic potential and the generated sequence is in (0,1) if $x_0 \in (0,1)$.

2. **Tent Map:** Similar to logistic map, tent map is a nonlinear function that takes as an input the previous output and generates chaotic sequences in (0,1) with the following form (Peitgen, Jürgens, & Saupe, 2006)

$$x^{t+1} = \begin{cases} x^t \,/\, 0.7, & if\ \ x^t < 0.7, \\ 10 \,/\, 3x^t\left(1 - x^t\right), & otherwise. \end{cases}$$

3. **Cubic Map:** The famous cubic map which has a wide range of applications is formally defined as (Rogers & Whitley, 1983)

$x^{t+1} = \pi\, x^t\, (1 - (x^t)^2)$

with the parameter $\pi = 2.59$. The generated chaotic sequence is in (0, 1).

4. **Gauss Map:** This map is generally utilized for testing the features of the generated sequences by using the following equations

$x^{t+1} = G(x^t)$

$$G\left(x\right) = \begin{cases} 0, & for\ \ x = 0, \\ \dfrac{1}{x} mod\left(1\right), & for\, x \epsilon \left(0,1\right). \end{cases}$$

5. **Circle Map:** This is an iterative nonlinear function that maps the circle onto itself and has the behavior of a simple sinusoidal oscillator (Essl, 2006). The simple circle map function is as follows

$$x^{t+1} = x^t + b - (a/2\pi)\sin(2\pi x^t) \bmod \quad (1)$$

6. **Sinusoidal Map:** The equation for the famous sinusoidal iterator is the following (Peitgen et al., 2006)

$x^{t+1} = a(x^t)^2 \sin(\pi x^\tau)$.

If the sequence begins with $x^0 = 0.7$ and we also have $a = 2.3$, then the equation is simplified as

$x^{t+1} = \sin(\pi x^t)$.

7. **Henon Map:** This nonlinear map is also used for testing purposes and has two parameters a and b. The representation of the map is described as

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$x^{t+1} = 1 - a(x^t)^2 + bx^{t-1})$

where a=1.4 and b=0.3 are the suggested parameters.

8. **Lozi Map:** This map is the simplified version of the Henon map with a two-dimensional piecewise linear function given by

$$(x^{t+1}, y^{t+1}) = H\left(x^t, y^t\right)$$

$$H\left(x^t, y^t\right) = \left(1 + y^t - a\left|x^t\right|, bx^t\right)$$

where a=1.7 and b=0.5 are the suggested parameters.

9. **Baker's Map:** Another two dimensional function represents Baker's map which defined as (Fridrich, 1998)

$$B\left(x\right) = \begin{cases} \left(2x, 2y\right), & for\, 0 \le x < 0.5, \\ \left(2 - 2x, 1 - \dfrac{y}{2}\right), & for\, 0.5 \le x < 1. \end{cases}$$

A one dimension of the Baker's map gives another representation which is similar to tent map given by

$$x^{t+1} = \begin{cases} 2x^t, & for\, 0 \le x^t < 0.5, \\ 2 - 2x^t, & for\, 0.5 \le x^t < 1. \end{cases}$$

10. **Arnold's Cat Map:** A two dimensional map is proposed in (Arnold & Avez, 1968) where an image of a cat is used to illustrate the chaotic behavior of the map. The representation of this map is

$$x^{t+1} = x^t + y^t \bmod \quad (1)$$

$$x^{t+1} = x^t + 2y^t \bmod \quad (1)$$

where $x^t \in (0,1)$ and $y^t \in (0,1)$.

11. **Lorenz Map (Lorenz Attractor):** Lorenz Map first introduced in (Lorenz, 1963) for a simplified model for atmospheric convection. Lorenz map is similar to logistic map and accepted as a generalization of circle homeomorphisms from zero entropy to positive entropy.(Cui & Ding, 2015) So the Lorenz map has two leaves, or branches, separated by a singular point at y=0.(Gilmore, 2015) Lorenz system based on definition of the system with three non-linear equations

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$\frac{dx}{dt} = a\left(y - x\right),\ \frac{dy}{dt} = x\left(b - z\right) - y,\ \frac{dz}{dt} = xy - cz$$

where a, b and c are constants. An example set is (10,28,8/3) and (28,46.92,4).

STATISTICAL TESTS FOR RANDOMNESS

Every PRNG should have a reasonably long period before its output sequence repeats itself, and its output sequence should pass statistical test suites so that it sufficiently resembles a truly random sequence (Keller & Wiese, 2007). The generated random numbers must have forward unpredictability property such that it is not possible to predict the future evolution of the system. Also, it must be shown that the generated pseudo-random numbers cannot be distinguished from truly random numbers. Hence, the statistical tests are proposed to justify the reliability of the generated random bit sequences.

If a pseudo-random number generator passes the next bit test then it passes any statistical test, i.e., no randomized polynomial-time algorithm can distinguish the generated pseudo-random numbers from truly random numbers (Graham, 1995). An ideal random number generator is infinite, aperiodic, uniform, uncorrelated, and computationally efficient (Phatak & Rao, 1995). For each type of algorithm, the selection of the suitable chaotic map and its parameters (i.e. initial values and control parameters) to generate the random numbers must depend on the statistical test results. Quantitative performance measurement of the chaotic driven random selections is provided by incurring most famous statistical tests (Beker & Piper, 1982; Chan, Chan, & Cheng, 2001) which are enumerated below.

1. **Frequency (Monobit) Test:** The frequencies of the 0 and 1 values in a binary sequence are tested such that they must be approximately equal. The frequency test checks that the number of ones in the sequence is not significantly different from half of the length of the sequence (Kanso & Smaoui, 2009)
2. **Serial (Two-Bit) Test:** Random sequences must have the property that value of each element of a sequence must be independent from the value of its predecessor. Hence, the frequencies of the different transitions in a binary sequence such that 11, 10, 01 and 00 must be approximately equal.
3. **Runs Test:** To test the hypothesis of the independence, the arrangement of numbers in a sequence is examined. After the serial test is passed, the number of runs in the sequence also must be checked to be a dichotomous sequence of values.
4. **Poker Test:** This test determines whether the sequences of different lengths each appear approximately the same number of times in the sequence as it is expected from a random sequence (Karimi et al., 2013). Obviously, the frequency test and the serial test are contained in the poker test.

USE OF CHAOTIC RANDOMNESS IN SELECTED ALGORITHMS

In this section of the chapter, the possible use of Chaotic pseudo-random number generators (CPRNG) for proper metaheuristic and artificial intelligence algorithms has explained. As mentioned in the beginning of this chapter there are vast number of research papers and books about different implementations and descriptions of the algorithms therefore instead of providing the detailed technical and theoretical

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

background only sample solutions for adapting CPRNG's is provided. (Draco, Petrowski, Siarry, & Taillard, 2006; Glover & Kochenberger, 2003; Gonzalez, 2007; Luke, 2009; Souza, Omkar, & Senthilnath, 2012; Talbi, 2009; Yegnanarayana, 2009).

Metaheuristic and artificial intelligence optimization algorithms search for an optimum in the problem domain without trapping in to local optimums, avoiding fall in to infinite loops and careful about early victories due to memorization or other issues. The main reason behind those is optimization is different from modelling or learning. Most of the optimizations based on searching for an unknown solution set. This uncertainty in algorithm sometimes gives a candidate optimum with the possibility of the local optimum accepted as a global or searching for the neighbors sometimes results in missing the optimum next to your current location. So there are also optimum parameters in the algorithm which searches the optimum of the problem. Also computational and algorithmic complexity is another concern while working with computers, high performance computing machines as well as the supercomputers and cloud computing where the computational limits far beyond normal acceptance. A balance of amount of resources like CPU, ram and computational time versus the possible algorithms available has to be obtained in order to solve the optimization problem.

In most of the methods two common function types can be used, one of them is to produce a direct random variable with "CPRNG" methods and normalize if necessary and the second one is "CPRNGTweak" method that will change and modify the original or candidate value with a degree of modification and a random number obtained from CPRNG.

CPRNG Generation

```
CPRNGRandom <- PRNGRandom number from ordinary equations
t <- initialization for chaotic map
repeat
test_status <- true
repeat t times do
 map(t+1) <- chaotic_map_function(map(t))
 CPRNGRandom <-map(t+1)
 repeat all_tests do
 if test fails then
 test_status=false
 break
until test_status is true or time limit reached
return CPRNGRandom
```

Simple Methods and Random Search

Simples search algorithms like Gradient Ascent, Newton's Method, Steepest Ascent Hill-Climbing and Random Search methods are the easy and simple methods to start with. All of these methods start with an initial random point or candidate point within the search domain. The rest is more or less in common, search the optimality and the validity of the candidate, if the stopping criteria is not satisfied then change the candidate a little with a random affect and return to the loop. A search based on pure random candi-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

date for each time is not a good way of searching, instead using a brute force or combinatorial search is better. Searching in in single, parallel or multi node computing adapting by replacing the random number operator with CPRNG will improve the solution diversity and performance.

Improved Hill-Climbing with Random Restarts with CPRNG

```
time_interval <- distribution of possible time intervals
candidate <- A Random CPRNG Candidate
optimum <- candidate
repeat
time <- a random time in near future from time_interval set
repeat
 mid_candidate <- CPRNGTweak (candidate)
if Quality(mid_candidate) > Quality (candidate)
then
candidate <- mid_candidate
 until stopping criteria satisfied with candidate, time is up or timeout
 if Quality(candidate) > Quality (optimum)
then
 optimum <- candidate
 candidate <- A Random CPRNG Candidate
until stopping criteria satisfied with optimum or timeout
Improved Random Search with CPRNG
optimum <- A Random CPRNG Candidate
repeat
candidate <- A random CPRNG Candidate
if quality(candidate) > quality (optimum)
then
optimum <- candidate
until stopping criteria satisfied with optimum or timeout
```

Simulated Annealing (SA)

Simulated Annealing is another global search algorithm for global optimization similar the Hill-Climbing with the difference of changing the candidate and intermediate candidate. Annealing is actually a heat application process in manufacturing and materials science that changes the physical and chemical properties of a material in order to obtain harder more ductile products. The Annealing process initially heating up the material above a suitable critical temperature where the material starts recrystallization and forms better bigger crystal structures which provides the ductility and hardness. After that the process continues with waiting for a proper time which allows slowly cooling so the crystal atoms becomes stable and forms new crystal structure.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Improved Simulated Annealing with CPRNG

```
T <- a high temperature value //melting
candidate <- A Random CPRNG Candidate
optimum <- candidate
repeat
mid_candidate <- CPRNGTweak (candidate)
if quality(mid_candidate) > quality(candidate) or
CPRNG[0..1] < exp((quality(mid_candidate)-quality(candidate))/T)
then
candidate <- mid_candidate
decrease t //cooling
if quality(candidate) > quality (optimum)
then
optimum <- candidate
until stopping criteria satisfied with optimum, T<=0 or timeout
```

Tabu Search (TS)

Tabu Search is a search algorithm using an imaginary memory of the recently tested candidate solutions by putting them on a list called tabu list with a predefined capacity m of storing them. Like memories, tabu lists are volatile based on the time passed or list fully filled with candidates with FIFO principle. It is based on a local search of the neighbors and checking the list if visited before or not. n modified and tweaked candidate children is generated in each step but the ones in the memory is not considered.

Improved Tabu Search CPRNG

```
m <- tabu list length //memory size
n <- number of tweaked candidates desired for sample the gradient
candidate <- some initial candidate solution CPRNG Candidate
optimum <- candidate
tabu_list <-{} a tabu list of length m //FIFO based
enqueuer candidate to memory
repeat
if tabu_list.LEN = m
then
apply FIFO remove the oldest from tabu_list
mid_candidate <- CPRNGTweak (candidate)
for n-1 times do
dummy_candidate <- CPRNGTweak (candidate)
if dummy_candidate not in tabu_list and
quality(W)> quality(mid_candidate) or mid_candidate is in tabu_list
then
mid_candidate <- dummy_candidate
```


EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

```
if mid_candidate is not a member of memory and
quality(mid_candidate)> quality(candidate)
then
candidate <- mid_candidate
 enqueue mid_candidate to memory
if quality(candidate) > quality (optimum)
then
optimum <- candidate
until stopping criteria satisfied with optimum or timeout
```

Particle Swarm Optimization (PSO)

Almost the same as the inspiration of the nerve cells particle swarm optimization motivated/brought about from the behaviors of the different huge numbers in the nature like birds, fishes, ants. Particle swarm optimization thought about/believed as population based optimization and can be improved by using chaotic random number modified members, crossovers and mutations. In general PSO a set of computer instructions is used for optimization where candidate solution sets or working limits/guidelines for the population modelled according to the general behaviors. These candidate solution sets which are called particles in solution swarm, live together and change and get better at the same time based on knowledge sharing with close-by particles. Each particle acts individually and also in coordination with the other particles. So each particle creates a solution based on the problem statement and initial starting point. In each turn all the particles changes their location, limits/guidelines, settings by using a speed vector to reach its next position in the solution space. Each particle changes its speed in order to reach a better position than the previous one by using its own best previous experience data and also the best previous experience data from the neighbors within the swarm.

Particles update their velocities as shown in equations below

$$v_{t+1}^{i} = w_t v_t^i + c_1 r_1 \left(p_t^i - x_t^i \right) + c_2 r_2 \left(p_t^g - x_t^i \right)$$

and the new positions according to

$$x_{t+1}^{i} = x_t^i + v_{t+1}^i$$

where x_t^i represents the current position of particle i in solution space and subscript t indicates an iteration, p_t^i is the best-found position of particle i up to iteration count t and represents the cognitive contribution to the search velocity v. Each component of v_t^i can be limited to the range [$-v_{\max}$,$v_{\max}$] to control excessive roaming of particles within the problem domain and creating unfeasible solutions. p_t^g is the global best-found position among all particles in the problem domain up to iteration count t and forms the swarm information flow to the velocity vector; r_1 and r_2 are random numbers uniformly distributed in the interval (0,1), where the chaotic random numbers will be replaced, while c_1 and c_2 are the

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

cognitive and social scaling parameters depending on the structure of problem. Finally w_t is the particle inertia, which is reduced dynamically to decrease the search area in a gradual fashion. The variable w_t is updated as

$$w_t = \left(w_{max} - w_{min}\right) * \frac{t_{max} - t}{t_{max}} + w_{min}$$

where, $w_{\max}$ and $w_{\min}$ denote the maximum and minimum of w_t respectively; $t_{\max}$ is a given number of maximum iterations. So the swarm members repeated the communication and changed their parameters until a time limit or a stopping criteria values reached.

Improved Particle Swarm Optimization

```
for each candidate particle in the candidate particle set
Initialize particle with working parameters
End
repeat
for each particle
Calculate fitness value for better
if the fitness value is better than the best fitness value (p_Best) in history
set current value as the new p_Best
End
Choose the particle with the best fitness value of all the particles as the
g_Best
for each particle
Calculate particle adapting velocity using
update particle working parameters r_1 and r_2 using CPRNG
end
until stopping criteria satisfied with optimum or timeout
```

Genetic Algorithms (GA)

Genetic algorithm (GA), is the most widely used version of evolutionary algorithms in the metaheuristic. Similar to the inspiration from the nature as well as ANN and PSO, genetic algorithms used the structure of evolution in order to reach the perfect solution that fits in to the energy environment. As in the nature every individual member of the has set of properties and working parameters which differs that device among the rest of the solutions in the domain.

These properties and working parameters corresponds to the chromosomes or genotypes similar to the ones in the living organisms (John, 1992).

As the nature changes and gets better in the time domain with new generation reproduced or composed of their parents or family relatives or things that existed long, long ago genotypes by crossover and change. Each new next generation member has a possible chance to better fit to its surrounding conditions. The rules of related to tiny chemical assembly instructions inside of living things sets of

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

computer instructions can be changed to fit optimization problems. Most of the numerical adaptations of the problems numerical values decoded as binary strings so easy to make different levels and sources of crossovers and mutations (Kumar & Mohan, 2011). The common steps of genetic algorithm is as follows

1. **Representations:** Proper coding of the optimization problem in to mathematical chromosomes while the decision variables within a solution (chromosome) are genes. The possible values of variables (genes) are the alleles and the position of an element (gene) within a chromosome is named locus. Most common representation is based on bits. Binary, discrete and permutation representations are available in the literature.
2. **Population Initialization:** Using the representation coding an initial state of the problem settings created and represented as defined on the first step. As to continue from the previous step, initial population contains some power stations are operational and some are not in a particular initial timing.
3. **Objective Function:** Proper coding of the optimization problem in to mathematical chromosomes which will be used to calculate the goodness of the solution.
4. **Selection Strategy:** The best parameter sets or working plans which fit in to the objective function selected for the next round of reproduction. In the rank based selections, without looking the proportions best n solutions selected for the reproduction. Roulette Wheel Selection, Stochastic Universal Sampling, Tournament Selection, Rank-Based Selection are common selection methodologies.
5. **Reproduction with CPNRG:** Using the current solutions sets best solutions will produce the next generation alternatives by the use of chaotic random numbers
    a. **Chaotic Crossover:** Crossover is used to increase the diversity of the possible solutions.
    b. **Chaotic Mutations:** New genetic material into the gene at some minor ratio. With a small probability, randomly chosen bits of the offspring genotypes change from '0' to '1' and vice versa for binary coded strings. This will increase the possibility of the new searches in the domain.
6. **Replacement Strategy:** The best solutions which fit in to the objective function selected for the next round of reproduction
7. **Stopping Criteria:** The best solutions which fit in to objective function selected for the next round of reproduction

Like most of the other artificial optimization methods genetic algorithm can be used for complex problems.

Improved Genetic Algorithms with CPRNG

```
generate an initial population with CPRNG
repeat
Repeat
If crossover condition satisfied Then
select parent chromosomes
choose crossover parameters
perform crossover with CPRNGCrossover
End if
```


EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

```
If mutation condition satisfied Then
choose mutation points
perform mutation with CPRNGMutation
End if
evaluate fitness of offspring
Until sufficient offspring created;
select new population;
until stopping criteria satisfied with optimum or timeout
```

Ant Colony Algorithm (ACO)

Ant Colony Optimization (ACO) is an another swarm intelligence based metaheuristic algorithm that is inspired by the pheromone trail laying and tracking of the remaining of ants while finding their way to food or home. Use of artificial ants in ACO for optimization problems are stochastic solutions that build possible candidate solutions for the optimization settings and parameters of the smart grid by adapting the artificial pheromone in other words fitness information. ACO mostly used for discrete optimization problems. The real ants uses stigmergy which is a form of indirect communication by modifying the environment for further usage or communication. In case of the ACO, stigmergy based on the pheromone trails leaved by the ants while passing through a path or way.

Improved Ant Colony Optimization with CPRNG

```
Initialize the base efficiency and fitness (attractiveness), τ, and visibil-
ity, η, for each solution set ;
repeat
for each ant do
choose CPRNGTweak(previous equation) the next state to move into
add that move to the tabu list for each smart agent
repeat until each smart agent completed a solution for dynamic response
End
For each ant that completed a solution do
update fitness τ for each edge that the ant traversed;
End
if (local best solution better than global solution) then
save local best solution as global solution;
End
until stopping criteria satisfied with optimum or timeout
```

Chaotic Crossover

In crossover is the operation of mixing and matching the genes of two parents in to new generations. Mixing and matching is totally random based on some rules. A good quality random numbers will increase the variety of the possible new generations and increase the search domain. In the general application there are three possible crossovers available as one-point, two-point and uniform crossovers. In all of these crossovers random locations selected for crossover.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Improved Single Point Crossover with CPRNG

```
parent_a_dna <- genetic coding of the problem to the representation
parent_b_dna <- genetic coding of the problem to the representation
c <- a CPRNG random integer chosen uniformly from 1 to length of dna
if c<> 1 then
for I from 1 to c-1 do
swap the values of parent_a_dna(i) and parent_a_dna(i)
Improved Two Point Crossover with CPRNG
parent_a_dna <- genetic coding of the problem to the representation
parent_b_dna <- genetic coding of the problem to the representation
c1 <- a CPRNG random integer chosen uniformly from 1 to length of dna
c2 <- a CPRNG random integer chosen uniformly from 1 to length of dna
if c1>c2 then
swap c1,c2
if c1<>c2 then
for i from c1 to c2-1 do
swap the values of parent_a_dna(i) and parent_a_dna(i)
```

Improved Uniform Crossover with CPRNG

```
swap_probability < - setting a 1/length_of_dna, <=0.5
parent_a_dna <- genetic coding of the problem to the representation
parent_b_dna <- genetic coding of the problem to the representation
for i from 1 to length_of_dna do
if p>= CPRNG[0..1] then
swap the values of parent_a_dna(i)and parent_a_dna(i)
```

Chaotic Mutation

Mutation is an operator used to maintain many different kinds of possible solutions from one generation of a problem set to the next. It is the same as change of values in a from its initial state. In mutation, the solution may change wholly from the previous solution. Hence GA can come to a better solution or sometimes an unfeasible solution. If the degree of mutation it is set too high, the optimization will turn into a random search. A simple mutation operator is the bit-flip operator for binary coding's.

Improved Bit-Flip Mutation with CPRNG

```
flip_probability < - setting a probability of flipping a bit
parent_dna <- genetic coding of the problem to the representation
for i from 1 to length_of_dna do
if p>= CPRNG[0..1] then
flip the values of parent_a_dna(i) as 1>0, 0>1
```


EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

FUTURE RESEARCH DIRECTIONS

The application of true random number generations with modified chaotic maps planned by the authors for the future research directions. The main idea behind the progress is to modify the nature inspired events and conditions with chaotic map functions and generate a hybrid version of semi TRNGs and semi CPRNGS.

CONCLUSION

Metaheuristic and artificial intelligence algorithms are widely used in different optimization, modelling and learning problems. Stochastic optimization and search algorithms require the use of random number for initialization of the problem set and tweaks of the current variables to the next calculation step. Increasing the uniqueness and the quality of the random numbers will eventually increase the processing time and the goodness of the solution. This is a more obvious requirement in parallel and grid computing based calculations where the probability of the coincidence of the same random numbers occurring in the system increased in using ordinary PRNGs. The use of TRNG is a possibility with increasing the computational complexity which costs energy and time. This chapter makes a common review of the subjects and gives examples of selected methodologies. As can be seen from the selected examples any algorithm that starts with random initial position, random variables and random modifications within the routine can be modified with chaotic pseudo random number generators and functions.

REFERENCES

Afrabandpey, H., Ghaffari, M., Mirzaei, A., & Safayani, M. (2014). A novel bat algorithm based on chaos for optimization tasks. *Intelligent Systems (ICIS), 2014 Iranian Conference on* (pp. 1-6) IEEE. 10.1109/IranianCIS.2014.6802527

Akhshani, A., Akhavan, A., Mobaraki, A., Lim, S. C., & Hassan, Z. (2014). Pseudo random number generator based on quantum chaotic map. *Communications in Nonlinear Science and Numerical Simulation*, *19*(1), 101–111. doi:10.1016/j.cnsns.2013.06.017

Alatas, B. (2010). Chaotic bee colony algorithms for global numerical optimization. *Expert Systems with Applications*, *37*(8), 5682–5687. doi:10.1016/j.eswa.2010.02.042

Arnold, V. I., & Avez, A. (1968). *Ergodic problems of classical mechanics*. Retrieved from WA Benjamin.

Bastos-Filho, C. J., Oliveira, M. A., Nascimento, D. N., & Ramos, A. D. (2010). Impact of the random number generator quality on particle swarm optimization algorithm running on graphic processor units. *Hybrid Intelligent Systems (HIS), 2010 10th International Conference on* (pp. 85-90) IEEE. 10.1109/HIS.2010.5601073

Beker, H., & Piper, F. (1982). *Cipher systems: the protection of communications*. Northwood Books.

Caponetto, R., Fortuna, L., Fazzino, S., & Xibilia, M. G. (2003). Chaotic sequences to improve the performance of evolutionary algorithms. *Evolutionary Computation*. *IEEE Transactions on*, *7*(3), 289–304.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chan, C. K., Chan, C. K., & Cheng, L. M. (2001). *Software Generation of Random Numbers by Using Neural Network. Artificial Neural Nets and Genetic Algorithms*. Springer. doi:10.1007/978-3-7091-6230-9_51

Cicek, I., Pusane, A. E., & Dundar, G. (2014). A novel design method for discrete time chaos based true random number generators. *Integration, the VLSI Journal, 47*(1), 38-47. doi:10.1016/j.vlsi.2013.06.003

Cui, H., & Ding, Y. (2015). Renormalization and conjugacy of piecewise linear Lorenz maps. *Advances in Mathematics*, *271*, 235–272. doi:10.1016/j.aim.2014.11.024

D'Angelo, G., & Marzolla, M. (2014). New trends in parallel and distributed simulation: From many-cores to Cloud Computing. *Simulation Modelling Practice and Theory*, *49*, 320–335. doi:10.1016/j.simpat.2014.06.007

dos Santos Coelho, L. (2009). Reliability and redundancy optimization by means of a chaotic differential evolution approach. *Chaos, Solitons, and Fractals*, *41*(2), 594–602. doi:10.1016/j.chaos.2008.02.028

Draco, J., Petrowski, A., Siarry, P., & Taillard, E. (2006). *Metaheuristics for hard optimization: methods and case studies*. Springer Science & Business Media.

Eberhart, R. C., & Shi, Y. (2001). Tracking and optimizing dynamic systems with particle swarms. *Evolutionary Computation, 2001.Proceedings of the 2001 Congress on* (vol. 1, pp. 94-100) IEEE. 10.1109/CEC.2001.934376

Essl, G. (2006). Circle maps as a simple oscillators for complex behavior: Ii. experiments. In *Proceedings of the International Conference on Digital Audio Effects (DAFx)*.

Francois, M., Grosges, T., Barchiesi, D., & Erra, R. (2014). Pseudo-random number generator based on mixing of three chaotic maps. *Communications in Nonlinear Science and Numerical Simulation*, *19*(4), 887–895. doi:10.1016/j.cnsns.2013.08.032

Fridrich, J. (1998). Symmetric ciphers based on two-dimensional chaotic maps. *International Journal of Bifurcation and Chaos in Applied Sciences and Engineering*, *8*(6), 1259–1284. doi:10.1142/S021812749800098X

Gentle, J. E., Hardle, W., & Mori, Y. (Eds.). (2004). Handbook of Computational Statistics. Springer.

Gilli, M., Maringer, D., & Schumann, E. (2011). Generating Random Numbers. In M.G.M. Schumann (Ed.), Numerical Methods and Optimization in Finance (pp. 119-158). San Diego, CA: Academic Press.

Gilmore, R. (2015). Explosions in Lorenz maps. *Chaos, Solitons, and Fractals*, *76*, 130–140. doi:10.1016/j.chaos.2015.03.020

Glover, F., & Kochenberger, G. A. (2003). *Handbook of metaheuristics Springer Science & Business Media*.

Gonzalez, T. F. (2007). *Handbook of approximation algorithms and metaheuristics CRC Press*. doi:10.1201/9781420010749

Graham, R. L. (1995). *Handbook of combinatorics* (1st ed.). Elsevier.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Jahan, M. V., & Akbarzadeh, T. (2012). Hybrid local search algorithm via evolutionary avalanches for spin glass based portfolio selection. *Egyptian Informatics Journal*, *13*(2), 65–73. doi:10.1016/j.eij.2012.04.002

Jin, C. H., & Ryu, H. G. (2012). Performance evaluation of chaotic CDSK modulation system with different chaotic maps. *ICT Convergence (ICTC), 2012 International Conference on* (pp. 603-606) IEEE. 10.1109/ICTC.2012.6387115

John, H. H. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT Press.

Juan, A. A., Faulin, J., Grasman, S. E., Rabe, M., & Figueira, G. (2015). A review of simheuristics: Extending metaheuristics to deal with stochastic combinatorial optimization problems. *Operations Research Perspectives*, *2*, 62–72. doi:10.1016/j.orp.2015.03.001

Kanso, A., & Smaoui, N. (2009). Logistic chaotic maps for binary numbers generations. *Chaos, Solitons, and Fractals*, *40*(5), 2557–2568. doi:10.1016/j.chaos.2007.10.049

Karimi, H., Hosseini, S. M., & Jahan, M. V. (2013). On the combination of self-organized systems to generate pseudo-random numbers. *Information Sciences*, *221*, 371–388. doi:10.1016/j.ins.2012.09.029

Keller, J. A., & Wiese, H. (2007). Period lengths of chaotic pseudo-random number generators. *Proceedings of the Fourth IASTED International Conference on Communication, Network and Information Security* (pp. 7-11). ACTA Press.

Kim, S., Bok, J., & Ryu, H. G. (2013). Performance evaluation of DCSK system with chaotic maps. *Information Networking (ICOIN), 2013 International Conference on* (pp. 556-559). IEEE.

Kumar, J., Shukla, S., Prakash, D., Mishra, P., & Kumar, S. (2011). Random Number Generator Using Various Techniques Through VHDL. *International Journal of Computer Applications in Engineering Sciences, 1*.

Kumar, V. S., & Mohan, M. R. (2011). A genetic algorithm solution to the optimal short-term hydrothermal scheduling. *International Journal of Electrical Power & Energy Systems*, *33*(4), 827–835. doi:10.1016/j.ijepes.2010.11.008

Lee, L. P., & Wong, K. W. (2004). A random number generator based on elliptic curve operations. *Computers & Mathematics with Applications, 47*(2-3), 217-226. doi:10.1016/S0898-1221(04)90018-1

Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the Atmospheric Sciences*, *20*(2), 130–141. doi:10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2

Lozi, R., & Taralova, I. (2014). From chaos to randomness via geometric undersampling. *ESAIM: Proceedings and Surveys, 46*, 177-195. Retrieved from EDP Sciences.

Luke, S. (2009). *Essentials of Metaheuristics. A Set of Undergraduate Lecture Notes*. Zeroth Edition.

Marton, K., Suciu, A., & Ignat, I. (2010). Randomness in digital cryptography: A survey. Romanian Journal of Information Science and Technology. *ROMJIST*, *13*(3), 219–240.

May, R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature*, *261*(5560), 459–467. doi:10.1038/261459a0 PMID:934280

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Mukhopadhyay, S., & Banerjee, S. (2012). Global optimization of an optical chaotic system by chaotic multi swarm particle swarm optimization. *Expert Systems with Applications*, *39*(1), 917–924. doi:10.1016/j.eswa.2011.07.089

Peitgen, H. O., Jürgens, H., & Saupe, D. (2006). *Chaos and fractals: new frontiers of science*. Springer Science & Business Media.

Phatak, S. C., & Rao, S. S. (1995). Logistic map: A possible random-number generator. *Physical Review E: Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, *51*(4), 3670–3678. doi:10.1103/PhysRevE.51.3670 PMID:9963048

Pivoluska, M., & Plesch, M. (2015). *Device Independent Random Number Generation*. arXiv:1502.06393

Reese, A. (2009). Random number generators in genetic algorithms for unconstrained and constrained optimization. Nonlinear Analysis: Theory. *Methods & Applications*, *71*(12), e679–e692.

Reyad, O., & Kotulski, Z. (2015). On Pseudo-Random Number Generators Using Elliptic Curves and Chaotic Systems. *Applications of Mathematics*, *9*(1), 31–38.

Rogers, T. D., & Whitley, D. C. (1983). Chaos in the cubic mapping. *Mathematical Modelling*, *4*(1), 9–25. doi:10.1016/0270-0255(83)90030-1

Souza, C., Omkar, S. N., & Senthilnath, J. (2012). Pickup and delivery problem using metaheuristics techniques. *Expert Systems with Applications*, *39*(1), 328–334. doi:10.1016/j.eswa.2011.07.022

Szczepanski, J., Wajnryb, E., Amiga, J. M., Sanchez-Vives, M. V., & Slater, M. (2004). Biometric random number generators. *Computers & Security*, *23*(1), 77–84. doi:10.1016/S0167-4048(04)00064-1

Talbi, E. G. (2009). Metaheuristics: from design to implementation (74th ed.). John Wiley & Sons.

Tsoi, K. H., Leung, K. H., & Leong, P. H. W. (2003). Compact FPGA-based true and pseudo random number generators. *Field-Programmable Custom Computing Machines, 2003.FCCM 2003.11th Annual IEEE Symposium on* (pp. 51-61). IEEE.

Uchida, A. (2012). *Optical communication with chaotic lasers: applications of nonlinear dynamics and synchronization*. John Wiley & Sons.

Wang, Q., Yu, S., Ding, W., & Leng, M. (2008). Generating high-quality random numbers by cellular automata with PSO. *Natural Computation, 2008.ICNC'08.Fourth International Conference on* (vol. 7, pp. 430-433). IEEE. 10.1109/ICNC.2008.560

Yegnanarayana, B. (2009). *Artificial neural networks*. PHI Learning Pvt. Ltd.

ADDITIONAL READING

Afrabandpey, H., Ghaffari, M., Mirzaei, A., & Safayani, M. (2014, February). A novel bat algorithm based on chaos for optimization tasks. In Intelligent Systems (ICIS), 2014 Iranian Conference on (pp. 1-6). IEEE. 10.1109/IranianCIS.2014.6802527

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Alligood, K. T., Sauer, T. D., & Yorke, J. A. (1997). [Springer Berlin Heidelberg.]. *Chaos (Woodbury, N.Y.)*, 105–147.

Bastos-Filho, C. J., Andrade, J., Pita, M. R., & Ramos, A. D. (2009, October). Impact of the quality of random numbers generators on the performance of particle swarm optimization. In Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on (pp. 4988-4993). IEEE. 10.1109/ICSMC.2009.5346366

Ding, K., & Tan, Y. (2014, July). Comparison of random number generators in Particle Swarm Optimization algorithm. In Evolutionary Computation (CEC), 2014 IEEE Congress on (pp. 2664-2671). IEEE. 10.1109/CEC.2014.6900461

Feng, Y., Teng, G. F., Wang, A. X., & Yao, Y. M. (2007, September). Chaotic inertia weight in particle swarm optimization. In Innovative Computing, Information and Control, 2007. ICICIC'07. Second International Conference on(pp. 475-475). IEEE. 10.1109/ICICIC.2007.209

Kazimipour, B., Li, X., & Qin, A. K. (2014, July). A review of population initialization techniques for evolutionary algorithms. In Evolutionary Computation (CEC), 2014 IEEE Congress on (pp. 2585-2592). IEEE. 10.1109/CEC.2014.6900618

Kromer, P., Snael, V., & Zelinka, I. (2013, October). On the Use of Chaos in Nature-Inspired Optimization Methods. In Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on (pp. 1684-1689). IEEE. 10.1109/SMC.2013.290

Krömer, P., Zelinka, I., & Snášel, V. (2014). Behaviour of pseudo-random and chaotic sources of stochasticity in nature-inspired optimization methods. *Soft Computing*, *18*(4), 619–629. doi:10.100700500-014-1223-y

Lin, H., & Yim, S. C. S. (1996). Analysis of a nonlinear system exhibiting chaotic, noisy chaotic, and random behaviors. *Journal of Applied Mechanics*, *63*(2), 509–516. doi:10.1115/1.2788897

Niederreiter, H. (1978). Quasi-Monte Carlo methods and pseudo-random numbers. *Bulletin of the American Mathematical Society*, *84*(6), 957–1041. doi:10.1090/S0002-9904-1978-14532-7

Parejo, J. A., Ruiz-Cortés, A., Lozano, S., & Fernandez, P. (2012). Metaheuristic optimization frameworks: A survey and benchmarking. *Soft Computing*, *16*(3), 527–561. doi:10.100700500-011-0754-8

Skiadas, C. H., & Skiadas, C. (2008). *Chaotic modelling and simulation: analysis of chaotic models, attractors and forms*. CRC Press. doi:10.1201/9781420079012

Stojanovski, T., & Kocarev, L. (2001). Chaos-based random number generators-part I: analysis [cryptography]. Circuits and Systems I: Fundamental Theory and Applications. *IEEE Transactions on*, *48*(3), 281–288.

Stojanovski, T., Pihl, J., & Kocarev, L. (2001). Chaos-based random number generators. Part II: practical realization. Circuits and Systems I: Fundamental Theory and Applications. *IEEE Transactions on*, *48*(3), 382–385.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Timmer, J., Rust, H., Horbelt, W., & Voss, H. U. (2000). Parametric, nonparametric and parametric modelling of a chaotic circuit time series. *Physics Letters. [Part A]*, *274*(3), 123–134. doi:10.1016/S0375-9601(00)00548-X

Toso, R. F., & Resende, M. G. (2015). A C++ application programming interface for biased random-key genetic algorithms. *Optimization Methods and Software*, *30*(1), 81–93. doi:10.1080/10556788.2014.890197

KEY TERMS AND DEFINITIONS

Chaotic Crossover: An evolutionary crossover based on the chaotic random numbers.

Chaotic Mutation: An evolutionary mutation based on the chaotic random numbers.

Chaotic Pseudo-Random Number Generators (CPRNG): PRNG based on chaotic maps and chaotic functions.

CPRNG Tweak Function: A modification function that uses the original value and a random number from CPRNG to tweak and modify the value.

Pseudo-Random Number Generators (PRNGs): A process for generating random numbers based on deterministic algorithms.

Random Number Generation: A process for generating tested random numbers for different purposes.

True-Random Number Generators (TRNGs): A process for generating random numbers based on non-deterministic random number generation.

This research was previously published in Intelligent Techniques for Data Analysis in Diverse Settings edited by Numan Celebi; pages 207-227; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 18
The Media-Sphere as Dream
Researching the Contextual Unconscious of Collectives

Stephen Brock Schafer
Digipen Institute of Technology, USA

ABSTRACT

The psychological nature of the electronic media environment is a virtual reality that—according to Jungian principles—is dreamlike. Perhaps it can be analyzed with Jung's Analytical Psychology. Science is experiencing a paradigm shift into a reality of mediated illusion, and psychological research on this illusion has become the human imperative. It may be stipulated that physics has abolished matter, conceding that "reality is organized mind stuff." If cosmos is structured holographically and the brain is structured holonomically, it is probable that "mind stuff" is structured holographically. The Jungian concept of Psyche is a good place to begin researching the Media-sphere as mind stuff. Cognitive sciences are probing the brain and nervous system in search of the template for cognitive organization, and the salient features have already emerged. It appears that both conscious and unconsciousness cognitive dimensions have dramatic form. This dreamlike structure can be employed to analyze the media dream, and to foster coherent psychological states in contextual collectives.

INTRODUCTION

Carl Jung used symbolic patterns occurring in dreams to promote understanding of contextual meaning residing in a patient's unconscious. Jung equated the patterns of the unconscious with atomic structure and called them archetypes. Dream symbols were understood to be projections from unconscious dimensions, but because they were relatively conscious, Jung called the projections archetypal representations which—as projections—mirrored their source. In this way, discovery of unconscious meaning not only becomes possible, it becomes measurable according to degrees of coherence. Based on a number of correlations between Jungian dreams and media dreams, the hypothesis of this chapter is that the metaverse constitutes a dreamscape, and its holographic images constitute contextual media dreams.

DOI: 10.4018/978-1-7998-0951-7.ch018

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

As this chapter will document, according to Jungian depth analysis, a series of dreams will disclose unconscious patterns of meaning of which a patient is unaware. The psychiatrist can use principles of Analytical Psychology to extrapolate the meaning of dream symbols for the purpose of providing feedback to the patient as to the meaning of dream symbols. Given this feedback, the patient is positioned to gain meaningful insight that can have healing affect—much like biofeedback. Though Jung's research was empirically based, more recent cognitive research on heart-rate coherence—to be documented later—has verified that media dreams are susceptible to scientific measurement. Therefore, just as a psychiatrist provides feedback to a patient and as advanced algorithms provide feedback as to a computer user's preferences, research addressed to specific (contextual) target groups could provide mediated feedback designed to foster healing insight to contextual collectives

According to Jung, dream sequences (Jung, 1933, p 14) provide a medium for communication between conscious and unconscious cognitive dimensions:

Dreams give information about the secrets of the inner life and reveal to the dreamer hidden factors of his personality. As long as these are undiscovered, they disturb his waking life and betray themselves only in the form of symptoms. This means that we cannot effectively treat the patient from the side of consciousness alone, but must bring about a change in and through the unconscious. (Jung, 1933, p. 16)

So—in Jungian practice—dreams are not negative. Even if they are nightmares, dreams have a benign function because they hold the keys to meaningful insight that can result in multi-functional (physical, mental, emotional, spiritual) healing. Accordingly, even if the mediated dreams of the metaverse—psychologically multi-functional images—are symbolically saturated with sex, violence, and nightmarish threat, these patterns must be understood as positive, potentially therapeutic messages from the contextual unconscious. If recognized as such, these "dreams" can promote meaningful insight that could lead to unprecedented healing of collective personae. Contextual collectives are much like target groups and come in many formats—great and small—such as Rockets fans, Coca Cola drinkers, stay-at-home-pops, African Americans, white supremacists, environmentalists, Republicans, Buddhists, females, Germans, etc..

It would be safe to say that technological media research in the context of holographic cognitive function—though incipient—is advancing apace. The technological media has long been the prime mover in the creation of modern human reality, but now it becomes clear that using that immense power wisely will not be easy. There can be no doubt that this media power is rooted in the capacity to reach into the *cognitive unconscious* of individuals and demographic groups in order to alter their perceptions of reality by reframing archetypal patterns. The fact that commercial and political motives have skewed the programming of mass media for at least half a century need not be argued.

In advertising, politics, and education, though much is already known about the psychological power of the mass media, proportional efforts to manifest positive potentials for the power of psychological mass media have not been forthcoming. The primary purpose of this chapter is to discuss these potentials in necessarily innovative terms: Dreams, functional coherence, learning, healing, and even transcendence.

Taking a prospective view of benign media purposes, Internet potentials for education, healing, and maintaining coherent states in dimensions of the cognitive unconscious are just beginning to emerge. In *Amusing Ourselves to Death*, (1995) Neil Postman argued that any research on media should recognize that each medium has its own personality, its deficits, and its potentials. In *The Media Equation*, (1996) Stanford researchers Byron Reeves and Clifford Nass concluded that at unconscious levels and in all demographics, people perceive media as "real"—that is to say—unconsciously—all people experience

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

media in terms of personality and psychological experience based in social interaction. These incipient works began to establish parameters for a millennial study of the media dream. More recently, the Immersive Education Initiative, an offshoot of MediaGrid (2015), continues a global effort to advance understanding of the uses of technological media in educational contexts.

Jung says that "dreams have a purpose" (Jacobi, 1973), so if the media dream has Jungian veracity, the media dreamscape is imbedded with the holographic patterns of that purpose in fractal form. Most importantly, the path to Jungian *Individuation*—symbolized by the archetypal Journey of the Hero as process—consummation of which may be understood as the ultimate purpose of all human evolutionary (transcendent) experience, lies latent in the digital/virtual field that is energetically correlated with the Jungian collective unconscious.

In other words, if the mathematics of Fourier transforms can demonstrate recursive correlation between energy configurations (archetypes) in the collective unconscious and the energy images they project in dreams (archetypal representations) energy patterns in digital programs (algorithms) can be correlated with images in the metaverse (dreamscape). How would this be accomplished? With research, of course! The chapter will argue that the use of psychecology games as dream analogs allows for the use of Fourier transforms to interrogate correlations between real-time neuro-biological images (EEG, ECG, MEG, fMRI, etc) and images on the game monitor. Even in 1991, research had already confirmed that some human sensory perceptions are based on Fourier mathematics (Talbot, 1991, pp. 27-29; Lashley, 1950, pp 454-82; Collier, Burkhardt, & lin, 1971; Pollen & Tracktenberg, 1972, p. 109; DeValois, Karen K. & DeValois, Russel L., 1979, pp. 483-505; Chopra) and it is well known that *Fourier Transforms* serve as transducers in the digital dreamscape.

If the correlation between dreams and monitor images stands, all the benefits of Jungian dream analysis may accrue to analysis of the media dream. Then what? Ultimately, such analysis could lead to a documentable understanding of unconscious dimensions of collective personae including a definition of contextual purpose—collective *self realization*—as well as discovery of contextual psychological obstacles that obscure and block that state of harmonious self-realization. Analyzing the media dream "successfully" could eventually lead to harmonization of Earth's psyche and the literal introduction of the vaunted "Golden Age."

If correlates can be established between the holographic-dramatic structure of dreams and the holographic-dramatic structure of mediated images, that structure provides a framework for research with which to analyze digital "dream" illusions as if they were "psychic" dream images. In the emergent reality of psychological electronic media, a *Psych*ecology (an ecology of psyche) may be contemplated as a dreamscape. Ample research evidence exists (see below) in order to hypothesize that—like Jungian Psyche—the *metaverse* is holographic and that its computational patterns correspond to archetypal patterns of Psyche. The terms psychecology and metaverse are similar, but the metaverse is manmade and increasingly digital, while the psychecology is primordial. If this is so—and this is why we must refer to myth—archetypal images projected from the psychecology become archetypal representations in the metaverse. Because projections are essentially recursive, the patterns imbedded in archetypal representations can be used like patterns imbedded in dream images. So archetypal patterns can be employed to discover meaning, to diagnose collective mental disorder, and to apply therapy using media feedback targeted to fostering coherent states in contextual subject groups.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

BACKGROUND

It may be stipulated that psychological dynamics of advertising and propaganda have been very successful, but that the profit motive underlying advertising and the power motive underlying propaganda have been largely responsible for fostering the horrors of the nuclear age, the arms race, wars and rumors of wars, the insanity of the Cold War, black ops and covert political action, fear-based dehumanization, genocide, and devaluation of natural life. (Curtis, 2009; Moyers, 1987)

It may be argued that advertising and political agendas have monopolized the technological media. Notwithstanding the power of all media as catalysts for human evolution, the modern media have functioned as prime movers throughout the latter half of the twentieth century and have amply demonstrated their psychological power to alter human values, choices, and behaviors using psychological tools. (Curtis, 2009) Increasingly, sub-optimization, power-lust, and profit have characterized media policy to a catastrophic degree, and political advertising has intruded itself on national governance. Though opinions vary, the recent Supreme Court decision known as Citizens United—a U. S. constitutional case dealing with the regulation of campaign spending by organizations—is having a profoundly detremental influence on U. S. elections because wealthy corporations can now buy unlimited advertising.

For a variety of reasons—well argued by Adam Curtis (2009)—psychological media power has been closely held by the elites of academia, industry, and government, but these entities—apparently preferring the negative Freudian perspective and fear of the collective unconscious (Curtis)—have refused to embrace many of the positive evolutionary neuro-cognitive potentials of mass media. Even the Public Broadcasting System has been cannibalized and exists in a constant state of attack. One consequence of this sub-optimizing stance is that media potentials for learning and healing have neither been supported nor trickled down and been assimilated by the general populace.

There can be no doubt that today incipient human "cognition"—including thoughts, feelings, memories, and perceptions—that expresses itself through language, science, art, law, technology, and all things cultural—has a powerful influence upon the course of planetary evolution. Due to cognitive plasticity, the digital media have a profound and evolving influence on human perceptions of reality, morality, ethics, and learning. Meanwhile, the Internet provides a constant stream of intelligence relative to the nearly incredible advances in computer technology. The MIT Review is especially adept at reviewing new intelligence. Recently, The Review reported that large corporate entities like Face Book have begun expending vast sums in order to better understand cognitive process for neuro-marketing purposes. The rapidly evolving ecosystems associated with personal data is creating an entirely new field of scientific study, (Geril, Philippe, 2014c; Claes, Zane, 2014; Simonite, Tom, 2015d & 2014; 2015; Demers, 2015)

Industrial-governmental interest in "mapping" the brain is becoming pervasive. Based on a twelve year effort involving dozens of research institutes using supercomputers to create a simulation of how the human brain works, the European Union agreed last year to invest more than one billion Euros in the Human Brain Project (HBP). However, a letter from dissenting scientists has charged that the project employs opaque management and the pursuit of goals not widely shared by neuroscientists. The letter says, "We believe the HBP is not a well-conceived or implemented project and that it is ill suited to be the centerpiece of European neuroscience." (Regaaldo, Antonio, 2014; Simonite, Tom, 2014c)

The European Union is not alone. Governments, including those of the United States and China, have all launched large neuroscience projects to study the brain. But having roughly 86 billion neurons and trillions of connections, the brain is so massively complex that there is little consensus on how to

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

study it. (Human brain project, 2014) Related to this interest in cognitive dynamics is Microsoft's heavy investment in Holographic research and the dynamics of machine intelligence. (Simonite, Tom, 2014) A new trajectory of research on deep cognitive dynamics in the metaverse is the modern human mandate.

EVOLUTION OF THE MEDIA AS DREAM CONCEPT

Long ago, in *The Tempest*, Shakespeare observed that, "We are such stuff as dreams are made on." More recently, Sir Arthur Eddington said, "The stuff of the world is mind-stuff," (Foster, 1985, p7-12) Deepak Chopra has argued that, "We are experiencing the final stages of the overthrow of the superstition of materialism," (Chopra, 1995) and Lewis Carroll asked, "Is all our life, then, but a dream? Taken together, these statements sound a keynote not only for study, but for survival in the emergent age of virtual illusions.

The importance of dreams has reverberated from time immemorial, and characteristics of the dreamscape resound in the metaverse that is synonymous with the media dream. Primary correlates between dreams and media are language, metaphor, and symbolism. In this regard, myth provides important precedent not to be ignored by scientists. Not withstanding Fritjof Capra in his *Tau of Physics* (1976), many physicists including Arthur Eddington, James Jeans, Bertrand Russell, Alfred North Whitehead, and J. Robert Oppenheimer have made note of the correlations between quantum physics and myth. (Foster, 1985) According to most myth, the source of dreams is other-dimensional—a "spiritual" dimension that employs symbolic language as a medium between humans and spirit. The Bible Hub quotes John 1:1 (The Holy Bible), "In the beginning was the Word, and the Word was with God, and the Word was God." That's saying a lot, and one commentator, Mathew Henry, accurately clarifies the matter saying, "The plainest reason why the Son of God is called the Word, seems to be, that as our words explain our minds to others, so was the Son of God sent in order to reveal his Father's mind to the world." (Bible Hub, 2015) Without going into great theological commentary and analysis, if they are understood as symbols, words are the quintessential media. This principle is basically true of the trinities in most myths. As documented by Manley T. Hall (2003), Huston Smith (1958), and Joseph Campbell (1968), it is safe to say that all myths are based on symbolic systems, and the "Word" takes many oracular forms including Odinian runes, the Confucian I Ching, the Cabalistic Sephiroth, Pythagorean numbers, Alchemical symbolism, Hermetics, Tarot cards, the Zodiac, ad infinitum.

What the ancients understood as creation by Spirit, moderns may understand as creation with energy. Not coincidentally, the two pillars upon which Jungian theory are based—psychic totality (coherence) and psychic energism—are also consistent with fundamental principles of physics including the unified field, non-locality/synchronicity, and interactivity. Such factors constitute fundamental correlations between physics and myth. As to cognitive correlation among individuals and collective personae, Jung says:

For every piece of conscious life that loses its importance and value—so runs the law—there arises a compensation in the unconscious. We may see in this an analogy to the conservation of energy in the physical world, for our psychic processes have a quantitative aspect also. No psychic value can disappear without being replaced by another of equivalent intensity. This is a rule which finds its pragmatic sanction in the daily practice of the psychotherapist; it is repeatedly verified and never fails. Now the doctor in me refuses point blank to consider the life of a people as something that does not conform to psychological law. A people, in the doctor's eyes, presents only a somewhat more complex picture of psychic life than the individual. (Jung 1933, p 209-210)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

For purposes of digital knowledge engineering, Jungian dreams appear to be holographic. Though opinions differ, as mentioned above, exploration of the holographic model has been comprehensively researched by Michael Talbot. (Talbot, 1991) If the basic principles of Psyche are unity and energy (Jacoby, 1973, p. 2), and if dreams are the archetypal representations of archetypal energy and information patterns, the "conversation" between archetypal patterns and projected images as archetypal representations may be understood as the basic form of mediation. Relative to discovery of the meaning embedded in dreams, that meaning may be divulged—as in any conversation—with insight and an understanding of the contextual language—the Word. If the media-sphere may be understood as a holographic conversation in contextual-symbolic dream language between unconscious and conscious dimensions, archetypal patterns of the collective unconscious may be understood to correspond with unconscious algorithmic patterns of digital software articulated as Fourier Transforms. The structure and purpose of projected images in the metaverse can then be correlated with dreams that Jung defined as archetypal representations.

Like film projections, archetypal representations are recursively structured and constitute a symbolic language between conscious and unconscious dimensions of Psyche. In other words, the dream is an accurate projection of unconscious multifunctional patterns. The term multi-functional refers to the four classical human functions (physical, emotional, mental, spiritual) that Jung called Sensing, Feeling, Thinking, and Intuiting. Because they consist of frequencies of energy, these functions cannot actually be separated. Human functions can be correlated with Aristotle's dramatic functions. The psychological functions constantly influence one another in ways consistent with Aristotle's definition of "dramatic unities," and these dramatic structures and unities can be used as a fundamental point of reference for cognitive research. For example, George Lakoff explains the chemical processes in the nervous system relative to semantic "scripts". (Lakoff, 2008) Lakoff cites Roger Shank and Robert Abelson (Lakoff, 2008, p. 249) in his paraphrase, "Complex narratives—the kinds we find in everyone's life story, as well as in fairy tales, novels, and dramas—are made up of smaller narratives with very simple structures." Lakoff says:

The neural circuitry needed to create frame structures is relatively simple [neural binding circuitry, neural signatures, event structures] and so frames tend to structure a huge amount of our thought…Each frame has roles (like a cast of characters), relations between the roles, and scenarios carried out by those playing the roles." (Lakoff, 2008, p. 22; Goffman, 1974) There is a protagonist, the person whose point of view is being taken. The events are good and bad things that happen. And there are appropriate emotions that fit certain kinds of events in the scenarios. (Lakoff, 2008, p. 23)

Citing Goffman, Dr. Lakoff also observes that words are all defined relative to conceptual frames: "Groups of related words, called 'semantic fields,' are defined with respect to the same frame. Thus words like 'cost,' sell,' 'goods,' price,' buy,' and so on are defined with respect to a single frame," (Lakoff, 2008, p 22; Goffman) and the roles of Buyer, Seller, Goods, and Money, form a narrative field context for the frame. Sounding remarkably like Jung who observed that dreams have dramatic action that can meaningfully be broken down into the elements of a Greek play (Jacoby, 1973, p. 83), both Jung and Lakoff understand the importance of the functional unities of cognition.

Symbolic patterns in dreams can be analyzed to discover their message—their purpose—so perhaps, digital images in the media-sphere can be analyzed like dream images in order to diagnose and treat collective mental disorder. Diagnosis would depend on "engineering" the knowledge management software according to principles of cognitive framing, and treatment would involve multifunctional

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

mediated biofeedback based on extensive research. By employing the compensational processes Jung calls the Amplification Method, a Jungian psychiatrist can mentor a patient by providing feedback as to the meaning of dream symbols. If successful, the process leads to meaningful (healing/coherent) insight on the part of the patient. In cognitive terms, unconscious cognitive frames can be adjusted into more coherent frames by providing media feedback. The psychological feedback/association process has been validated by advertising and propaganda, but the process of Jungian feedback is much different, more sophisticated, and differently motivated.

Adjusting cognitive frames in the collective unconscious of targeted personae with feedback in the form of interactive media content could foster meaningful insight and the achievement of coherent states among contextual collectives. The Jungian process of compensation is essentially a process of metaphorical extension, but it is based upon intense empathy with the patient on the part of the psychiatrist. The question of whether such empathy can be incorporated into digital simulations must be addressed with research, and such research lies within the realm of dramatic unities, their integrated archetypal functionality, and their degree of coherence. Because the psychiatrist must have extensive and empathetic knowledge about the patient, s/he is enabled to provide feedback to the dreamer relative to the dreamer's contextual psychic state. Feedback takes the form of suggestions about the meaning of symbolic patterns in the patient's dreams. When a patient (or learner) has insight as to the meaning of a dream, insight becomes learning and healing. According to this same process, media content targeted to contextual subject groups could foster a collective state of coherence (meaningful insight).

The dream speaks in images, and gives expression to instincts, that are derived from the most primitive levels of nature. Consciousness all too easily departs from the law of nature; but it can be brought again into harmony with the latter by the assimilation of unconscious contents. By fostering this [amplification] process we lead the patient to the rediscovery of the law of his own being. …The way of successive assimilations reaches far beyond the curative results that specifically concern the doctor. It leads in the end to that distant goal (which may perhaps have been the first urge to life), the bringing into reality of the whole human being—that is, individuation. (Jung 1933, p 26)

A fundamental assumption in the chapter is that the *Psych*ecology and--as archetypal representation—the media-sphere function in accordance with the rules of dream dynamics observed by Carl Jung. Directly relevant to this media dream-likeness is that humans have finally verified that so-called "material reality" is—in fact—illusion. Therefore, the universal axiom of myth and philosophy has also become scientific fact verified by quantum physics, and echoed by cognitive neuroscience. So, hypothetically, the *Media Dream* can be researched from the perspective of its curative potential. (Schafer & Yu; 2011) It may be understood that precisely like Jungian dreams, mediated digital dreams arise from the collective *unconscious* that is directly correlated with the digital unconscious and ultimately exists as a unified field of energy and information where human functions are unified according to principles of dramatic unity. Both cognitive and digital patterns take the form of frequencies of energetic force that, according to cognitive-psychiatric research, clothe themselves in linguistic-dramatic images and are projected as "dreams" (illusions).

During the past twenty years, research in neurological plasticity has reached a degree of agreement about the existence of this narrative-metaphorical structure of Psyche by studying the electro-chemical dynamics of the nervous system, integrating them with principles of physics, and researching neurobiological processes with technologies such as EEG, ECG, fMRI, MEG, etc. Drama is a more accurate

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

term than narrative for cognitive-linguistic structure because drama is electro-chemically interactive; i.e., actors on the stage and people in the audience perceive drama according to a fourfold matrix of functional components that have always been used to gauge human behavior and experience. To reiterate, Carl Jung called these functions reason, emotion, sensing, and intuition. Jung observed that dreams have dramatic structure, and modern cognitive researchers like George Lakoff have observed that the cognitive unconscious has dramatic structure to which we all refer our projected realities. The roles we play in life have their origin at unconscious levels and provide us with a template for understanding ourselves, our context, our exposition (statement of the problem), and our *lysis* (resolution of the problem). In other words, we live out our unconscious narratives, and—for example—we use the same part of our brains to see, imagine we are seeing, remember seeing, and understand language about seeing. (Lakoff, 2008, p. 39) The four functions are integrated in drama, and this dynamic explains why—when people have vicarious experience such as dreams, film, and video games—they experience emotions appropriate to the film/game story and access unconscious moral patterns. (Lakoff, 2008, p. 63) This effect is usually understood as immersion or suspension of disbelief.

While research on this emergent reality is yet in its infancy, it is certainly apparent that it constitutes a paradigm shift. Therefore, a better understanding of psychological health, morality, ethics, and laws appropriate to the shift become the human mandate. Researching the reality of the media dream according to dream dynamics becomes the human imperative. Though the most recent cognitive research seems to be entirely consistent with the principles of psyche and dream analysis established by Carl Jung, his research opus is more comprehensive and has provided a universal (and often unattributed) framework for computer engineering—memory—and analysis of cognitive research (neural networks, connectionism, holographic computing, etc.). As a recommendation and suggestion of future prospects, this paper addresses the Media Dream primarily from the standpoint of its potential for discovery of unconscious meaning at psychological levels of cultural collectives. Such meaning resides in the imbedded symbolism of mediated images.

Contextual Dream Patterns in War and Entertainment

A significant trend in current virtual reality research is sponsored by the military, and battlefield training courses are intended to be as realistic as possible—a clear example of the paradigm shift into the reality of the metaverse. From the cognitive perspective addressed above, these battle exercises are perceived by trainees as real (Reeves & Nass). Being "immersive" they also impact trainees at unconscious functional cognitive levels. In other words, the contextual group of trainees experiences virtual training exercises as real experience, and this experience changes unconscious "archetypal" cognitive frames at all functional levels including values. Presently, these training "dreams" are being programmed precisely for these reasons. The military wants its trainees to have "warfigher" values.

Warfighters

Just one example of this may be found online at the Naval Office of Research site as reported by EUROSIS, "ARLINGTON, Va.—Marines were able to turn a lush golf course into a hostile battleground complete with tanks, mortar fire and smoke. (Crawford, Katherine H., 2015) This "dream" was made possible by a newly developed augmented reality technology. "The Augmented Immersive Team Trainer (AITT) moved one step closer to its ultimate goal of giving warfighters glasses that can be worn to en-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

hance augmented reality training scenarios." (Crawford) The reporter makes a distinction between the terms "augmented reality" and "virtual reality". The former refers to "a live environment onto which virtual images are superimposed. Think of watching a football game on television and seeing the yellow first-down line added to the screen you're watching." (Crawford) A "virtual reality" is different because it "is a wholly computer-generated environment in which users immerse themselves." However, according to Peter Squire, a program officer with ONR's Expeditionary Maneuver Warfare and Combating Terrorism Department, says that the AITT system, "which can support a wide array of live, virtual and constructive training scenarios, is 'only limited by the imaginations of those using it,'" (Crawford) For the purposes of this chapter, this is a clear sign of the trend that employs virtual reality to alter unconscious cognitive frames.

Virtual Humans

Another example of the trend to employ contextual dreams to alter experience-based cognitive frames at unconscious levels is the following. Online, we find that: "Research and technology are essential for providing the best capabilities to our warfighters." (Bellman, Orli, 2015) An online article, Researchers Create Virtual Humans, quotes Dr. Laurel Allender, director of the ARL Human Research and Engineering Directorate as saying, "This is especially so for the immersive and live-training environments we are developing to achieve squad overmatch and to optimize Soldier performance, both mentally and physically." The Army and USC are, "exploring the potential of developing a flexible multi-modal human-robot dialogue that includes natural language, along with text, images and video processing." (Bellman)

If this online article is analyzed as part of a more general dream sequence, it may disclose unconscious meaning. The more general sequence has to do with the scientific passion for virtual realities and robotics. We should be asking the question, "Where will this passion lead?" With the emergence of the brave new world of illusion being created by technological media, "virtual humans" are just one form of the denizens projected to inhabit the Metaverse. In the Metaverse, "real" people are expected to express themselves increasingly as avatars—in the same way as game players project their persona into a video game protagonist. The implicit concept underlying the Metaverse is that real life is a metaphor and that virtual life is reality. The term is being used to describe the concept of a future iteration of the internet, made up of persistent, shared, 3D virtual spaces linked into a perceived virtual universe—a virtual reality-based successor to the Internet. (Smart, J.M., Cascio, J. and Paffendorf, J., 2007) Whether "reality" is metaphorical or virtual, both are illusions, and the symbolism in the dreams can promote meaningful insight.

Artificial Intelligence and Robotics

Virtual humans and avatars exemplify the explosion of artificial intelligence applications which are enabling an incipient virtual age. Artificial intelligence algorithms underlie research on *robotics*, so Artificial intelligence and robotics are essentially synonymous. As suggested above, much AI research is being sponsored by governments—with research done by universities—for purposes of waging war. A conclusion is implicit; Just as the video game genre called *First-Person-Shooters,* in which the game player's avatar is first person, have achieved overwhelming popularity in video games, virtual realities based on warfare and "warfighters" have become a priority aspect of media age illusion. Because a

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

basic assumption of this chapter is that all experience, including virtual experience, alters the cognitive patterns in the personal and collective unconscious, the following example of the explosive trajectory of research in robotics becomes a prospective dream that is susceptible to analysis.

The US Defense Advanced Research Projects Agency (DARPA) has launched a new program to help small businesses, and individuals who have not previously worked with the federal government, develop new robotics capabilities. Dubbed Robotics Fast Track (RFT), the program aims to enable rapid, cost-effective development of new robotics capabilities designed to respond to, and even anticipate, quickly emerging warfighter requirements. (Geril, 2015a)

More ominous than the DARPA project is the following report from The U.S. Army that computer simulations improve lethality by understand the effects different caliber ammunition will have on targets. The research center at Picattiny Arsenal in New Jersey, "handles nearly 90 percent of the Army's lethality projects and all conventional ammunition for joint warfighters." (Kowal, Eric, 2015) The research center works with organizations like the U.S. Army Research Laboratory and the Survivability/Lethality Analysis Directorate at Aberdeen Proving Ground, Maryland.

Kowal points out that though large caliber system analysis has taken advantage of computational simulation tools for years, in order to improve consistency and performance with newly released smaller caliber enhanced performance rounds (EPR) a combination of modeling, simulation and testing was employed. "The use of newer, faster computers made critical portions of that effort possible. Some data was statistical, some physical and some probabilistic, but all done using physics-based models to simulate ballistics on soft targets" (Kowal). An "up-side" of simulating soft-target wounds might be that many who study survivability and lethality such as program executive officers and program managers, users, testers, the Army's independent evaluator and other "customers" would not see this as ominous because the research supports milestone decisions that might save lives. Another aspect of the "up-side" is that prototypes can cost millions of dollars a piece. "Modeling and simulation can save money by avoiding costly prototypes and testing" (Kowal).

Obviously, DARPA is not the only government agency interested in warfighters. "Several US government agencies including the army, navy, the department of homeland security, and the National Aeronautics and Space Administration have expressed interest in validating successful RFT projects and performers for use in their operations" (Geril, 2015a).

Research on robotics is not all about warfare. The Europe 2020 Initiative: Digital Agenda for Europe, "Scientists have created underwater robot swarms that function like schools of fish, exchanging information to monitor the environment, searching, maintaining, exploring and harvesting resources in underwater habitats." (Cocoro, anon., 2015) In ten experimental demonstrations, autonomous robots interact with each other and exchange information, resulting in a cognitive system that is aware of its environment. The European Union supported the COCORO project that explored and developed collective cognition in autonomous robots. COCORO's scientists used nature as a model. For example, observing how honeybees cluster, helped scientists develop the BEECLUST algorithm that they used to aggregate robots at a specific location. "Robot swarms not only look like schools of fish, they behave like them" Cocoro, par. 2).

Increasingly, the research on "robotics" involves the use of algorithms based on cognitive function. (Simonite, Tom, 2015b). As Neil Jacobstein, head of the Singularity University told the BBC:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Artificial intelligence agents are already involved in every aspect of our lives - they keep our inboxes free of spam, they help us make our web transactions, they fly our planes and if Google gets its way will also soon drive our cars for us. "AI's are embedded in the fabric of our everyday lives. They are used in medicine, in law, in design and throughout automotive industry." And each day the algorithms that power away, making decisions behind the scenes, are getting smarter. It means that one of the biggest quests of the modern world - the search to make machines as intelligent as humans - could be getting tantalizingly close. (Wakefield, Jane, 2014)

In another recent report, the MIT Technology Review reports on an algorithm that can help robots heal themselves without human intervention. Researchers say that:

The robots learn with an "evolutionary algorithm.' This means that it will run repeated trial and error steps—building on a list it created before being deployed that details the things it can do and how valuable each one is—to determine a new way to get around. It's like a simplified version of what humans do: if you sprain an ankle, you use memory and experimentation to figure out how to hobble around in the most efficient and least painful way. (Garling, 2015, par. 6)

Though researchers' examples such as the Mars Rover or search and rescue robots are still far from real-world scenarios such as fighting fires or rescuing people, they are a significant aspect of the metaverse. Researchers Jean-Baptiste Mouret and Antoine Cully say these learning algorithms will soon be tested outside the lab on bigger robots. Perhaps they will be tested on virtual humans; this raises considerable ontological, epistemological, and ethical questions relative to the reality of the metaverse. "While watching the six-legged robot figure out a new method of movement inevitably brings about visions of the Terminator and evil robots, Mouret says researchers have plenty of time to build in proper safeguards to prevent robots from implementing behaviors that could harm people" (Garling).

Thinking of contextual collectives in terms of target groups is more a hop than a quantum leap. The use of focus groups in advertising to tap buy buttons in target groups is longstanding and very successful procedure. Now, AI research on demographics—from virtual humans to autonomous robot swarms—employs smart algorithms based on cognitive mapping. The Virtual Worlds Report for March 2015 defines "contextual" demographics, attitudes, and preferences: "The aim of this study was to enhance our understanding of demographics, attitudes, activities and play preferences across a variety of non-game, social virtual worlds, also referred to as metaverses" (Geril, 2015b).

Entertainment

In another framework of the psychological media sphere, the entertainment industry has always been about illusion addressed to contextual collectives. From musicals to war films, drama is based on appeal to contextual collectives, and these collectives are multi-faceted. For example, films of the WWII era were propagandistic and appealed to entire cultures as well as to subsidiary collectives. It may be stipulated that entertainment—to use a vastly general term—has always been a powerful source for acculturation. The idea that the entertainment industry has enormous influence on culture at large: cultural values, individual & cultural choices, and behaviors, need not be extensively documented. (Curtis, 2009)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Entertainment and the Internet are becoming ever more closely linked as the age of illusion unfolds in a kaleidoscopic phantasmagoria of wireless applications. In answer to the question, "Where is the entertainment industry going?" we discover that Star Trek has provided the pre-eminent paradigm, in the form of the 'holodeck', the virtual reality facility on the Starship Enterprise. The holodeck was a box of a room where crewmembers could immerse themselves in an array of virtual illusions, like distant planets, and interact with aliens without ever leaving their spacecraft. However, mapping consumer demand against industrial innovation is questing beyond the holodeck:

Producing immersive entertainment experiences will require a whole suite of tools and underpinning technologies: in computer vision, positional and spatial audio, smart content creation, compression and real time rendering. Other advances will be needed in capabilities outside the range of even the leading audio-visual production studios. These will include 3-D display, augmented reality, holography, surround sound and haptic technologies. (Geril, 2015b)

Whether a virtual reality is a dream, a memory, a fantasy, a film, television, the Internet, or a video game, the research-based understanding of virtual-cognitive reality is that it has dramatic structure—even at quantum levels. According to Jung and Lakoff, the cognitive structures in unconscious dimensions are *projected* into so-called "real-life" contexts. (Jacobi, 1973; Lakoff, 2008; Goffman, Erving, 1974; Schafer & Yu, 2011) As already mentioned, Jung called the structural patterns of the unconscious *archetypes*, and Lakoff calls them *cognitive frames*. Cognitive-level processes that explain precisely *why* dramatic structure is important for an understanding of virtual illusion—of both conscious and unconscious (immersive) dimensions of *reality*— are not well-known except in fields of cognitive research. Nevertheless, researchers have been exploring linguistic issues of the unconscious for decades. Whether understood from cognitive levels or not, in diverse fields ranging from law enforcement to psychiatry, the value of narrative is practical and well known. Always interested in the poetic dimensions of drama, now the entertainment industry—especially the game industry—is becoming interested in cognitive dimensions as well. For example:

21st century storytelling will require not just enabling technologies, but also story tellers: professionals with expertise in interaction, human experience and story design to create content and experiences to be developed and consumed on these platforms. This in turn may necessitate new organizational structures, management systems and business models to enable innovations and the merging of technologies across films, games and media industries. (Boyd, Frank, 2015)

Recursive Holography and the Metaverse

From these brief references to massive cultural trends, contemplating the media illusion as a dreamscape and using dream analysis to psychoanalyze collective personae for purposes of harmonization and healing does not seem entirely crazy. An increasingly popular term for the media dream is the metaverse which is the global holographic projection of interactive images including Internet, film, video games, television and all the technological paraphernalia presently available. Our fundamental hypothesis is that the digital software (AI) sources of mediated *dream images* is analogous to the archetypal energy/information that is the cognitive source of dreams. Accordingly, the holographic projections of computer engineered archetypes (programs) are analogous to what Carl Jung calls "archetypal representations."

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

As discussed above, an important aspect of such representations—especially images—is that they are essentially recursive; i.e. they have the same dramatic structure as "archetypes" that inhabit the personal and collective unconscious. Though research is necessary to verify, the translation from archetype to archetypal representation appears to be based on recursive frequency analyzers known as Fourier transforms (Talbot, 1991, p. 27). It has been verified that human sensory mechanisms are based on Fourier mathematics (Talbot, p. 27; DeValois, Karen, DeValois, Russell L., Yund W. W., 1979 pp. 483-505), but the specific connection to archetypal-quantum structure remains to be validated.

The term "recursive" may be understood in terms of the "fractal" forms of physics, or in terms of the Gestalt perceptual principle of "Invariance". Invariance is the property of perception whereby simple geometrical objects are recognized independent of rotation, translation, and scale. Another psychological aspect of recursion includes such perceptions defined by Gestaltists as "Multistability" or multistable perception (the tendency of ambiguous perceptual experiences to pop back and forth unstably between two or more alternative interpretations), and "Reification" (the *constructive* or *generative* aspect of perception, by which the experienced percept contains more explicit spatial information than the sensory stimulus on which it is based).

This recursive dynamic of perception allows for the computational modeling of what appears to be randomly fluctuating symbols and pixilated images in both dreams and computer monitors. In other words, a quantum flux of archetypes that inhabit the personal and collective unconsciousness as defined by Carl Jung are projected as recursive archetypal representations in a neurobiological-holographic media-sphere. (Schafer & Yu, 2011; Schafer, 2011; Schafer, 2012)

Research on issues of the relationship between the holographic cosmos and the holonomic (holographic) brain is longstanding, and it's early proponents (Penfield, Wilder, 1975; Lashley, Karl, 1950, pp.454-82; Pribram, Karl, 1977, p. 75; Pribram, Karl, 1977, p. 123 & 169; Goleman, Daniel, 1971, pp. 71-72; Collier, J., Burckhardt, C. B.,& lin, L. H., 1971; van Heerden, Pieter, 1970, pp. 410-11; Pietsch, Paul, 1981, p. 78; Pollen, Daniel A. & Tractenberg, Michael C., 1972, p. 109; DeValois, Karen K., DeValois, Russell L., & Yund, W.W., 11979, pp. 483-505; Dossey, Larry, 1982, pp. 108-9; Restak, Richard, 1981, p. 19; Restak, Richard, 1979, p. 253; et al.) are well documented in Michael Talbot's *The Holographic Universe* (1991)

Holographic Coherence

The most recent research evidence that verifies this connection between quantum dimensions and their projections as "consciousness" is the discovery by Stuart Hameroff and Sir Roger Penrose of quantum vibrations in "microtubules" inside brain neurons. "They suggest that EEG rhythms (brain waves) also derive from deeper level microtubule vibrations, and that from a practical standpoint, treating brain microtubule vibrations could benefit a host of mental, neurological, and cognitive conditions." (Hameroff & Penrose, 2014) These researchers based their discovery on two decades of previous research on "orchestrated objective reduction" (Orch Or) which suggests that quantum vibrational computations (microtubules) are orchestrated by synaptic inputs and memory stored in microtubules. (Hameroff & Penrose, 2014) The term "orchestration" suggests the potential to harmonize conscious patterns with archetypal patterns according to the principle of coherence. Coherence is a measurable harmonious neurobiological state that has been studied extensively by the Heartmath Institute. (McCraty, Rollin & Childre, Doc (2010); McCraty, Deyhle, & Childre (2012).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Coherent states appear to have the same characteristics as Carl Jung's "Compensational" states and "Flow". (Csíkszentmihályi, 1990) Compensation is the fundamental Jungian principle that makes discovery of unconscious meaning possible using the process of symbolic analysis on projected archetypal energy patterns in dream images. As images, dreams and computer graphics have structural correlates. The entirety of Carl Jung's psychiatric theory and practice is based on symbolic images in dreams, so interpretation of the underlying meaning in dream-computer images measured in terms of coherence has identical potential.

Correlated Images

Dream images and computer monitor images are directly correlated as to structure, perception, and leaning-healing potential. These correspondences are a primary component for qualifying psychecology games as a dream analog. (Schafer, 2011; Schafer, 2012) Therefore, psychecology games can be used as a cognitive research instrument in conjunction with MEG, EEG, ECG, fMRI, etc. for measuring real-time neurobiological activity in subject game players. (Schafer & Yu, 2011; Schafer, 2011; Schafer, 2012; Trafton, Anne, 2014) Research methodologies must be contextualized to various media and to myriads of demographics, but the dynamics of dream analysis remain constant. Though the media dream includes multi-functional sensory data that is integrated at levels of the unconscious, this chapter emphasizes the primarily visual symbolism in images.

Jung is not the only observer to say that the changing human perceptions of reality—what he calls the Spirit of the Age (Jung, 1933, pp. 175,196-220)—have been thrust upon us so rapidly that throughout the twentieth century it has disoriented and traumatized human collectives throughout the world. At base, the changes derive from discoveries in the realms of physics, psychology, and computer science. It is a well established fact that the psychological dynamics and structure of advertising and propaganda entail utilizing unconscious demographic patterns to sell their "products". This chapter develops the hypothesis that because the discoveries in these sciences have been incorporated into the holographic, psychological metaverse, its contextual projections can be analyzed like dreams.

THE MEDIA DREAM: REALITY AS ILLUSION

Researching the nature and dynamics of this illusory paradox of human beingness, human creativity, and human responsibility is critical to planetary survival. The arguments and details of why human existence—specifically material existence—is an illusion and a paradox is an issue of physics and psychology that has been well developed by many of humanity's ablest exemplars. Our argument emphasizes Jungian dream parameters as a basis for addressing the importance of the media dream as illusion that appears real. Based on the two pillars of Jungian psychiatry—psychic totality and psychic energy—this chapter postulates correlations among the unified fields of physics, psyche, and the technological media-sphere. (Schafer & Yu, 2011; Schafer, 2012) Though ongoing research constantly adds to the verification of the hypothesis, it is already an authentic premise in diversified scientific fields and is being addressed by an increasingly evolutionary neuroscience. Moreover, scientists and scholars including Eddington, Jeans, Whitehead, Capra, Radins, Oppenheimer, and many others have observed the connections that exist between myth and physics and which are central to Jungian principles of symbolic-metaphorical dream analysis and the process of Compensation through Amplification. Many have recognized the

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

correlations between physics and psychology. Perhaps the most prominent of these is the scholar of myth, Joseph Campbell who asserted that, "Myth is psychology." Campbell was referring specifically to Jungian psychology.

Functional Interactivity

Based in physics, the most promising candidate for dream dynamics and structure is holography—a science that is just coming into its own. (Talbot, 1991) Accordingly, because of their corresponding structure and dynamics, it may be argued that both dreams and media dreams are holographic. Both realities share the principles of energy patterns, wholeness, linguistic structure, functionality, and—most importantly—interactivity. (Schafer & Yu, 2011) The entire framework of Jungian analysis and treatment is founded on his observation that dreams are interactive. (Jacoby, 1973) In other words, the symbolism and context of dreams change according to a patient's evolving need for psychic meaning. These changing patterns resulting from interactive experience manifest as bio-energetic phenomena in the nervous system. We will detail this research in a later chapter. For now, this same bio-energetic interactivity can be imaged and measured in real time as a player makes choices leading to a game-win while playing drama-based video games.

Because non-local bio-energetic processes are now being applied to research using a plethora of electronic devices that contribute to understanding the cognitive connection to holographic media, the collective unconscious is being impacted with the digital unconscious generated by complex algorithms. The fundamentally interactive nature of dreams has already become nearly pervasive—though not recognized—in the media dream. In ways that are correlated, increasing interactivity in media domains allows people to use the Internet, television, smart phones, and a growing array of apps to access contextual meaning from almost anywhere in the virtual domain of the holographic media dream—action at a distance. The electro-chemical neural processes associated with this pervasive interactivity may be calculated mathematically employing sophisticated engineering sets and components such as finite element analysis, fluid dynamics, Discreet Fourier Transforms, efficient algorithms of fast Fourier Transforms, and a variety of other calculations for articulating the paradigm shift. (Schafer, 2011) To reiterate, for decades, scientists from various fields have speculated that the human brain and the universe itself are holographic. (Talbot, 1991; Bohm, 1951; Mishlov, 1998) So holography becomes an increasingly relevant metaphorical model that has the potential to verify the media dream hypothesis.

Deepak Chopra has explained reality is "a socially programmed hypnosis," and promised that ours is an age of the "overthrow of the superstition of materialism." (Chopra, 1995) In a scientific environment where Wilder Penfield has searched for the "thinker of the thought," and others have researched "cellular memory," we have made important strides in the aggressive study of cognition and consciousness, (Chopra, 1995), but we have not yet accepted the ramifications or the psychological impact of the media dream as a primary aspect of human reality. Our research has studied this premise from a variety of angles including the enhancement of psychological coherence in contextual collectives (Schafer, 2011), serious games for articulating the paradigm shift (Schafer, 2011), ethical design of social influence marketing (Schafer & Palamedes, 2011), the use of drama-based video games as a research instrument for creation of meaningful media (Schafer & Yu, 2011), and optimizing cognitive coherence and learning in contextual collectives. (Schafer & Yu, 2011) The arguments in these publications and many conference presentations assert the concept that the technological media-sphere is dynamically and structurally analogous with the unified field of energy and information known to physics and has virtually the same dynamics as dreams defined by the Analytical Psychology of Carl Jung.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

As mentioned above, this analog extends to the structure and dynamics of psychecology games. Defining target groups with subjective data for advertising purposes has been very successful. Using more sophisticated personality profiling like Myers-Briggs and Big 5 profiles to define game player functionality would contribute greater scientific authenticity to the process of defining target (contextual) groups. During drama-based video game play, player choice patterns could be stored and studied using sophisticated headsets or virtual reality technologies to track electro-magnetic and chemical activity in the nervous system. Such research will then contribute to understanding of distributive functions and unconscious motivations in collective personae. Correlating neuro-linguistic dynamics experienced by game players with neuro-linguistic dynamics experienced by users of the media dream could lead to a much better understanding of the "positive" uses of media content, media policy, and media education. This would provide a key that could open the portal of opportunity. Opening the door and crossing the threshold of opportunity would constitute a critical initiation in the human perception and understanding of what is real. We would begin to appreciate the reality of illusion as portrayed in the Media Dream. This would be our crisis of departure from the superstition of materialism. Now we have an opportunity to address the psychological crisis of the media age before we lose track of reality and become immersed in psychotic reaction to illusion—the symbols and symptoms of which are now rampant in the Media Dream. (Schafer, 2012)

According to the dynamics of holography, our technological media provide insight—like a hall of mirrors—to the collective unconscious, while video games have the potential to provide highly specific access to neurobiological function. If neural processes can be correlated with archetypal states of the cognitive unconscious where archetypes form a magnetic field and energy center recursive with the projection of images, a wealth of Jungian principle and research will become available for timely use. "The dream is a statement, uninfluenced by consciousness, expressing the dreamer's inner truth and 'reality as it really is.'" (Jacoby, 1973) "Thus for Jung the manifest content of a dream is not a façade but a fact which always reveals what the unconscious has to say about the situation and which always says exactly what the unconscious means." (Jacoby, 1973) Therefore, the imbedded symbolism in images can be analyzed—as it is in Jungian dream analysis—to access contextual meaning that resides in both the cognitive unconscious and the digital unconscious. (Schafer, 2011) Such access to the personal and collective unconscious can be employed to reinforce cognitive coherence. (Schafer, 2012)

In today's reality of statistics and target groups, it may be stipulated that—at least metaphorically—human groups have "personalities," (ethnic, religious, political, corporate) that are defined by normative-conceptual cognitive structures imbedded in the energy and information fields of the collective unconscious. Jung discusses cognitive structures in terms of archetypal patterns, and George Lakoff discusses them in terms of "neural recruitment" which is a process for strengthening neural synapses. (Lakoff, 2008) Employing the process of Jungian Amplification (symbolic-metaphorical extension), these energy/information patterns of the unconscious are susceptible to harmonization (measured in terms of heart rate variability as a signifier of coherence).

By all indications, Jungian Compensation is synonymous with "coherence". It is our contention that with the application of Jungian principles to the Media Dream, coherent psychological states can be induced in collective personae by pre-measuring how target groups can be predisposed to functionally coherent states. McCraty and Childre discuss the nature of coherence, how it can be measured and reinforced with biofeedback, and applied to harmonization of psyche in all its functions—physical, emotional, rational, and spiritual. "In globally coherent systems such as human beings, there is an incredible amount of activity at every level of magnification or scale that spans more than two-thirds of

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

the 73 known octaves of the electromagnetic spectrum." (McCraty & Childre, 2010) The functions cited by McCraty and Childre correspond to the Jungian functions of psyche—sensing, feeling, thinking, and intuiting. It is these functions that can be integrated into experimental design in order to foster cognitive coherence. As cited by Michael Talbot in *The Holographic Universe*, researchers (DeValois, 1979 pp. 483-505) have discovered that the brain uses Fourier mathematics to translate patterns of the sensory function. For example, the visual cortex responds to frequencies of various wave forms. (Talbot, p. 28) Subsequently, their research has been expanded to confirm that the other sensory functions are based on Fourier mathematics (Talbot, 1991). Moreover, though much more research is necessary to expand our understanding of the dynamic, Lakoff confirms that the functions are integrated by the process of narrative and metaphor. "There is a protagonist, the person whose point of view is being taken. The events are good and bad things that happen. And *there are appropriate emotions* that fit certain kinds of events in the scenarios (italics mine)." (Lakoff, 2008, p. 23)

Characteristics of compensation and coherence are also consistent with characteristics of the "Flow" phenomenon researched by Mihály Csíkszentmihályi. "Flow is the mental state of operation in which the person is fully immersed in what he or she is doing by a feeling of energized focus, full involvement, and success in the process of the activity." (Csíkszentmihályi, 1990) Components of flow include clear goals, concentration and focus, diminished self-consciousness, loss of the sense of time passage, and a sense of personal control and intrinsically rewarding action are aspects of immersion in games and—perhaps—suggestions of contact with an altered cognitive state.

As this dynamic of functionally coherent, compensational, flow becomes more generally accepted—making it a global fact of everyday reality—the human reality will shift in benign ways. Emergent changes in human perceptions of reality and human responsibility will evoke the imperative to generate coherent patterns relative to morality, law, responsibility, privacy, and an entire gamut of social interactions that are presently obscure. Gradually, as we achieve this new understanding of the "real", we will mark a breakthrough in the human sense of responsibility for the direct power humans wield with the energy of their thought and desires. Certain oft-reiterated truths will be clarified: "Thoughts are things." (Plato); "As a man *thinketh* in his heart, so is he." (Christ); "A man is what he thinks." (Buddha); "Our thoughts are epochs in our lives." (Thoreau); "Nurture your minds with great thoughts. To believe in the heroic makes heroes." (Disraeli); "The greatest events of an age are its best thoughts. Thought finds its way into action." (Boice). Greater understanding of the media dream may mark a crossing of the initiatory threshold comparable to the discovery of fire.

From these brief references to massive cultural trends, contemplating the media illusion as a dreamscape and using dream analysis to psychoanalyze collective personae for purposes of harmonization and healing does not seem entirely crazy. A popular term for the media dream is the metaverse which is the global holographic projection of interactive images including Internet, film, video games, television and all the technological paraphernalia presently available. Our fundamental hypothesis is that the digital software, AI, sources of mediated *dream images* is analogous to the archetypal energy/information that is the cognitive source of dreams. Accordingly, the holographic projections of computer engineered archetypes (programs) are analogous to what Carl Jung calls "archetypal representations." An important aspect of such representations—especially images—is that they are essentially recursive; i.e. they have the same dramatic structure as "archetypes" that inhabit the personal and collective unconscious. Though research is necessary to verify, the translation from archetype to archetypal representation appears to be based on recursive frequency analyzers known as Fourier transforms (Talbot, 1991, p. 27).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

It has been verified that human sensory mechanisms are based on Fourier mathematics (Talbot, p. 27; DeValois, Karen, DeValois, Russell L., Yund W. W., 1979 pp.483-505), but the specific connection to archetypal-quantum structure remains to be validated.

The term "recursive" may be understood in terms of the "fractal" forms of physics, or in terms of the Gestalt perceptual principle of "Invariance". Invariance is the property of perception whereby simple geometrical objects are recognized independent of rotation, translation, and scale. Another psychological aspect of recursion includes such perceptions defined by Gestaltists as "Multistability" or multistable perception (the tendency of ambiguous perceptual experiences to pop back and forth unstably between two or more alternative interpretations), and "Reification" (the *constructive* or *generative* aspect of perception, by which the experienced percept contains more explicit spatial information than the sensory stimulus on which it is based).

This recursive dynamic of perception allows for the computational modeling of what appears to be randomly fluctuating symbols and pixilated images in both dreams and computer monitors. In other words, a quantum flux of archetypes that inhabit the personal and collective unconsciousness as defined by Carl Jung are projected as recursive archetypal representations in a neurobiological-holographic media-sphere. (Schafer & Yu, 2011; Schafer, 2011; Schafer, 2012)

Research on issues of the relationship between the holographic cosmos and the holonomic (holographic) brain is longstanding, and it's early proponents (Penfield, Wilder, 1975; Lashley, Karl, 1950, pp.454-82; Pribram, Karl, 1977, p. 75; Pribram, Karl, 1977, p. 123 & 169; Goleman, Daniel, 1971, pp. 71-72; Collier, J., Burckhardt, C. B.,& lin, L. H., 1971; van Heerden, Pieter, 1970, pp. 410-11; Pietsch, Paul, 1981, p. 78; Pollen, Daniel A. & Tractenberg, Michael C., 1972, p. 109; DeValois, Karen K., DeValois, Russell L., & Yund, W.W., 11979, pp. 483-505; Dossey, Larry, 1982, pp. 108-9; Restak, Richard, 1981, p. 19; Restak, Richard, 1979, p. 253; et al.) are well documented in Michael Talbot's *The Holographic Universe* (1991).

The most recent research evidence that verifies this connection between quantum dimensions and their projections as "consciousness" is the discovery by Stuart Hameroff and Sir Roger Penrose of quantum vibrations in "microtubules" inside brain neurons. "They suggest that EEG rhythms (brain waves) also derive from deeper level microtubule vibrations, and that from a practical standpoint, treating brain microtubule vibrations could benefit a host of mental, neurological, and cognitive conditions." (Hameroff & Penrose, 2014) These researchers based their discovery on two decades of previous research on "orchestrated objective reduction" (Orch Or) which suggests that quantum vibrational computations (microtubules) are orchestrated by synaptic inputs and memory stored in microtubules. (Hameroff & Penrose, 2014) The term "orchestration" suggests the potential to harmonize conscious patterns with archetypal patterns according to the principle of coherence. Coherence is a measurable harmonious neurobiological state that has been studied extensively by the Heartmath Institute. (McCraty, Rollin & Childre, Doc (2010); McCraty, Deyhle, & Childre (2012).

Coherent states appear to have the same characteristics as Carl Jung's "Compensational" states and "Flow". (Csíkszentmihályi, 1990) Compensation is the fundamental Jungian principle that makes discovery of unconscious meaning possible using the process of symbolic analysis on projected archetypal energy patterns in dream images. As images, dreams and computer graphics have structural correlates. The entirety of Carl Jung's psychiatric theory and practice is based on symbolic images in dreams, so interpretation of the underlying meaning in dream-computer images measured in terms of coherence has identical potential.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

This chapter argues that dream images and computer monitor images are directly correlated as to structure, perception, and leaning-healing potential. These correspondences are a primary component for qualifying psychecology games as a dream analog. (Schafer, 2011; Schafer, 2012) Therefore, psychecology games can be used as a cognitive research instrument in conjunction with MEG, EEG, ECG, fMRI, etc. for measuring real-time neurobiological activity in subject game players. (Schafer & Yu, 2011; Schafer, 2011; Schafer, 2012; Trafton, Anne, 2014) Research methodologies must be contextualized to various media and to a myriad demographics, but the dynamics of dream analysis remains constant. Though the media dream includes multi-functional sensory data that is integrated at levels of the unconscious, this chapter emphasizes the primarily visual symbolism in images.

Jung is not the only observer to say that the changing human perceptions of reality—what he calls the Spirit of the Age (Jung, 1933, pp. 175,196-220)—have been thrust upon us so rapidly that throughout the twentieth century it has disoriented and traumatized human collectives throughout the world. At base, the changes derive from discoveries in the realms of physics, psychology, and computer science. It is a well established fact that the psychological dynamics and structure of advertising and propaganda entail utilizing unconscious demographic patterns to sell their "products". This chapter develops the hypothesis that because the discoveries in these sciences have been incorporated into the holographic, psychological metaverse, its contextual projections can be analyzed like dreams.

CONCLUSION

Jungian Dream Analysis

The unconscious aspect of *psyche* is much like a *unified field of physics*. According to Jung:

Sooner or later nuclear physics and the psychology of the unconscious will draw closer together as both of them, independently of one another and from opposite directions, push forward into transcendental territory, the one with the concept of the atom, the other with that of the archetype. (Aion, p. 69; from Jacoby, 1973)

Dreams employ a symbolic language that communicates messages between that part of the dreamer's brain that is conscious and that part that is unconscious. Psychecology games are structurally analogous with the structure of Psyche. Due to the interactive loop between conscious-unconscious player choices a complex cognitive process becomes a recursive monitor image—the interactive process articulates the psychological process. This interactivity represents an important breakthrough that incorporates total functionality in the mediating process.

Functionality should be understood in terms of agency and the consequences of action. In the sense that players make psychological choices that are both conscious and unconscious and take action based on the images (and other sensory data) in their perceived reality, video games can be understood as an authentic simulation of a "real" existence in which causal behavior is generated and consequences of the behavior seem to materialize as images. So doing, players are creating their functional four-dimensional reality based largely on unconscious memories of psychological experiences. Using perceptions of images, they are employing cognitive alchemy and symbolic language to translate the unknown into the known. Imbedded unconscious meaning is translated into meaningful insight—precisely the compensational objective of Jung's Amplification Process.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

RECOMMENDATIONS FOR FUTURE RESEARCH

An abundance of current research on Artificial Intelligence, psychological dynamics of global target groups, and robotics is underway. However, the frameworks of research are generally antiquated; i.e. they do not employ a comprehensive understanding of dream dynamics at work in the virtual media-sphere.

Contextual Ecosystems

As reported by EUROSIS—The European Multidisciplinary Society for Modeling and Simulation Technology—and directly relevant to analyzing contextual groups with digital technology, we hear that Facebook has "Like Farms". Emiliano De Cristofaro and colleagues (2014) from the University College London, targeted Facebook users in the US, France, India, Egypt, and worldwide. By measuring "likes" activity on an array of Facebook pages, researchers correlated user profiles with measures of page-popularity. Data relative to understanding "likes" of Facebook users translates to traditional personality profiling of target groups, but does not sufficiently factor unconscious dynamics into conclusions. Though an intuitive sense of the scope of unconscious issues exists, the importance of data relative to page popularity has commercial motivation and is being applied primarily to promote business opportunities. Such studies of contextual collective personae should be employed in more profound ways to promote cognitive coherence.

The study of online social networks has revolutionized the way social scientists understand human interaction on a grand scale. It is based on the assumption that the fundamental unit of interaction is the social tie that exists between individuals. This tie can be a message that one person has sent to another, that one person follows another, that one person "likes" another and so on. (8/1/2014)

The MIT Technology Review (1-11-2015) reports that computer scientists in the emergent science of human-data interaction say, "The rapidly evolving ecosystems associated with personal data are creating an entirely new field of scientific study." (Ecosystems, 2015) Again, the article goes on to illustrate the marketing potentials of this new science instead of the more significant potentials for addressing cognitive coherence:

Back in 2013, the UK supermarket giant, Tesco, announced that it was installing face recognition software in 450 of its stores that would identify customers as male or female, guess their age and measure how long they looked at an ad displayed on a screen below the camera. Tesco would then give the data to advertisers to show them how well their advertising worked and allow them to target their ads more carefully. (Ecosystems, 2015)

A remarkable online article defines three neuromarketing lessons derived from video games. The article explains that:

Games are uniquely adept at leveraging human psychology to motivate behavior. We play them for hours on end and enter into a state of flow with an ease not found in other fields. This is no accident: during my time studying game design and working in the field, neurological language like "dopamine hits" and "social proof" were commonplace. In this post, I'll examine some of the common tools used by video game designers in order [to] insure engagement and ultimately drive sales. (Claes, 2014)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Another recent article is about *The New Statistics*. In their abstract, authors say that, in response to heightened concern that our published research literature is incomplete and untrustworthy, "We need to make substantial changes to how we conduct research." The article goes on to explain that null hypothesis significance testing (NHST) has recognized flaws, and recommends preferred techniques including estimation based on effect sizes, confidence intervals, and meta-analysis (Cumming, 2013).

Robotics and Artificial Intelligence

Recent advances in robotics research constantly confirm the functional interactivity of physics, mathematics, and cognition. In another fascinating article from the MIT Review (Sept 22, 2014), "Eugene Izhikevich thinks you shouldn't have to write code in order to teach robots new tricks. "It should be more like training a dog," he says. "Instead of programming, you show it consistent examples of desired behavior." For example, to teach a robot to pick up trash, a mentor uses a remote control to repeatedly guide the robot's gripper to perform that task. After just minutes of repetition, the robot takes the initiative and starts doing the task for itself. Izhikevich's startup, *Brain Corporation*, based in San Diego, has developed an operating system for robots called *BrainOS* to make that possible. (Simonite, 2014)

Global Coherence Initiative

Perhaps the most exciting and relevant development in the vast arena of current research is the Global Coherence Initiative (GHI) spearheaded by the Institute of HeartMath. There is strong support for the existence in the earth's ionosphere of a global information field that connects all living systems and consciousness. "Every cell in our bodies is bathed in an external and internal environment of fluctuating invisible magnetic forces that can affect virtually every cell and circuit in biological systems." (McCraty, Deyhle, & Childre, 2012). Progress on the establishment of a global network of monitoring stations for studying these electro magnetic fluctuations portends a quantum leap in the direction of understanding of the media dream and how humans interact with it to create reality.

The scope of integration among all earth's inhabitants is becoming a conscious fact throughout the human population. Though vested interests seem always to marginalize the importance of the natural environment, the pervasiveness and popularity of media paraphernalia must eventually lead to a widespread appreciation of their dynamic impact and integration at all functional levels of human reality. We know something about the sun's influences on the earth—like the fact that without the sun, all life on earth would disappear. However, scientists have been expanding our knowledge with regularity. "The most likely mechanism for explaining how solar and geomagnetic influences affect human health and behavior are a coupling between the human nervous system and resonating geomagnetic frequencies called Schumann resonances, which occur in the earth-ionosphere resonant cavity and Alfvén waves. (McCraty, et al 2012) HeartMath researchers say that, "It is well established that these resonant frequencies directly overlap with those of the human brain and cardiovascular system" (McCraty, et al 2012).

This section on the need for further research is rather long in order to emphasize the vast scope and significance of the subject. The scope of necessary research is ontological, and research must be framed to contemplate the very nature of the paradigm shift into a virtual reality. The chapter takes the perspective—as does Joseph Campbell--that "mythology is psychology", that the Jungian archetypes of mythology are the selfsame energy and information patterns of physics, that the media-sphere or metaverse is analogous with psyche, and that it constitutes a holographic projection of media dreams from the collective unconscious.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Relative to more research, the scope of the chapter is both metaphysical and practical. Realization of the existence of a virtual *correspondence* between individuals and collectives—a sort of fractal bio-energetic relationship between the individual and the group as illustrated in such phenomena as focus groups, target groups, or statistical entities—marks the beginning of a more positive scientific understanding of the cognitive reality of the mediated dream. Many scientists have addressed this subject, and David Foster has provided an excellent overview. (Foster, 1985) Today, the marketing industry applies the correspondence principle constantly. Carl Jung understood the correspondence between individual and collective psyche as a fundamental principle of his therapeutic method. Accordingly, with appropriate adjustments in research methodology, therapy that works on individual levels may also be hypothesized to work at collective levels.

Now the doctor in me refuses point blank to consider the life of a people as something that does not conform to psychological law… "A people, in the doctor's eyes, presents only a somewhat more complex picture of psychic life than the individual," to quote Swedenborg. And so we can draw a parallel: just as in me, a single human being, the darkness calls forth the helpful light, so does it also in the psychic life of a people. (Jung 1933, p 209-210)

The therapeutic efficacy of narrative/drama is ancient and well-documented. (White, 2000) Phenomena such as catharsis have long been studied, so—given the robust interest in the technologies of the media-sphere—drama-based video games and Massively Multi-Player Online Games (MMOG) may provide a portal relative to collective healing and such new-paradigm experiences as epiphany, "metanoia" (Markley, 1972) or psychological critical mass.

REFERENCES

Bellman, O. (2015). *Researchers create virtual humans*. Playa Vista, CA: USC Institute for Creative Technologies. Retrieved from http://www.army.mil/article/147822

Bohm, D. (1989). Quantum Theory. New York: Prentice Hall. (Original work published in 1951). Retrieved from http://en.wikipedia.org/wiki/David_Bohm

Boyd, F. (2015). Immersive Media: To the Holodeck and Beyond. Retrieved from https://www.linkedin.com/grp/post/1039687-6004890688691453955

Campbell, J. (1973). *The hero with a thousand faces*. Princeton, NJ: Princeton University Press.

Capra, F. (1976). *The tau of physics*. Boulder, CO: Shambhala Publications, Inc.

Chopra, D. (1995). Body, Mind, and Soul: The mystery and the magic (VHS). Public Broadcasting Service/Turner Home Entertainment: Atlanta, GA.

Claes, Z. (2014). Three neuromarketing lessons derived from video games. *Neuromarketing*. Retrieved from http://www.neurosciencemarketing.com/blog/articles/neuromarketing-video-games.html

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

COCORO. anon. (2015). COCORO: robot swarms use collective cognition to perform tasks.European Commission: Digital Agenda for Europe. Retrieved from https://ec.europa.eu/digital-agenda/en/news/cocoro-robot-swarms-use-collective-cognition-perform-tasks

Collier, J., Burckhardt, C. B., & Lin, L. H. (1971). *Optical Holography*. New York: Academic Press.

Crawford, K. H. (2015). Here's looking at you: ONR tests new glasses for augmented reality system with Marines. *Office of Naval Research.* Retrieved from http://www.onr.navy.mil/en/Media-Center/Press-Releases/2015/Augmented-Reality-Demo-With-Marines.aspx

Csíkszentmihályi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York: Harper and Row.

Cumming, G. (2013). *Substantial changes to how we conduct research needed*. Statistical Cognition Laboratory, La Trobe University. Retrieved from http://pss.sagepub.com/content/25/1/7.full

Curtis, A. (2009). *Century of the self: Happiness machines for free.* British Broadcasting Corporation, Freedocumentaries.com. Retrieved from http://freedocumentaries.org/film.php?id=140

De Cristofaro, E. (2014). Facebook "Likes". *The European Multidisciplinary Society for Modeling and Simulation Technology.* Retrieved from http://www.technologyreview.com/view/530961/the-hidden-world-of-facebook-like-farms/

Demers, T. (2014). Using gamefication to drive brand awareness. *Small Business Trends.* Retrieved from http://smallbiztrends.com/2014/03/using-gamification-brand-awareness.html

DeValois, Karen, & DeValois, Russell, and Yund, W. W. (1979). Responses of striate cortex cells to grating and checkerboard patterns. *The Journal of Psychology*, *291*, 483–505. PMID:113531

Dossey, L. (1982). *Space, time, and medicine* (pp. 108–109). Boston: New Science Library.

Ecosystems associated with personal data is creating an entirely new field of scientific study (2015). The MIT Technology Review. Retrieved from http://www.technologyreview.com/view/533901/the-emerging-science-of-human-data-interaction/)

Eddington, A. (1935). *The Nature of the physical world*. Cambridge University Press.

Foster, D. (1985). *The philosophical scientists*. USA: Dorset Press.

Garling, C. (2015). An algorithm that can help robots walk off injuries. MIT Technology Review. Retrieved from http://www.technologyreview.com/news/537791/an-algorithm-that-can-help-robots-walk-off-injuries/?utm_campaign=newsletters&utm_source=newsletter-daily-all&utm_medium=email&utm_content=20150528#

Geril, P. (2014c). Startup Europe- Capacity building for Business Angels (BAs) in ICT/ A digital Agenda for Europe. *Newsroom*. Retrieved from https://ec.europa.eu/digital-agenda/en/news/startup-europe-capacity-building-business-angels-bas-ict

Geril, P. (2015b). The virtual worlds 2015 survey report. *EUROSIS, The European Multidisciplinary Society for Modelling and Simulation Technology*. Retrieved from https://www.linkedin.com/grp/post/1039687-5988685234479075329

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Goffman, E. (1974). *Frame analysis: An essay on the organization of experience*. New York: Harper and Row.

Goleman, D. (1979, February). Holographic memory: interview of Karl Pribram. *Psychology Today*, *12*(9), 71–72.

Greitemeyer, T., & Mügge, D. (2014). Video games do affect social outcomes: A meta-analytical review of the effects of violent and prosocial video game play. *Sage Publishing*. Retrieved from http://psp.sagepub.com/content/40/5/578.full

Hall, M. P. (2003). *The Secret Teachings of All Ages*. New York: Jeremy P. Tarcher/Penguin.

Hameroff, S., & Penrose, S. R. Discovery of quantum vibrations in 'mocrotubules' corroborates theory of consciousness. Retrieved 1/16/2014 http://phys.org/news/2014-01-discovery-quantum-vibrations-microtubules-corroborates.html)

How Yahoo Research Labs Studies Culture as a Formal Computational Concept (2014). *MIT Technology Review*. Retrieved from http://www.technologyreview.com/view/529521/how-yahoo-research-labs-studies-culture-

Hubb, B. (2015). John 1:1. Retrieved from http://biblehub.com/john/1-1.htm

Human behavior as a formal computational concept (2015). *MIT Technology Review*. Retrieved 1/25/2015 http://www.technologyreview.com/view/529521/how-yahoo-research-labs-studies-culture-as-a-formal-computational-concept/

Human Brain Project. (2014). *MIT Technology Review* July 7, 2014. http://www.technologyreview.com/news/528796/neuroscientists-object-to-europes-human-brain-project/

Jacobi, J. (1973). *The psychology of CG Jung*. London, USA: Yale University Press.

Jung, C. (1933). *Modern man in search of a soul*. Orlando, Fl: Harcourt Brace & Co.

Kleinman, A. (2014). Man successfully controls 2 prosthetic arms with just his thoughts. Huffington Post. Retrieved from http://www.huffingtonpost.com/2014/12/18/prosthetic-arms-thoughts_n_6348026.html

Kowal, E. (2015). Computer simulations improve lethality. The Official Home Page of the United States Army. Retrieved fromwww.army.mil

Lakoff, G. (2008). *The political mind* (pp. 83–84). USA: Viking Penguin.

Lashley, K. (1950). In search of the engram. In *Physiological Mechanisms in Animal Behavior* (pp. 454–482). New York: Academic Press.

Makuch, E. (2015). Microsoft has invested heavily in Holographic research and the dynamics of machine intelligence. *Gamespot.com*. Retrieved from http://www.gamespot.com/articles/windows-holographic-revealed/1100-6424791/

Markley, O. W. (1972). *Unity with diversity: Toward Social Policies for a future world order*. Center for the Study of Social Policy: Stanford Research Institute.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

McCraty, R., & Childre, D. (2010). Coherence: bridging personal, social, and global health. Alternative Therapies, 16(4).

McCraty, Deyhle, & Childre (2012). The Global Coherence Initiative: Creating a coherent planetary standing wave. *Global Advances in Health & Medicine*, 1(1). Retrieved from www.gahmj.com

MediaGrid, Immersive Education Initiative. (n. d.). Retrieved from http://immersiveeducation.org/

Mishlove, J. (1998). The Holographic Brain with Karl Pribram, MA; Ph.D. *TWM.co.nz*.

Mortier, R. (2015). *The emerging science of human-data interaction*. MIT Technology Review.

Moyers, B. (1987). The secret government: The constitution in crisis. *Public Broadcasting System.* Retrieved from https://search.yahoo.com/search?fr=mcafee&type=B011US739D20150419&p=The+Secret+Government,+Bill+Moyers

Myers, M. (2015). Augmenting Agriculture with Social Impact Games has produced a digital prototype of a new social impact board game Bumper Crop. *The Arts Humanities Research Council research project Play to Grow.* Retrieved from https://play.google.com/store/apps/details?id=foam.bumpercrop

Pattern recognition in Go (2014). *MIT Technology Review 12-15-2014.* Retrieved from http://www.technologyreview.com/view/533496/why-neural-networks-look-set-to-thrash-the-best-human-go-players-for-the-first-time/

Penfield, W. (1975). *The mystery of the mind; a critical study of consciousness and the human brain*. Princeton, N.J.: Princeton University Press.

Philippe, G. (2015a). The AI Revolution: The Road to Superintelligence. *EUROSIS, TheEuropean Multidisciplinary Society for Modelling and Simulation Technology.* Quoted from Private Group DARPA seeks non-traditional entities to develop new robotics capabilities. *Army Technology.com.* Retrieved from http://www.army-technology.com/news/newsdarpa-seeks-non-traditional-entities-to-develop-new-robotics-capabilities-4580219?WT.mc_id=DN_News#.VVtNUbpwZLQ.linkedin

Pietsch, P. (1972, May). Shufflebrain. *Harper's Magazine* 244, p. 66.

Pietsch, P. (1981). *Shufflebrain: The quest for the hologramic mind* (p. 78). Boston: Houghton Mifflin.

Pister, K. (2015). Smart Dust. *The European Multidisciplinary Society for Modeling and Simulation Technology.* Retrieved from https://medium.com/war-is-boring/smart-dust-is-getting-smarter-4b062abd7769

Pollen, D. A., & Tractenberg, M. C. (1972, May 12). Alpha rhythm and eye movements in eidetic imagery. *Nature*, *237*(5350), 109–112. doi:10.1038/237109a0 PMID:4555444

Postman, N. (1985). *Amusing ourselves to death*. USA: Viking Penguin Inc.

Pribram, K. (1969, January). The neurophysiology of remembering. *Scientific American*, *220*(1), 75. doi:10.1038cientificamerican0169-73 PMID:4972532

Pribram, K. (1977). Language of the brain (pp. 123 & 169). Monterey, CA: Wadsworth Publishing.

Reeves, B., & Nass, C. (1996). The media equation. Stanford: Center For the study of Language and Information & Cambridge University Press.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Regaaldo, A. (2014). Neuroscientists Object to Europe's Human Brain Project. *MIT Technology Review.* Retrieved from http://www.technologyreview.com/news/528796/neuroscientists-object-to-europes-human-brain-project/

Restak, R. (1979). *The brain* (p. 253). New York: Warner Books.

Restak, R. (1981). Brain power: A new theory. *Science Digest*, (March): 19.

Schafer, S. (2011). Articulating the paradigm shift: Serious games for psychological healing of the collective persona. In M.M. Cruz-Cunha (Ed.), *Handbook of Research on serious games as educational business and research tools* (Ch. 19). Hershey, PA, USA: IGI Global. Retrieved from www.igi-global.com

Schafer, S. (2012). Optimizing Cognitive Coherence, Learning, and Psychological Healing with Drama-based Games. In J. Gackenbach (Ed.), *Video Game Play and Consciousness.* Nova Science Publishers.

Schafer, S., & Palamides, T. (2011). Designing an Ethical structure for Social Influence Marketing. In M.M. Cruz-Cunha, P. Gonçalves, N. Lopes, & G.D. Putnik (Eds.), *Handbook of research on Business Social Networking: Organizational, Managerial, and Technological Dimensions*. Hershey, PA, USA: IGI Global. Retrieved from www.igi-global.com

Schafer, S., & Yu, G. (2011). *Meaningful video games: drama-based video games as transformational experience*. Hershey, PA, USA: IGI Global. doi:10.4018/978-1-60960-567-4.ch019

Schroeder, R. (2014). Big data and the brave new world of social media research. *Big Data & Society*. Retrieved from http://bds.sagepub.com/content/1/2/2053951714563194.abstract

Simonite, T. (2014). Digital Summit: Microsoft's Quantum Search for the "Next Transistor". MIT Technology Review. Retrieved from http://www.technologyreview.com/news/528256/digital-summit-microsofts-quantum-search-for-the-next-transistor/

Simonite, T. (2014a). Robots That Learn Through Repetition, Not Programming. *MIT Technology Review.* Retrieved from http://www.technologyreview.com/news/530871/robots-that-learn-through-repetition-not-programming/

Simonite, T. (2014c). Google's Brain-Inspired Software Describes What It Sees in Complex Images. *MIT Technology Review*. Retrieved from http://www.technologyreview.com/news/532666/googles-brain-inspired-software-describes-what-it-sees-in-complexeimages/?utm_campaign=newsletters&utm_source=newsletter-daily-all&utm_medium=email&utm_content=20141119

Simonite, T. (2015b). A Better Way to Build Brain-Inspired Chips. *MIT Review*. Retrieved from http://www.technologyreview.com/news/537211/a-better-way-to-build-brain-inspired-chips/

Simonite, T. (2015d). A startup's neural network can understand video. *MIT Review*. Retrieve from http://www.technologyreview.com/news/534631/a-startups-neural-network-can-understand-video/

Smart, J. M., Cascio, J., & Paffendorf, J. (2007). Metaverse Roadmap Overview, 2007. *Accelerated Studies Foundation*. Retrieved from http://en.wikipedia.org/wiki/Metaverse

Smith, H. (1991). *The World's Religions*. New York: Harper Collins.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Stoupel, E. (1993). Sudden cardiac deaths and ventricular extrasystoles on days of four levels of geomagnetic activity. *J Basic Physiol Pharmacol.*, *4*(4), 357–366. PMID:8664252

Talbot, M. (1991). *The holographic universe*. New York: Harper.

Cumming, G. (2013). The New Statistics. *Psychological Science*, 25(1), 7-29. Retrieved from http://pss.sagepub.com/content/25/1/7.full

The rapidly evolving ecosystems associated with personal data is creating an entirely new field of scientific study (2015). *MIT Technology Review.* Retrieved from http://www.technologyreview.com/view/533901/the-emerging-science-of-human-data-interaction/)

Toma, C., & Hancock, J. (2013). Self-affirmation underlies Facebook use. *Pers. Soc. Psychol. GBull.*, 39(3), 321-331. Retrieved from http://psp.sagepub.com/content/39/3/321.abstract

Trafton, A. (2014). New brain-scanning technique allows scientists to see when and where the brain processes visual information. *Medicalxpress.com*. Retrieved from http://medicalxpress.com/news/2014-01-brain-scanning-technique-scientists-brain-visual.html

Twist, J. (2004, March 24). Brain waves control video game. *BBC News.* Retrieved from http://news.bbc.co.uk/2/hi/technology/3485918.stm

Understanding human behavior as a Formal Computational Concept (2014). *MIT Technology Review & Yahoo Research Labs*. Retrieved from http://www.technologyreview.com/view/529521/how-yahoo-research-labs-studies-culture-as-a-formal-computational-concept/

VanHeerden, P. (1970, July 25). Models for the brain. *Nature*, *227*(5256), 410–411. doi:10.1038/227410b0 PMID:5428449

Wakefield, J. (2014). Singularity: The robots are coming to steal our jobs. *BBC News.* Retrieved from http://www.bbc.com/news/technology-25000756

White, M. (2000). *Narrative Means to Therapeutic Ends; Maps of Narrative Practice. Reflections on Narrative Practice*. Adelaide, South Australia: Dulwich Centre Publications.

Why Neural Networks Look Set to Thrash the Best Human Go Players for the First Time. *MIT Technology Review*. Retrieved from http://www.technologyreview.com/view/533496/why-neural-networks-look-set-to-thrash-the-best-human-go-players-for-the-first-time/

KEY TERMS AND DEFINITIONS

Archetypal Representations: Jungian archetypes of mythology are the selfsame energy and information patterns of physics that constitute the source of the media-sphere or metaverse and are analogous with psyche. Archetypes are the psychoenergetic, generative source of holographic projections (archetypal representations) of media dreams from the collective unconscious.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Archetype: Jungian archetypes of mythology are the selfsame energy and information patterns of physics that constitute the source of the media-sphere or metaverse and are analogous with psyche. Archetypes are the psychoenergetic, generative source of holographic projections (archetypal representations) of media dreams from the collective unconscious.

Artificial Intelligence (AI): This field overlaps with electronics, computer science, robotics, mechatronics, nanotechnology, and bioengineering.

Avatar: The doctrine of avatars is a mythical concept in which a deity takes incarnation in human form that embodies a cosmic concept or philosophy in a person; i.e. Jesus embodies the concept of Love, and the Buddha embodies the concept of Compassion. In in the computer world, an avatar specifically refers to a character that represents an online user. Avatars are commonly used in multiplayer gaming, online communities, and Web forums.

Coherence: A single wavelength or frequency with all the waves in phase.

Collective Personae: According to Carl Jung, the persona is the social face the individual presents to the world—a kind of mask that is projected into the world and serves as a mirror of unconscious psychic patterns of individuation. A primary function of the persona is the evolutionary development of a viable, adaptive social persona. Jung was explicit as to the existence of collective personae (much like target groups in advertising) that are functional correspondences to individual personalities.

Compensation: A natural process aimed at establishing or maintaining balance within the psyche. The activity of consciousness is selective and requires direction, but direction requires the exclusion of everything irrelevant. According to Jung, the conscious orientation is one-sided. The psychic contents that are excluded and inhibited by the chosen direction sink into the unconscious, where they form a counterweight to the conscious orientation. As a rule, the unconscious compensation does not run counter to consciousness, but is rather a balancing or supplementing of the conscious orientation. In dreams, for instance, the unconscious supplies all those contents that are constellated by the conscious situation but are inhibited by conscious selection, although a knowledge of them would be indispensable for complete adaptation.

Contextuality: Jung's principle of compensation demands that dream interpretation must incorporate significant familiarity and understanding of the individual or patient who is doing the dreaming. In order for game software to provide authentic compensation, the game must be specifically programmed to the psychological profile or "context" of the player.

Holography: A method of reproducing a three-dimensional image of an object by means of light wave patterns recorded on a photographic plate or film. Holography is sometimes called lensless photography because no lenses are used to form the image. The plate or film with the recorded wave patterns is called a *hologram.* The light used to make a hologram must be coherent, i.e. of a single wavelength or frequency and with all the waves in phase. A coherent beam of light can be produced by a laser. The term is often used as a metaphorical model; i.e. "holographic universe" and "holographic brain".

Interactivity: In computers, interactivity is the dialog that occurs between a human being and a computer program. Programs that run without immediate user involvement are not interactive, but video games are usually thought of as fostering a great amount of interactivity. In the context of the chapter, interactivity extends to the four psychoenergetic Jungian functions (thinking, feeling, sensing, and intuiting). However, order entry applications, channel changing, and many other business applications are also interactive.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Media Dream: The media dream, a term coined by the author, goes beyond the technologically generated virtual realities called the metaverse and provides a comprehensive framework for correlating fundamental dimensions of reality by including multi-functional dynamics of Jungian dreams. The dreamscape of the holographic media-sphere that has the same psychoenergetic structure as Jungian dreams; i.e. dramatic, multifunctional structure.

Metaverse: The Metaverse is a virtual shared space—a collective—created by the convergence of virtually enhanced physical reality and "physically persistent virtual space, including the sum of all virtual worlds, augmented reality, and the internet. The word *metaverse* is a portmanteau of the prefix "meta" (meaning "beyond") and "universe" and is typically used to describe the concept of a future iteration of the internet, made up of persistent, shared, 3D virtual spaces linked into a perceived virtual universe. The term was coined in Neal Stephenson's 1992 science fiction novel *Snow Crash*, in which humans, as avatars, interact with each other and software agents, in a three-dimensional space that employs the real world as metaphor. Stephenson used the term to describe a virtual reality-based successor to the Internet. Concepts similar to the Metaverse have appeared under a variety of names in the cyberpunk genre of fiction as far back as 1981 in the novella *True Names*, Stephenson stated in the afterword to *Snow Crash* that after finishing the novel he learned about *Habitat*, an early MMORPG which resembled the Metaverse" (Smart, J.M., Cascio, J. and Paffendorf, J., 2007).

***Psych*ecology:** The unified field ecology of psychoenergetic, multidimensional reality.

***Psych*ecology Video Games:** These drama-based, psychoenergetic games should not be unduly associated with video games in their present state of development. Rather, they constitute an emergent game genre that is characterized by advanced principles of cognitive dynamics and story-telling art. These games have all the components of dreams, physics, the media, and interactivity with which to research the fourth dimension.

Psychological Functions: The classical four dimensions of human reality are 1) sensory/material, 2) emotional, 3) mental/rational and 4) spiritual. Jung called them 1) Sensing, 2) Feeling, 3) Mental, and 4) Intuitive. These dimensions are symbolized by 1) earth, 2) water, 3) air, and 4) fire.

Recursion: The process of repeating items in a self-similar way. For example, when the surfaces of two mirrors are exactly parallel with each other, the nested images that occur are a form of infinite recursion. The term has different meanings specific to a variety of disciplines ranging from linguistics to logic. The most common application of recursion is in mathematics and computer science, in which it refers to a method of defining functions. The term function refers to a relation between a set of inputs and a set of permissible outputs in which the function being defined is applied within its own definition. The functional relationship between Jungian archetypes and archetypal representations (projections) is analyzed in dreams according to symbolic images that are projections of unconscious psychoenergetic patterns. Recursion may also apply to metaphorical resonance that can be coded mathematically.

Robotics: A branch of engineering that involves the conception, design, manufacture, and operation of robots. This field overlaps with electronics, computer science, artificial intelligence, mechatronics, nanotechnology, and bioengineering.

This research was previously published in Exploring the Collective Unconscious in the Age of Digital Media edited by Stephen Brock Schafer; pages 232-260; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 19
On the Use of Bayesian Network in Crime Suspect Modelling and Legal Decision Support

O. E. Isafiade
University of Cape Town, South Africa

A. B. Bagula
University of the Western Cape, South Africa

S. Berman
University of Cape Town, South Africa

ABSTRACT

Predictive policing (Pp) relates to identifying potentially related offences, similar criminal attributes and potential criminal activity, in order to take actionable measures in deterring crime. Similarly, Legal Decision Making Process (LDMP) considers some level of probabilistic reasoning in deriving logical evidence from crime incidents. Bayesian Networks (BN) have great potential in contributing to the area of Pp and LDMP. Being based on probabilistic reasoning, they can assess uncertainty in crime related attributes and derive useful evidence based on crime incident observations or evidential data. For example, in a particular context of crime investigation, BN based inference could help collect useful evidence about a crime scenario or incident. Such evidence promotes effective legal decision making process and can assist public safety and security agencies in allocating resources in an optimal fashion. This chapter reports on various application areas of BN in the crime domain, highlights the potential of BN and presents "thought experiments" on how offender characteristics could be modelled for decision support in legal matters. The chapter further reports on the performance of empirical analysis in the legal decision support process, in order to elucidate the practical relevance and challenges of using BN in the crime domain.

DOI: 10.4018/978-1-7998-0951-7.ch019

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. INTRODUCTION

The foundation and motivation for repeat victimisation study (RVS) relates to the fact that repeat crime is much more likely to happen in a short interval of time after the first incident, and often within a close proximity (Evett et al., 2006 & Riesen et al., 2009). RVS is also associated with the notion that crime is mostly initiated and perpetuated "where and when" there is a "conducive" environment for a potential offender and the presence of a suitable (vulnerable) target victim. This notion is corroborated by the long-established routine activity theory (RAT) in criminology depicted in Figure 1 (Clarke et al, 1993), which stipulates that a crime can only be committed when three specific criteria are involved. These criteria are that there must be a motivated offender, a suitable target or vulnerable victim (available at an appropriate location and time), as well as the absence of a capable guardian. RAT is a sub-field of crime opportunity theory that focuses on situations or circumstances of crimes. Jury members also consider such circumstances when probing crime scenarios, in order to justify the legal decision process.

Furthermore, judges or jury members often make legal decisions based on their instinct and/or experience (Milgram A., 2013). Such legal decision making process considers some level of probabilistic reasoning in deriving logical evidence, although the mathematical rigour for such probabilistic evidence is sometimes ignored or ill-defined.

One aspect that stimulates interest is whether a crime X that previously occurred or repeatedly happens in location K is committed by the same individual suspect in that location, another location of close proximity or geographically close domestic dwellings, given observations of certain exculpatory or incriminating evidence or findings. For example, exculpatory or incriminating evidence could be footwear or footprint marks found at both points of entry of the crime scenes, particular items recovered from the points of entry at both crime scenes, type of victim information repeatedly reported and Modus Operandi (MO) involved. The degree of similarity between the evidence in two crimes can further motivate a number of research questions:

1. How does the evidence transfer between previous and current crimes?
2. To what extent can we attribute the crime observations or offender characteristics in the current crime to the previous one, for example in terms of MO reported?

Figure 1. A model of routine activity theory



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Such probing or investigative research questions are also at the intersection of efforts in predictive policing.

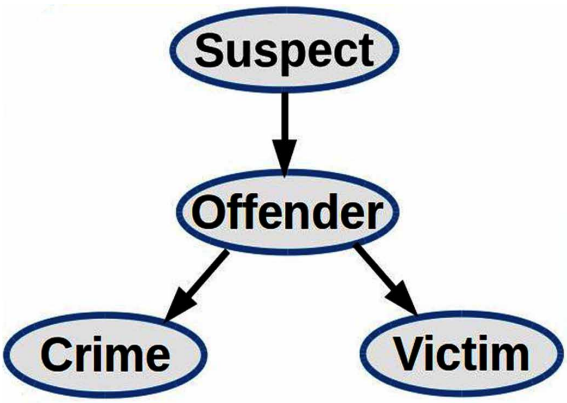
Predictive policing deals with the usage of analytical and predictive techniques in law enforcement to identify and curb potential criminal activity. Existing software technology called PredPol is a good sample application in the area of predictive policing (Predpol, 2013). However, PredPol does not predict a potential offender, but helps in identifying and preventing potential crime incidents by predicting the next probable crime location. Predictive policing approaches are categorised into four distinct methods, these are:

- Methods for predicting perpetrators' identities,
- Methods for predicting offenders' characteristics,
- Methods for predicting crimes,
- Methods for predicting victims of crime.

It is widely recognised that Bayesian Networks (BNs), also known as Bayesian belief networks (BBNs), or Probabilistic Directed Acyclic Graphical Model (PDAGM), have great capacity for contributing to the area of predictive policing and crime suspect modelling (Zoete et al, 2014). Being based on probabilistic reasoning, they can derive useful probabilistic dependencies and inferences about the crime subject. Such probabilistic inference can ultimately assist public safety agencies in identifying mitigation priorities and allocating security resources in an optimal fashion, during investigative crime procedures. The terms Bayesian network (BN) and Bayesian belief network (BBN) shall be used interchangeably throughout the chapter.

A BN is a graphical mathematical model that encodes probabilistic relationships and conditional independence relationships between crime variables in the graph structures. A BN can be used to model suspect/offender behaviour as presented in Figure 2, a simple (4-nodal) BN, where a crime "suspect" variable encodes information about who the "offender" would most likely be, and the offender has a direct influence on both crime type and victim(s) involved. The offender encodes information about the crime committed and has a direct impact on the victim.

Figure 2. A simple (4-nodal) Bayesian network



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

A Bayesian network is potentially useful in the following ways:

- Acquiring useful knowledge (probabilistic dependencies) from crime data.
- Assessing uncertainty in crime related attributes or characteristic features.
- Combining crime knowledge sources or observations.
- Deriving actionable crime evidence and mitigation priorities.

BNs are of practical significance in real life scenarios, especially where domain experts are not available or in limited supply for knowledge acquisition. Its usefulness is further evident in the fact that it is a compact representation of the joint probability distribution over the crime variables and encodes the conditional independence relationships between the variables in the graph structure (BN).

Current criminal justice systems and legal decision making processes are mostly dependent on a jury's intuition and experience. Judges have the best interest of the community at heart when they make legal decisions, but decisions made with instincts and/or experience have not significantly helped to promote the public safety of many nations (Milgram, 2013). However, a BN can assist a jury with reasoning under uncertainty (Sjoerd et al, 2014), in order to make a decision that makes the community and nation at large safer.

BNs have received widespread acceptance in academia and industry and proven useful as practical applications of artificial intelligence in solving real-life problems. For example, BNs have practically and successfully been adopted in speech recognition (Nefian et al., 2002), spam filtering/text mining (Raghuram et al., 2009), image classification systems (Jayech et al, 2010) and smart home activity prediction (Nazerfard, 2013), to mention a few.

Furthermore, BNs have over the years received much attention in the public safety and crime domain, and successfully been applied in the areas of offender behavioral modeling (Baumgartner et al., 2008), predicting victimisation (Riesen et al, 2009), crime linkage analysis (Zoete et al, 2014) and Forensic science (Sjoerd et al, 2014), among others. The BN approach to crime profiling typically draws on the fields of criminology, offender modeling and knowledge-based systems. The adoption of probabilistic networks make it practical to represent multidimensional relationships between all relevant crime or offender variables that may be identified at a crime scene. Hence, a valid BN model is a probabilistic directed acyclic graphical model (PDAGM) and can be used to predict unknown crime variables composing an offender profile based on observational variables from a crime scene.

1.2 Contributions and Outline

BNs have been employed for deriving crime knowledge from information in various capacities. This chapter explores various implementation of BN for crime knowledge discovery and delivery, and presents "thought" experiments that focus on understanding the underlying arguments regarding crime or offender modeling using BNs, as well as support in legal decision making process.

The organization of the chapter is as follows: Section 2 presents a brief review of background concepts on BN and further reports on current implementation or adoption of BN in the public safety and crime domain. Section 3 describes methods of exploring the application of BNs in modeling a simple offender scenario. In section 4, a more complex scenario is presented alongside the types of knowledge that can be derived from the BN. The performance of the empirical analysis in legal decision support is also reported. Section 5 summarizes and concludes the chapter.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. BACKGROUND CONCEPTS AND RELATED STUDY

2.1 Theory and Notation of Bayesian Network Models

A Bayesian belief network (BBN) is a directed acyclic graph (DAG), that encodes probabilistic quantities to derive statistical inference based on probabilities (Jensen, 1996). Essentially, a BBN combines graph theory and probability theory to map variable interdependences in a stochastic system. Nodes in the network (graph) depict variables, while edges represent conditional dependencies between those variables (Vlek et al., 2013a). The DAG $\nabla=\{\chi,\gamma\}$ comprises a set of directed arcs γ that denote the dependencies among a set of nodes (variables) $\chi=\{X_1,\ldots,X_n\}$ such that:

$$\gamma=\left\{\left(X_i,X_j\right);X_j\in\chi\wedge X_i\neq X_j,j>i\right\}.$$

Given a set of discrete random variables $X_1\ldots X_n$ in the Bayesian network, their joint probability density distribution is mathematically defined as follows, in Equation 1:

$$P\left(\chi\right)=p\left(X_1,\ldots,X_n\right)=\prod P\left(X_i=x_i\;/\;Parents\left(X_i\right)\right) \tag{1}$$

where *Parents*(X_i) represent the values of the parents of the node X_i with respect to the graph. Bayesian probability, from Bayesian statistics, is the underlying concept in Bayesian networks. Bayesian probability is based on the assumption that probabilities are dependent on prior events or incidents. It then follows that in a Bayesian crime network, the strengths of the conditional crime dependencies would be calculated with the help of the Bayes' Rule equation. The probabilities $P(X_i)$ are derived from Bayes' theorem (Equation 2) and correspond to degrees of belief of a crime observation. Bayes' rule allows the computation of the probability of a crime event, given prior knowledge of a variable that the event is conditionally dependent on. As new information regarding a crime is obtained, the probabilities of certain characteristics of the unidentified subject are updated with Bayes' Rule.

Bayes' theorem is mathematically stated in Equation 2 as follows:

$$P\left(X\;/\;Y\right)=\frac{P\left(Y\;/\;X\right)P\left(X\right)}{P\left(Y\right)} \tag{2}$$

where *X* and *Y* are "effect" and "cause" events respectively.

- $P(X)$ and $P(Y)$ are independent probabilities of *X* and *Y*.
- $P(X/Y)$ is the conditional probability of *X* given that *Y* is true.
- $P(Y/X)$ is the conditional probability of *Y* given that *X* is true.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

BN application in a crime domain typically relies on four major steps, which are

1. Identify crime variables of interest, which become nodes in the network.
2. Specify prior probabilities over observable crime scenarios or incidents.
3. Compute posterior probabilities based on prior specifications, using Bayes' rule and all observed crime information.
4. Make predictions or inference according to the posterior quantities.

Figure 3 is a depiction of how crime evidence is influenced by hypothesis (i.e. extracting inference from crime observations or scenarios).

Given the following:

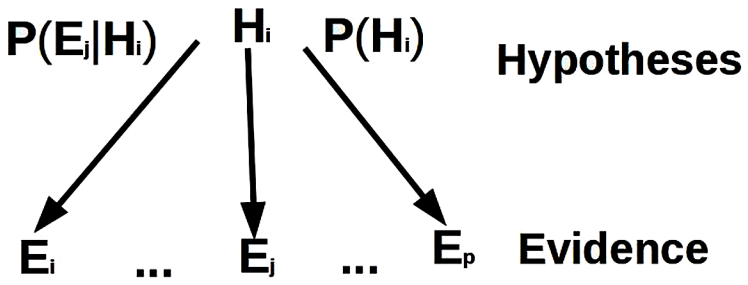
- Evidence/manifestations of crime: $E_1 \ldots E_p$,
- Hypotheses/observed crime features: $H_1 \ldots H_n$

Suppose we know the prior probability of the hypothesis $P(H_i)$ and the conditional probability $P(E_j/H_i)$ and want to compute the posterior probability $P(H_i/E_j)$, applying Bayes' theorem from Equation 2 results in Equation 3

$$P\left(H_i / E_j\right) = \frac{P\left(E_j / H_i\right) P\left(H_i\right)}{P\left(E_j\right)} \tag{3}$$

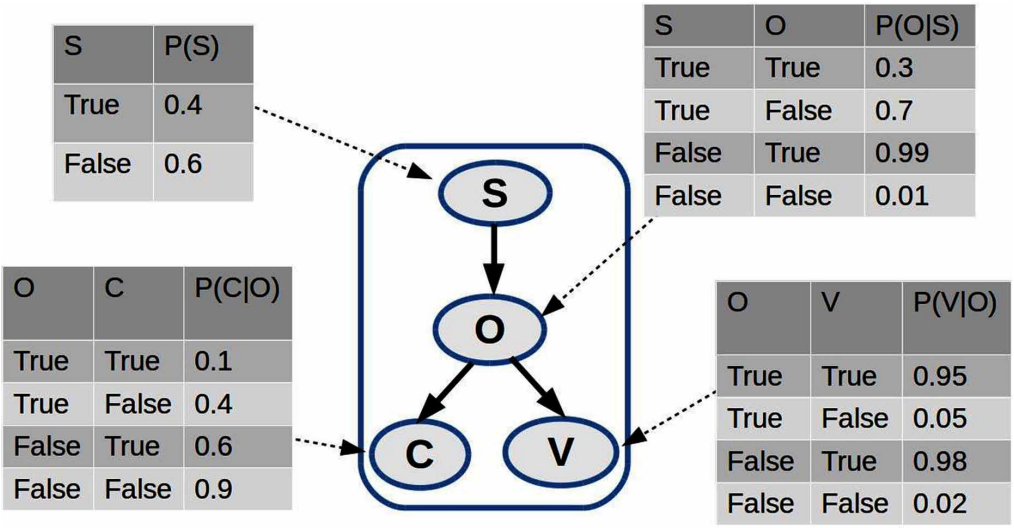
For a simple Bayesian investigative reasoning, the goal is to determine the hypothesis H_i with the highest posterior, i.e. $\text{Max}_i P(H_i/E_1 \ldots E_p)$. However, the simple Bayesian inference is limited in handling very large sets of crime information or data. In spite of their significant power and potential to address inferential procedures, there are some inherent limitations and liabilities to Bayesian networks (Fenton et al, 2014). It is computationally expensive, and sometimes specifying an appropriate or actual prior is not straightforward. Some researchers often just specify a convenient prior for computational purposes, not the actual prior. Further problem centers on the quality and extent of the prior beliefs used in Bayesian inference processing. A BBN is only as useful as the strength or reliability of such prior knowledge or belief. Either a pessimistic or an excessively optimistic expectation of the quality of these prior beliefs could distort the entire network and invalidate the results. Vlek et al. (Vlek et al, 2013c) noted that

Figure 3. Model setting of investigative Bayesian inference



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. A Bayesian network with CPTS for Figure 2



S	P(S)
True	0.4
False	0.6

S	O	P(O|S)
True	True	0.3
True	False	0.7
False	True	0.99
False	False	0.01

O	C	P(C|O)
True	True	0.1
True	False	0.4
False	True	0.6
False	False	0.9

O	V	P(V|O)
True	True	0.95
True	False	0.05
False	True	0.98
False	False	0.02

constructing a BN for complex scenarios is sometimes cumbersome and proposed a tool for exploring crime scenarios in a step-wise manner by building a BN in a step-by-step fashion. This notion exploits unfolding crime scenarios with variations. The idea only focuses on crime scenarios of interest at any given time of the investigative procedure and expands each scenario as crime event or evidence unfolds. Such an approach therefore results in a relatively more rigorous network of crime incident observations.

Using a Bayesian network to compute probabilities is called inference. Figure 4 presents a representation of Figure 2 with the conditional probability distribution for each variable of interest, where the names of the variables have been abbreviated to S = Suspect, Offender (O), Crime (C) and Victim (V). Each node X_i (in Figure 4) has a conditional probability table (CPT) representing its probability distribution $P(X_i/Parents(X_i))$, which quantifies the effect of the parents on the node or relationship between variables. The parameters are the probabilities in the CPTs. Therefore, for a given combination of values of the parents *O* in the analogy, the entries for $P(C=true/O)$ and $P(C=false/O)$ must add up to 1, satisfying one of the axioms of probability as shown in Equation 4 below (Vlek et al., 2013a).

$$P(C = true/O=false) + P(C = false/O = true) = 1 \quad (4)$$

In this case for example, *C* is the "evidence variable" while *O* is the query "variable". It is important to note that the probabilities in Figure 4 are chosen for illustrative purposes of a conditional probability table (CPT). Also, note that suspect information (depicted as "*S*" in the table) is the so-called "root node" with no parent(s), so there is only a single probability distribution for the node, *S*, while for nodes with parents there is a conditional distribution for each possible instantiation of its parents. Hence the varying CPTs in Figure 4.

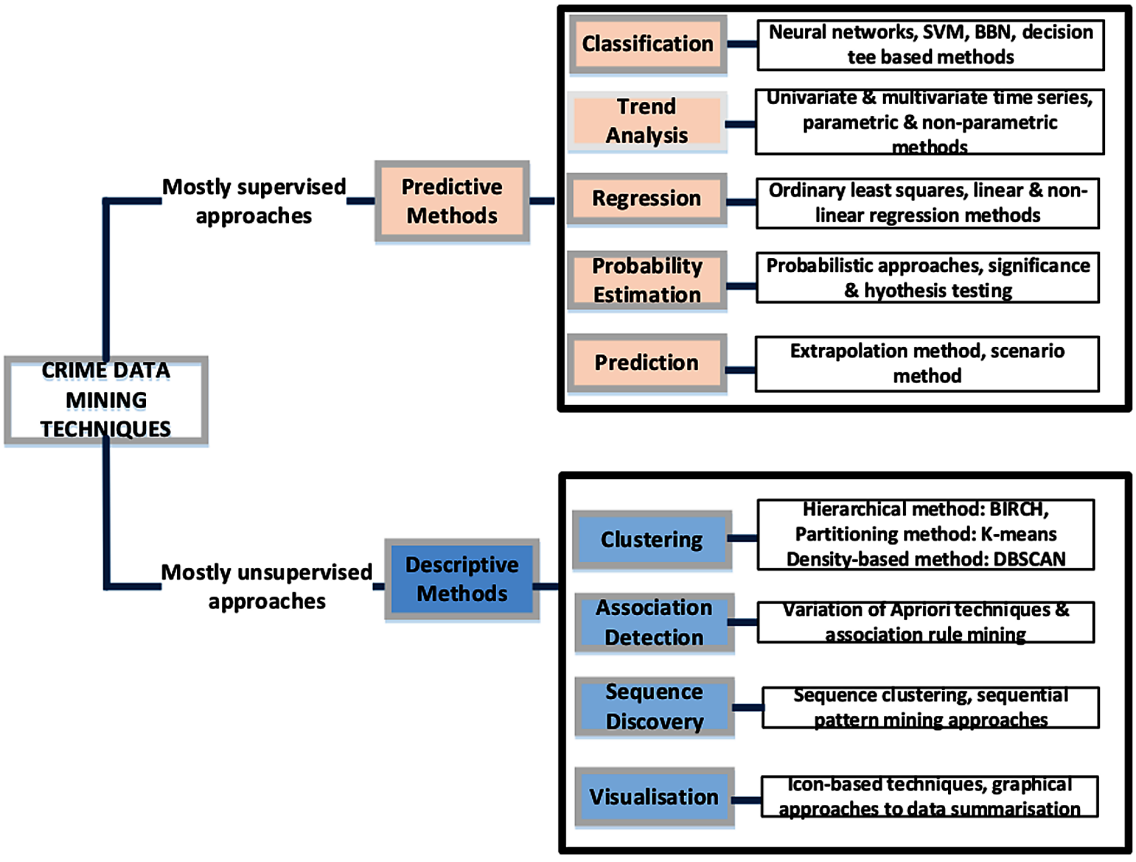
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.2 Practical Applications of BN Model in Crime Domain

There is a wealth of research in the field of crime data mining and with a number of different techniques and approaches for achieving different tasks in this field. Figure 5 presents different techniques or approaches of data mining available for exploration.

The strength of Bayesian Belief Networks (BBNs) as a powerful predictive probabilistic model (PPM) in the crime domain is evident in its different areas of real world application (Vlek et al, 2013b), which includes criminal profiling, forensic science, prediction of victimization characteristics, crime linkage and crime series detection, to mention a few. Furthermore, there exists a number of software technologies available for building and/or modeling BN (Korb et al, 2011). These include Netica, software used to program BNs in Java (Woodberry et al, 2015), CaMML (Causal discovery via Minimum Message Length), a machine learning program used for learning causal BNs from data (Chris et al, 1999), HUGIN model-based decision support software, and BayesiaLab software, to mention a few. It is important to note that most of these are however commercial (not free) software. There are a few open source BN tools available, such as the BN implementation in the Weka Machine Learning Workbench for data mining (Hall et al, 2009). We hereafter present a non-exhaustive summary of existing application areas of BN in the crime domain.

Figure 5. A taxonomy of crime data mining techniques



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.2.1 Criminal Profiling

BNs offer several advantages in criminal profiling and received significant attention in this area (Stahlschmidt et al, 2013). Criminal profiling deals with identifying the attributes or behavioural characteristics of an unknown offender, what motivates the crime committed and principal personality from a crime scene (Baumgartner et al, 2008). As such it is an important consideration in crime investigation since it provides useful gain in psychological and social profiling of the suspect. This generally promotes crime detection and criminal identification (CDCI), since identified characteristics of the offender can narrow the number of potential suspects and reduce investigative efforts by excluding those not showing the specific traits inferred. For example, if a crime investigator initially has three suspects on his list to investigate, and thereafter realises that only two of the suspects match all the crime observations he has gathered to date, he can easily focus more of his attention on those two whose characteristics correlate with his current observations.

Security agencies do background work by filtering or sniffing through a crime scene for observational evidence, because they have some degree of belief that careful observation and analysis of the crime scene can produce useful information that could help apprehend an unknown offender or understand "why" such crime occurred. Essentially, the three major approaches to modeling reasoning with evidence are argumentative, narrative and probabilistic approaches. The work of Vlek et al. (Vlek et al, 2014) reports that integrating these three approaches could arguably enhance the communication between an expert and a judge or jury. Their approach combined the three techniques (argumentative, narrative and probabilistic) in order to form a design method for constructing a Bayesian network based on narratives. Related research has however proposed techniques to represent narratives in a Bayesian network and to use narratives as a basis for systematizing the construction of a Bayesian network for a legal case (Sjoerd et al, 2015).

While Stahlschmidt et al (Stahlschmidt et al, 2013) report on using BN for criminal profiling and prediction in sex-related homicides, the work of Baumgartner et al. present a novel Bayesian network approach to criminal profiling (Baumgartner et al, 2008). Their approach learns a BN model from crime data, which is one of the options for constructing a BN, and then implements the model for profiling means. The variables used as nodes include "offender" and "evidence" variables. The offender variables are characteristics of the offender, which includes attributes such as military history and sex crime history. Evidence variables are characteristics of the crime scene, such as possibility of committing arson, stabbing the victim, and so on. The strengths of edges between nodes are calculated using a combination of a database of offender and crime data and expert knowledge from actual criminal profilers. The more data, the more iterations of Bayes' rule can be applied, and the more accurate the interdependences between offender variables and evidence variables. Once the preliminary conditional probabilities have been calculated, the network can be instantly used to narrow down criminal profiles in unsolved cases in an automated structure. Once evidence/crime scene data are supplied as input, characteristics/attributes that are closely correlated to the variables of interest are observed. Conditional probabilities are re-computed based on new crime data. Their research presents a significant use of BN and Bayes' rule in real world scenarios.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.2.2 Prediction of Victimisation Characteristics

Riesen et al. (Riesen et al, 2009) developed a Bayesian network classifier for predicting a victimisation attribute value from a crime data set. They identified that a scoring function, suitable structure-learning algorithm and parameter learning algorithm are necessary in order to induce a Bayesian network from a dataset. The dataset they considered was the National Crime Victimisation Survey dataset with 216,000 and 250 data points and attributes respectively. Their model implemented the local hill-climbing and k2 algorithms in the parameter and structural learning phase using the Weka machine learning software workbench, while ensuring reasonable bounds for time-space complexity (Hall et al, 2009).

Related work of Axelrad et al (Axelrad et al, 2013) focuses on predicting insider threat using a BN model. Their paper which consists of five different phases reports on a Bayesian network model for the prediction of level of interest in a malicious insider. Authors investigate the motivation and psychology of the malicious insider by first developing an initial model based on results in the research literature, thereby highlighting essential variables for the prediction of level of interest in a potentially malicious insider. Second, a survey was conducted to measure these predictive variables in a regular sample of normal participants. Third, the original model was used to construct a structural equation model. The model was updated based on a split-half sample of the empirical survey data and validated against the remaining half of the dataset. Fourth, the Bayesian network was updated in view of the results of the empirical analysis. Lastly, the updated model was used to develop an upper limit on the quality of model predictions of its own simulated data. When empirical data regarding psychological predictors were input to the model, predictions of counterproductive behavior approached the upper bound of model predictiveness.

To validate the BN constructed, a hypothesized structural equation model (SEM) was developed, which incorporated as many of the variables from the Bayesian network as it was feasible to measure. Then the data was split in half, randomly assigning cases to a development sample and a holdout sample. SEM was then fit to the first data sample and model adjusted in light of the results so as to test it on the holdout sample. Data was then aggregated to fit a final structural model.

2.2.3 Crime Linkage

Crime linkage focuses on identifying possibly related crime incidents in relation to perpetrator attribute characteristics, different (incriminating) crime scene observations such as Modus Operandi, and victim information. When two distinct similarities are identified in two or more different crimes, it becomes of interest to know whether the crimes were committed by the same culprit or group of culprits. Bayesian networks can be used to model varying evidential structures in linking two or more crimes using different crime attributes as metrics.

Space and time metrics have been considered as parameters for estimating crime linkage. Tonkin et al (Tonkin et al., 2012) used crime location and time between two crime events as yardsticks for identifying linked and unlinked crime cases. They focused on how far apart (in time) were two crimes committed or length of space between two related crimes. They resolved that temporal and/or spatial proximity and/or crime incident locations can achieve statistically significant levels of discrimination between crimes that are linked and unlinked. A related attempt by Zoete et al (Zoete et al, 2014) also reports on modeling crime linkage using Bayesian networks. Their work is motivated by the lack of adequate mathematical treatment in defining the degree of crime similarity in practical casework. They

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

investigated how evidential information derived in one case can influence the probability of hypotheses about who the offender is in another case. It was reported that the usual procedure in law enforcement where only cases with evidence pointing to a suspect are considered and other cases discarded, could be misleading because those discarded cases can generate evidence with important practical consequences if considered.

2.3 Limitations and Challenges of Bayesian Network

A BN has a significant computational cost (an NP hard task) because one needs to explore all parent crime nodes or variables, in order to explore an unknown crime network. Missing dependencies could affect crime results or evidence provided. It potentially requires initial knowledge or prior probabilities of certain crime variables, which may be extremely difficult to obtain and is sometimes falsified. Furthermore, it is sometimes insufficient to derive conclusions for a specific case from an initially developed network (Vlek et al, 2013c). Often there is a need to further explore sub-scenarios of crime in order to update the developed model or network and uncover new crime evidence, especially in a very complex BN model. Furthermore, since each crime is unique, there is a need to investigate how useful a scheme for a general situation is, when dealing with a peculiar crime situation.

3.0 BAYESIAN NETWORKS ADOPTION IN OFFENDER MODELLING

In this section we present a "thought experiment" on how a BN can be used to model a simple offender's behaviour and present a scenario of offender modelling in the subsequent section. We focus more on the logic of the argumentation which presents the essence of the problem and application of BN in reality. There exist various methods of obtaining Bayesian Network (e.g have an expert design it or learn it from data), but that is not the focus of this chapter. We refer the reader to comprehensive textbooks on Bayesian networks for more theoretical arguments or information on BN (Korb et al, 2011 & Woodberry et al, 2015).

3.1 A Simple Case of Offender Modelling

One of the major goals of offender modeling is to identify mitigation priorities in the crime investigation procedure. When related characteristics of an offender are identified it helps to narrow down an investigative procedure by eliminating suspects with characteristics unrelated to those in the identified profile list. Suppose two separate crimes have been committed, not necessarily by the same offender, and you are trying to determine whether an individual is a serial criminal who committed both (Evett at al, 2006 & Zoete, et al, 2013). You observe the following (how likely a suspect is a serial offender given certain conditions, for example prior offences, prior arrests):

- The suspect appears to have previously committed "crime "A".
- The suspect appears to have previously committed "crime "B".
- The suspect appears to have previously committed "crime "C".

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Note that we are not 100% certain that the person is a serial offender because of these "observations". Thus, we are dealing with uncertainty at this stage. While there is no similarity evidence at this point of observation, one would assume that knowledge of the culprit or offender in case A would not necessarily affect the uncertainty related to case B or C.

Consider the simple Bayesian offender modeling (from Figure 2). Given a set of observations (evidence (E)), we have the following hypothesis for a single crime scenario:

H1: The suspect (S) is the offender (O).
H2: The suspect (S) is not the offender (O).

3.1.1 Properties of a Hypothesis

1. Must be mutually exclusive (non-overlapping).
2. Must consider all possible scenarios (i.e. exhaustive).

In a "two-crime scenario" (crime A and crime B), however, the following hypothesis is considered (Zoete et al, 2014).

H1: The suspect (S) is the offender (O) in both crime cases.
H2: An unknown person (U) is the offender (O) in both cases.
H3: An unknown person (U) is the offender in the first case and the suspect is the offender in the second case.
H4: The suspect is the offender in the first case and an unknown person (U) is the offender in the second case.
H5: An unknown person (U) is the offender (O) in the first case and another unknown person (U') is the offender (O) in the second case.

In any case, the goal is to find the hypothesis (H_i) with the highest posterior probability. The resulting output further strengthens, weakens or nullifies the *a priori* belief regarding the initial hypothesis considered. In light of the fact that BN is a PPM, it has the capability of approximating the posterior probability distribution of any attribute of interest in a crime domain (Riesen et al, 2009).

In general, each of the five hypotheses (crime propositions) has a unique probability. In theory the probabilities must sum up to one. However, one is much more interested in how the probabilities would change as they are updated with corresponding evidence. The question that follows then is "how does the *posterior* probability $P\left(H_1 \,/\, E, A\right) = k'$ change or evolve as it is updated with the computation from the prior probability $P(H_1/A)=k$, where H_1 is the hypothesis, A is the crime scenario considered and E is the evidence.

Going by Bayes' rule, we obtain the following equation:

$P(H_1/E,A) \propto p(E/H_1,A)P(H_1/A)$.

The same expression could be derived for the remaining joint propositions or hypotheses. For example, if the *posterior* value $k' > k$, then it becomes clear that there is a higher chance that hypothesis

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

H_1 is true, i.e. the suspect is most likely to be the offender in both cases. The *posterior* probability is being updated as new evidence is derived, computed and updated; eventually an effective decision would result.

It is important to note that more work needs to be done in terms of modeling complex scenarios of crime using BN. Such effort has been demonstrated in the work of Vlek et al (Vlek et al, 2013c), which deals with unfolding (sub) crime scenarios in an attempt to make a scenario more precise. This precision is achieved by modeling several connected sub-scenarios to enhance the network and derive more inclusive evidence. They implemented a three-stage approach idiom which exploits "scenario", "sub-scenario" and "merged-scenario" idioms, in order to further complement and strengthen the derived evidence. This is one of the ways by which inference could be extracted from observations in a BN, which greatly relies on prior belief and observations from a crime scene or about a specific crime. Depending on the level or strength of the incriminating evidence, the jury may then be able to reach one of the following decisions:

1. Charge the suspect with case 1; drop case 2.
2. Charge the suspect with case 2; drop case 1.
3. Charge the suspect with both cases.
4. Drop both cases.

3.2 Complexity of Inference in Bayesian Belief Networks

Inference in a BBN corresponds to calculating the conditional probability, $P[A/B=b]$, where $A, B \subset \left\{X_1, \ldots, X_n\right\}$ are sets of latent and observed crime variables, respectively. Cooper (Cooper, G. F., 1990) proved that exact inference in BBNs is NP-hard. In general, the size of the problem is given by the total size of the probability tables needed to represent the BN. Dagum and Luby (Dagum, P., Luby, M.,1993) further showed that even approximate inference is NP-hard (note: the mathematical detail of the proof is not explicitly highlighted in this chapter).

There exist three main inference tasks for BBNs. These include:

1. **Inferring Unobserved Variables:** Since BN is a PDAGM, it can be used to answer probabilistic queries about variables and their relationships. It therefore helps in deriving probabilistic inference, which is a process that derives updated knowledge of the state of a subset of crime variables when other variables (the *evidence* variables) are observed.
2. **Parameter Learning:** It is necessary to specify for each node *X* the probability distribution (CPT) for *X* conditional upon *X*'s parent, in order to fully specify the BBN and therefore completely represent the joint probability distribution of the random variable. The conditional distributions sometimes include unknown crime parameters which must be estimated from the available crime data, sometimes using the maximum likelihood approach.
3. **Structure Learning:** The structures of BBNs are either learned from data or specified by an expert in the corresponding domain and afterwards used to perform inference. In applications where the task of defining the network is too complex for humans, the network structure and the parameters of the local distributions are usually learned from data. An alternative method of structural learning uses optimization based search, which requires a scoring function and a search strategy. A typical scoring function is *a posterior* probability of the structure given the training crime data.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.0 COMPLEX SCENARIOS OF BAYESIAN LEARNING AND OFFENDER MODELLING

In this section, we present relatively more complex scenarios of offender modeling using Bayesian networks and corresponding implications for deriving valid results.

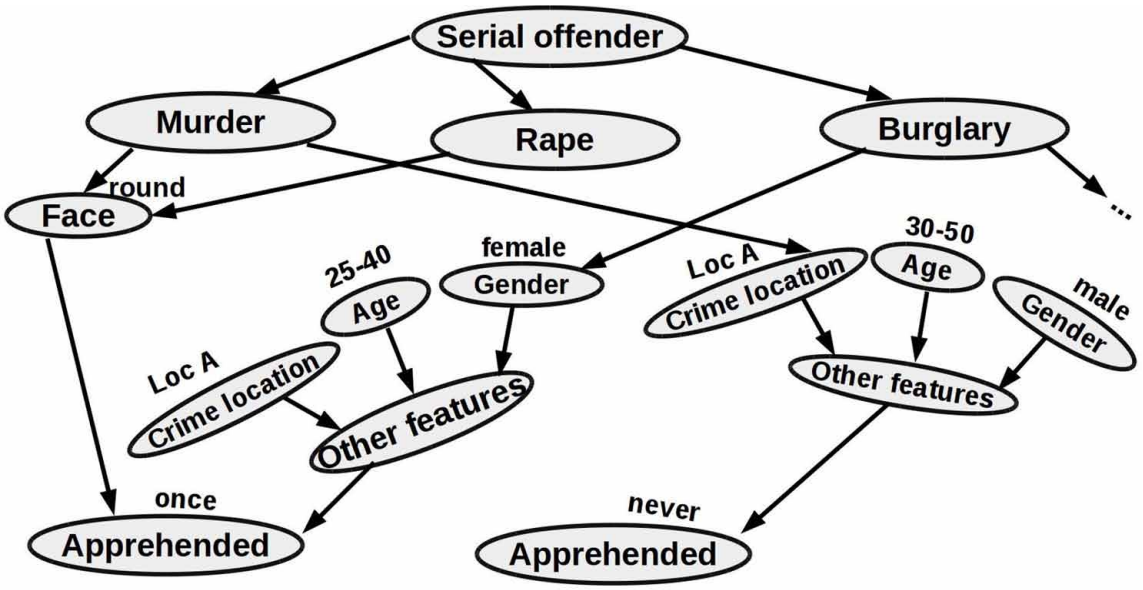
More generally, BN probabilistic reasoning can result in a powerful combination of predictive, diagnostic and explanatory reasoning. Considering the fact that several levels of scenarios could be explored in the figurative explanation of a relatively complex BN (Figure 6), which may also have different levels of precision, it is worth noting that the validity of evidence that can be derived would to a large extent depend on the level of this exploration.

The general approach to a complex crime scenario or offender information, such as in Figure 6 usually considers evaluating different kinds of evidence. Such evidence includes information from crime attractors (e.g. vulnerable victims), eye witnesses, scientific processes and results and alibis, to mention a few. Most of the evidence is usually not quantified but could be used by the jury to follow a quantitative means of reaching a fair and effective decision, among several.

The BN in Figure 6 can be used to derive useful inference such as the following:

1. Can combine evidence from various crime attributes.
2. Can predict a crime suspect from the observational crime features.
3. Can explain the model's belief in a crime investigation scenario.
4. Can identify specific characteristics of suspects based on the nature of their crimes.
5. Can model information such as the crime spatial dispersion pattern and suspect apprehension.

Figure 6. A more complex Bayesian network of suspect modelling



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.1 Generating Candidate and Extraction Rules

Candidate items in our "thought experiment" refer to attributes of crime incidents of interest, which a jury or crime analyst is very much interested in investigating or probing further. Consider five different crime scenarios or types as follows: Crime A, Crime B, Crime C, Crime D, Crime E, with Figure 7 depicting how each crime might be linked to a suspect.

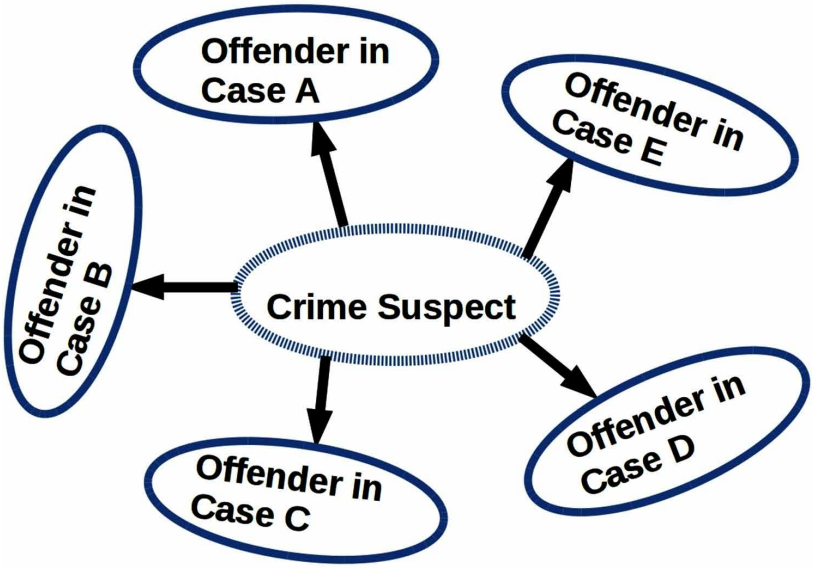
Now, for each scenario, suppose we are interested in how many crimes relate to a particular suspect (e.g identifying a serial offender), assuming we already have prior knowledge of certain incriminating evidence or observations from each crime scene. In real life scenarios, one may observe that residential burglars tend to return to previously victimized locations, or to their close vicinities, after having acquired information on the schedules of inhabitants, the properties, and possible surveillance systems. For example, in Figure 6, a once apprehended offender may be suspected of another related crime in a location of close proximity to the previous crime.

However, depending on how much incriminating evidence is attributed to a suspect, the jury may then be able to make one or more of the following decisions:

- The suspect is not guilty of any of the crimes.
- The suspect is guilty of only one crime, say crime k, $k \in \{A,B,C,D,E\}$.
- The suspect is guilty of only two of the crimes, say crimes A and B, etc.

We could model possible scenarios of allegations that can be made to a suspect in a 5-crime scenario as presented in Table 1. Since there are five crimes, then there are $S=2^5$ possible states related to the suspect in question. Variables A,B,C,D,E depict the five different crime types thought to have been committed by the suspect, S_i, $i \in \{1,2,\ldots,32\}$.

Figure 7. BN model for crime suspect decision support in real life scenarios



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Possible scenarios of suspect allegations in legal decision making

	A	B	C	D	E	GWM
S1	N	N	N	N	N	N/A
S2	N	N	N	N	G	1/5
S3	N	N	N	G	N	1/5
S4	N	N	N	G	G	2/5
S5	N	N	G	N	N	1/5
S6	N	N	G	N	G	2/5
S7	N	N	G	G	N	2/5
S8	N	N	G	G	G	3/5
S9	N	G	N	N	N	1/5
S10	N	G	N	N	G	2/5
S11	N	G	N	G	N	2/5
S12	N	G	N	G	G	3/5
S13	N	G	G	N	N	2/5
S14	N	G	G	N	G	3/5
S15	N	G	G	G	N	3/5
S16	N	G	G	G	G	4/5
S17	G	N	N	N	N	1/5
S18	G	N	N	N	G	2/5
S19	G	N	N	G	N	2/5
S20	G	N	N	G	G	3/5
S21	G	N	G	N	N	2/5
S22	G	N	G	N	G	3/5
S23	G	N	G	G	N	3/5
S24	G	N	G	G	G	4/5
S25	G	G	N	N	N	2/5
S26	G	G	N	N	G	3/5
S27	G	G	N	G	N	3/5
S28	G	G	N	G	G	4/5
S29	G	G	G	N	N	3/5
S30	G	G	G	N	G	4/5
S31	G	G	G	G	N	4/5
S32	G	G	G	G	G	5/5 = 1

Variables *G* and *N* in Table 1 represent *guilty*:*G* and *not guilty:N* respectively. Note that certain allegation scenarios states S_i, $i \in \{1,2,\ldots,32\}$ appear to share the same value in terms of general weight measure (GWM), where $0 < GWM \leq 1$. They may however not necessarily share the same aggregate weight measure (AWM). There are two "exceptional" cases S_1 having no weight since the suspect is not guilty of any of the (five) crimes and S_{32} where the suspect is guilty of all allegations. For example scenarios

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

S_2, S_3, S_5, S_9, S_{17} have a general weight measure (GWM) of 1 in 5 crimes. However, suppose the unique weight (w) for each crime is depicted as W_A, W_B, W_C, W_D, W_E respectively. Then each scenario S_i would have different weight values. A special case would then be if all the crimes have the same weight measure, in which there is a high probability that a suspect has committed a crime repeatedly five times. In summary, we have the following states or scenarios sharing the same GWM as presented in Figure 8 and Equations 5, 6, 7, and 8. However, note that every scenario has only one GWM.

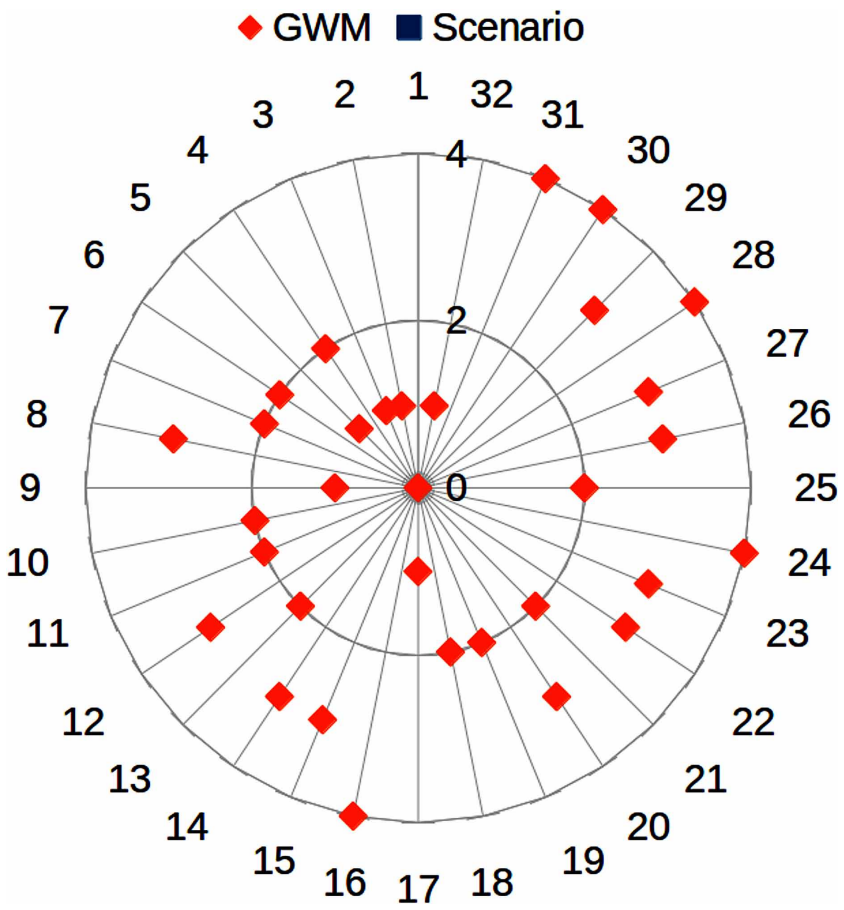
$$1/5 = \{S_2, S_3, S_5, S_9, S_{17}\} \tag{5}$$

$$2/5 = \{S_4, S_6, S_7, S_{10}, S_{11}, S_{13}, S_{18}, S_{19}, S_{21}, S_{25}\} \tag{6}$$

$$3/5 = \{S_8, S_{12}, S_{14}, S_{15}, S_{20}, S_{22}, S_{26}, S_{27}, S_{29}\} \tag{7}$$

$$4/5 = \{S_{16}, S_{24}, S_{28}, S_{30}, S_{31}\} \tag{8}$$

Figure 8. Correlated GWM of possible scenarios of suspect allegation



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Thus the Aggregate Weight Measure (AWM) for each state S_i is computed as the sum of the corresponding weights $c_W \in \{W_A, W_B, W_C, W_D, W_E\}$ for which the suspect is guilty (G), as depicted in Equation (9)

$$AWM_{\{S_i\}} = \sum_{G_i} \{C_W\} \tag{9}$$

The weight information could be regarded as indicating the gravity of the crime committed, which is further defined by the AWM for every suspect scenario S_i, especially for a serial offender. The suspect could then be charged for correspondingly identified scenario(s) accordingly. It is important to recognise that the reported idea does not eliminate or exclude the jury's intuition but serves to complement their overall decision.

5. SUMMARY AND CONCLUDING REMARKS

The public safety agencies collect a great deal of crime information which can be explored for evidence-based solutions in deterring crime. Also legal cases involve, to a great extent, reasoning with evidence. Thus, employing Bayesian networks (BNs) in the crime domain is a recognised application that can model different crime scenarios and assist in mitigation priorities or investigative procedures for solving crime, since it is based on probabilistic reasoning.

The strength of BNs as a Predictive Probabilistic Model (PPM) is evident in its wide acceptance and application in solving real world problems, both in academia and industry. A BN is particularly useful for classification and prediction tasks and has thus received significant attention in several domains, including the crime domain. This chapter has described various means by which BNs have been employed in the crime mining domain. The various applications areas discussed include crime suspect modeling, crime linkage, criminal profiling and repeat victimisation studies, which can assist in understanding criminal identification or apprehension, crime control and legal decision processes.

This chapter further highlighted the strengths and limitations of BNs and also reported a "thought experiment" on simple and "near-complex" offender modeling cases using BN probabilistic reasoning, with discussions of the logic of evidence interpretation and decision making. While BNs have been explored in different capacities in the crime domain, it is important to recognise that there is still room for exploring its capability in modeling complex crime scenarios and advancing the strength of the technique, as it has some inherent limitations such as computational difficulty in its current application. It is worth noting that inference from BNs is quite convenient and feasible in small to medium sized networks and could be very time consuming in large networks. However, researchers resort to various techniques for approximating inference, which are more reliable and faster.

Also the resulting BN model is not intended to make final decisions or conclusions about a crime case, but to assist in the investigative process and advise or assist a jury or judge in legal decision making. This means that we do not propose eliminating the judge's instinct and experience in legal decision making and crime control. We however present the fact that judge's instinct and experience can be complemented and strengthened using BN, as presented in current applications of BN in the crime domain. There is a great deal more to BNs when building an actual network of any complexity, which can be studied in depth in Bayesian Network focused textbooks. This work illustrates the future potential and opportuni-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ties of applying Bayesian Networks in this field and can be explored further in different forms in future practical research. This chapter however could further serve as a reference guide to understanding key issues and underlying framework of BN application in Crime Detection and Criminal Identification (CDCI) or crime mining field in general.

LIST OF NOTATIONS

BBN: Bayesian Belief Network
BN: Bayesian Network
CaMML: Causal discovery via MML
CPT: Conditional Probability Tables
CDCI: Crime Detection and Criminal Identification
DAG: Directed Acyclic Graph
DBN: Dynamic Bayesian Network
EDA: Exploratory Data Analysis
LDMP: Legal Decision Making Process
MO: Modus Operandi
PDAGM: Probabilistic Directed Acyclic Graphical Model
PPM: Predictive Probabilistic Model
RVS: Repeat Victimisation Study
SCDA: Spatial Crime Data Analysis

ACKNOWLEDGMENT

The authors gratefully appreciate the financial support from an AIMS-DAAD Fellowship, L'oréal-UNESCO Women in Science Fellowship and resources made available by the University of Cape Town and the University of the Western Cape.

REFERENCES

Axelrad, E. T., Sticha, P. J., Brdiczka, O., & Shen, J. (2013). A Bayesian Network Model for Predicting Insider Threats. In *Proceedings of Security and Privacy Workshops* (pp. 82–89). IEEE. doi:10.1109/SPW.2013.35

Baumgartner, K., Ferrari, S., & Palermo, G. (2008). Constructing Bayesian Networks for Criminal Profiling from Limited Data. Journal of Knowledge-Based Systems. doi:10.1016/j.knosys.2008.03.019

Baumgartner, K., Ferrari, S., & Salfati, C. (2005). Bayesian Network Modeling of Offender Behavior for Criminal Profiling. In *Proceedings of the 44th IEEE Conference on Decision and Control*. 10.1109/CDC.2005.1582571

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chris, S. W., & Kevin, B. K. (1999). Learning Linear Causal Models by MML Sampling. In *Causal Models and Intelligent Data Management* (pp. 81–11). Springer-Verlag.

Clarke, V. R., & Felson, M. M. (1993). Introduction: Criminology, Routine Activity, and Rational Choice. *Advances in Criminological Theory Routine Activity and Rational Choice*, *5*, 1–14.

Cooper, G. F. (1990). The computational complexity of probabilistic inference using Bayesian belief networks. *Artificial Intelligence*, *42*(2-3), 393–405. doi:10.1016/0004-3702(90)90060-D

Dagum, P., & Luby, M. (1993). Approximating probabilistic inference in Bayesian belief networks is NP-hard. *Artificial Intelligence*, *60*(1), 141–153. doi:10.1016/0004-3702(93)90036-B

Evett, L. W., Jackson, G., Lindley, D. V., & Meuwly, D. (2006). Logical Evaluation of Evidence When a Person is Suspected of Committing two Separate Offences. *Science & Justice*, *46*(1), 25–31. doi:10.1016/S1355-0306(06)71564-7 PMID:16878782

Fenton, N. E., Neil, M., & Hsu, A. (2014). Calculating and understanding the value of any type of match evidence when there are potential testing errors. *Artificial Intelligence and Law*, *22*(1), 1–28. doi:10.100710506-013-9147-x

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). "The WEKA data mining software: An update". In ACM Digital Library. *SIGKDD Explorations*, *11*(1), 10–18. doi:10.1145/1656274.1656278

Jayech, K., & Mahjoub, M. A. (2010). New approach using Bayesian Network to improve content based image classification systems. *International Journal of Computer Science Issues.*, *7*(6), 53–62.

Jensen, F. V. (1996). *Introduction to Bayesian Networks*. New York: Springer.

Korb, K. B., & Nicholson, A. E. (2011). Bayesian Artificial Intelligence. Chapman & Hall/CRC.

Milgram, A. (2013). Why smart statistics are the key to fighting crime. *TED Video*. Retrieved from https://www.ted.com/talks/anne_milgram_why_smart_statistics_are_the_key_to_fighting_crime

Nazerfard, E., & Cook, D. J. (2013). *Using Bayesian Networks for Daily Activity Prediction*. Association for the Advancement of Artificial Intelligence. Retrieved from www.aaai.org

Nefian, A. V., Liang, L., Pi, X., Liu, X., & Murphy, K. (2002). Dynamic Bayesian Networks for Audio-Visual Speech Recognition. *EURASIP Journal on Applied Signal Processing*, *11*, 1–15.

Papoulis, A. (1984). Bayes' Theorem in Statistics and "Bayes' Theorem in Statistics (Reexamined). In Probability, Random Variables, and Stochastic Processes (2nd ed.). New York: McGraw-Hill.

Predpol. (2013). *PredPol Predicts Gun Violence*. Retrieved from http://cortecs.org/wp-content/uploads/2014/10/predpol_gun-violence.pdf

Raghuram, S., Xia, Y., Palakal, M., Jones, J., Pecenka, D., Tinsley, E., ... Geesaman, J. (2009). Bridging Text Mining and Bayesian Networks. In *Proceedings of the IEEE International Conference on Network-Based Information Systems*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Riesen, M., & Serpen, G. (2009). A Bayesian Belief Network Classifier for Predicting Victimization in National Crime Victimization Survey. In *Proceedings of the IEEE International Conference on Artificial Intelligence*.

Sjoerd, T., Charles, M., Hendrik, P., Silja, R., & Bart, V. (2015) A Structure-guided Approach to Capturing Bayesian Reasoning about Legal Evidence in Argumentation. In *Proceedings of the ICAIL international conference.* ACM.

Sjoerd, T., Hendrik, P., Charles, M., Silja, R., & Bart, V. (2014) Extracting Legal Arguments from Forensic Bayesian Networks. In *Legal Knowledge and Information Systems. JURIX 2014: The Twenty-seventh Annual Conference*.

Stahlschmidt, S., Tausendteufel, H., & Härdle, W. K. (2013). Bayesian Networks for Sex-related homicides: Structure Learning and Prediction. *Journal of Applied Statistics*, *40*(6), 1155–1171. doi:10.1080/02664763.2013.780235

Tonkin, M., Woodhams, J., Bull, R., & Bond, J. W. (2012). Behavioural Case Linkage with Solved and Unsolved Crime Crimes. *Forensic Science International*, *222*(1-3), 146–153. doi:10.1016/j.forsciint.2012.05.017 PMID:22704669

Vlek, C., Prakken, H., Renooij, S., & Verheij, B. (2013a). Modeling Crime Scenarios in a Bayesian Network. *The 14th International Conference on Artificial Intelligence and Law (ICAIL 2013). Proceedings of the Conference*, (pp. 150-159). New York: ACM.

Vlek, C., Prakken, H., Renooij, S., & Verheij, B. (2013b). Representing and Evaluating Legal Narratives with Sub-scenarios in a Bayesian Network. *Workshop on Computational Models of Narrative*. OASICS Schloss Dagstuhl Publishing.

Vlek, C., Prakken, H., Renooij, S., & Verheij, B. (2014). Building Bayesian networks for legal evidence with narratives: a case study evaluation. Journal of Artificial Intelligence and Law, 22(4), 375-421. doi:10.100710506-014-9161-7

Vlek, C. S., Prakken, H., Renooij, S., & Verheij, B. (2013c). Unfolding Crime Scenarios with Variations: A method for building a Bayesian network for legal narratives. In Legal Knowledge and Information Systems - JURIX.

Woodberry, O., & Mascaro, S. (2015). *Programming Bayesian Network Solutions with Netica*. Bayesian Intelligence. Retrieved from http://bayesian-intelligence.com/programming-bns/

Zoete, J. D., Sjerps, M., Lagnado, D., & Fenton, N. (2014). Modelling Crime Linkage with Bayesian networks. *Journal of Science and Justice, Elsevier, Ireland*, *7*(5), 1–9. PMID:25934374

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

Bayesian Network: A directed acyclic graph that encodes probabilistic reasoning of an event or a compact representation of probability distribution over a set of propositional random variables.

Crime Linkage: The process of verifying whether two or more crimes, say A and B, committed at possibly separate times and/or locations are linked by some related features.

Criminal Profiling: The process of identifying behavioural characteristics of a crime offender and motivations for committing the crime.

Crime Pattern: An observable recurring trend of crime attributes or characteristics.

Crime Tactical Analysis: A crime analysis that is carefully planned in order to combat a specific crime.

Dynamic Bayesian Networks (DBNs): These are Bayesian networks that include directed edges pointing in the direction of time.

Modus Operandi: This characterises the process or mode of operation of an offender or group of offenders.

Probabilistic Reasoning: The process of summarising uncertainty from various sources or using probability theory and information about interdependencies to derive statistical inference.

Serial Crime: A set of two or more related crimes identified to have been committed in a successive time interval and mostly within close proximity.

Serial Offender: An offender that has been identified to commit related crime offences in a successive time interval, mostly with the same Modus Operandi.

Victimization Prediction: The process of identifying principal attributes, characteristics or features that surrounds a crime victim and using such observation to identify potential victims.

This research was previously published in Data Mining Trends and Applications in Criminal Science and Investigations edited by Omowunmi E. Isafiade and Antoine B. Bagula; pages 143-168; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 20
Information Extraction for Call for Paper

Laurent Issertial
Osaka Prefecture University, Japan

Hiroshi Tsuji
Osaka Prefecture University, Japan

ABSTRACT

This paper proposes a system called CFP Manager specialized on IT field and designed to ease the process of searching conference suitable to one's need. At present, the handling of CFP faces two problems: for emails, the huge quantity of CFP received can be easily skimmed through. For websites, the reviewing of some of the main CFP aggregators available online points out the lack of usable criteria. This system proposes to answer to these problems via its architecture consisting of three components: firstly an Information Extraction module extracting relevant information (as date, location, etc...) from CFP using rule based text mining algorithm. The second component enriches the now extracted data with external one from ontology models. Finally the last one displays the said data and allows the end user to perform complex queries on the CFP dataset and thus allow him to only access to CFP suitable for him. In order to validate the authors' proposal, they eventually process the well-known precision / recall metric on our information extraction component with an average of 0.95 for precision and 0.91 for recall on three different 100 CFP dataset. This paper finally discusses the validity of our approach by confronting our system for different queries with two systems already available online (WikiCFP and IEEE Conference Search) and basic text searching approach standing for searching in an email box. On a 100 CFP dataset with the wide variety of usable data and the possibility to perform complex queries we surpass basic text searching method and WikiCFP by not returning the false positive usually returned by them and find a result close to the IEEE system.

DOI: 10.4018/978-1-7998-0951-7.ch020

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

With today's technology any type of information accessibility has greatly increased and this is also true for our domain of study: Call For Papers (CFP). It has never been so easy to find information on conference usually via numerous websites or simply by email. However, this profusion of information has also its downside. Handling CFP can be seen as a really time consuming task. Regarding the emails, many CFP are received every day and most of them will only be skimmed through or forgotten by lack of time or attention. As for the web, it is possible to find many websites where you can browse a wide range of CFPs such as WikiCFP (2014) or ConferencePartner (2014), or others bound to different institutions like IEEE (2014) or ACM (2012). Nevertheless these sites have several weak points. They do not allow the user to use CFPs obtained by external sources (e.g. Email) ; thus a new research have to be done on a totally different database, sometimes with less information you could actually find in your CFP email. Moreover, conferences research parameters can be really basic and not handle a request more complex than searching different strings of characters (e.g. conference about X topic in Y country with a Z deadline). Those are the issues we propose to handle with our system: CFP Manager.

In one of our previous work (Issertial & Tsuji, 2011), we proposed the concept of a text mining system able to extract relevant information from a group of CFP. In another one concerning visualized comparison of CFP datasets (Issertial, Saga, & Tsuji, 2012), we introduced the idea of extracted data enrichment via the utilization of ontology, another one (Issertial & Tsuji, 2013) focused on the query and interface system along with the different kind of output that can be obtained from the set of relevant data previously collected. This paper proposes to focus on the case of a final user who's looking for information about conferences suited to him and how our proposed system will ease this process. We will do so by reviewing current web-services proposing this functionality and comparing their theoretical results with our proposed system on different queries. This will be introduced along with the description of our system and numerical experimentations proving its good working. CFP will be subject to rule based text mining algorithms in order to extract relevant information. These very data will be enriched via ontology models with concepts close to the ones found in the CFP. Finally, using the previously collected data, we form the base of a user interface system with an intuitive query system allowing the user to perform complex queries.

PROBLEM DESCRIPTION CFP HANDLING PROBLEM

Our problematic can be divided in three parts; how to extract, manage and display relevant data. CFP is already an active subject of research to make the handling of CFP easier. Nevertheless these researches focus on displaying (Sacco, 2005), gather (Jaimez-Gonzalez & Luna-Ramirez, 2013) and creating end-user application (Xin, Li, & Tang, 2007) or web crawler (Yang, 2009) from available sources and do not take an Information Extraction approach as we intend to do. The Information extraction field is a well-known area and many techniques and algorithms have been used to solve this kind of problem like classifiers (WHISK), sequence model (Hidden Markov Chain) or written rules (NLP algorithm) (Feldman & Sanger, 2007) historically support by completion based conferences as MUC (Message Understanding Conference) (Chinchor & Sundheim, 1993). The application of IE to CFP can be traced back up to 1998 (Lazarinis, 1998). From this time researches focusing on extracting relevant information from CFP applying different algorithms as SVM (Support Vector Machine) (Ireson et al., 2005), CRF

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

(Conditional Random Field) (Scnheider, 2006; Cox, Nicolson, Finkel, Manning, & Langley, 2005) or rule induction (Ireson et al., 2005; Li, Rastan, Shepherd, & Paik, 2013) can be found. These researches give relatively good results in term of accuracy but always lake of diversity and precision in term of extracted data. For example even if deadline related information is extracted, precise information like which type of deadline for which type of paper is not provided.

CFP could be seen as a simple type of document, they are semi-structured data possessing internal data structures as shown in Figure 1; for example important dates as paper submission, deadline, etc… are often represented in a paragraph titled *Important Dates* with similar structures (type of event and date). Conference date and location are often found together with similar pattern of location enumeration. Even conference titles share some writing patterns. Therefore we decide to use rule based algorithms searching these patterns to extract relevant data.

In addition, we can also acknowledge that information contained in data can be extended to more general concepts but not expressed for several reasons. This can be seen in two different points in CFP files. The first one is the obvious data not mentioned by lack of usefulness; a good example could be location of conference. Location level upper than country are never mentioned in a CFP for evident reasons. However this information could be sometimes useful for queries. The second type is 'generalization': the way very specific information can be generalized to a wider term. In the CFP case, it can be found in conference topics when the topics are rather specific in order to be as exhaustive as possible. However if we could generalize these terms, we could discover and observe relations between conferences easily. Thus, we enhance the extracted data via external data stored in two ontologies, one for locations and another for topics as shown in one of our previous work (Issertial et al., 2012), as this kind of knowledge representation is close to our needs and have been used numerous times to represent conference linked subject (Sure, Bloehdorn, Haase, Hartmann, & Oberle, 2005; Jeong & Kim, 2010). This feature should improve the capacity to make meaningful queries for the user.

Figure 1. Example of a CFP


CALL FOR PAPERS

The Eighteenth IEEE International Conference on Engineering of Complex Computer Systems (ICECCS 2013)

17-19 July 2013
National University of Singapore, Singapore
http://www.comp.nus.edu.sg/~patiiceccs2013
==========
==========

IMPORTANT DATES

Abstract submission: February 1, 2013
Paper submission deadline: February 15, 2013
Workshop proposal submission: February 15, 2013
Notification of acceptance: March 30, 2013
Camera-ready material for publication: April 20, 2013
Checking for Production: April 30, 2013
Conference date: July 17-19, 2013

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Current Systems Review

There are already plenty of systems handling and presenting conference and CFP to users searching for conference suiting their interest. Consequently we decide to take into account some of these systems and to compare them to our system in terms of data usability and functionality, namely WikiCFP (2014), Semrex (Call For Paper) (Yuan et al., 2013), Conference Partner (2014), Conference Alerts (2014), ACM Conference Calendar (2014), IEEE Conference Search (2014), and Microsoft Academic Search (Jacso, 2011). The first three are general CFP aggregator for computer science related conferences. The following one is the same kind but with a broader field, from social science to economics. The following two gather only the conferences linked to their own organization respectively ACM and IEEE. The last one is just slightly related to CFP and focus more on publication. We finally add the basic research that can be executed in an email box as shown in Table 1.

As we can see, the strong point of our system is the variety of data handled by it, especially the different events' date. Except for the IEEE Conference Search, most of the systems are limited to gather CFP without providing the necessary tools to explore them. As for the topic row, it shows the range of topics available as a list and which also are linked to the CFP, i.e. the term don't have to appear in the CFP itself. However some mistakes or double can be found, for example with "AI" and "Artificial Intelligence" categories in WikiCFP, or many nonsensical categories in Semrex. Concerning the additional functionality, like aforesaid, beside IEEE (who also allows the use of wildcard) these systems do not give the opportunity to use Boolean operators and thus to perform complex queries. Finally our system shows three theoretical weak points with the impossibility to find if a conference exists in series, to use wildcard in queries, and to display if a deadline is extended or not.

Table 1. CFP Manager and available systems comparison

	CFP Manager	WikiCFP	Call For Paper SemreX	Conference Partner	Conference Alerts	ACM	IEEE	Microsoft Academic	Email
Topics	○ (1149)	○ (297) *1	△ (355)*2	×	△*3	△	△	○ (24)	×
City	○	○	×	×	○	×	○	×	△
Continent / Region	○	×	×	×	○	×	○	×	×
Deadline (Draft)	○	×	○	×	×	×	×	×	×
Deadline (Final)	○	×	×	×	×	×	○	×	×
Conference venue	○	×	×	×	○	○	○	×	×
Conference Series	×	○	×	△	×	×	×	○	×
Extended	×	×	×	○	×	×	×	×	△
Boolean operator	○	×	×	×	×	×	○	×	○
Wildcard	×	×	×	×	×	×	○	×	×

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

SYSTEM OVERVIEW

As shown in Figure 2, the CFP Manager system is composed of three main parts, information extraction from CFP, enrichment of these data via external ontologies and the query system.

Information Extraction

The information extraction part is based on the GATE framework developed by the University of Sheffield for text mining (Cunningham, Maynard, Bontcheva, & Tablan, 2002) as shown on Figure 3. Before using the text mining part the collected documents go through a pre-processing phase in order to extract the text body from an email document or a webpage. It also appeared that some CFPs are multi lingual so we extract the English part for the following phases, for this purpose we use basic pattern recognition and regular expression. From this point we use successively different GATE modules for text tagging that can be based on information tagged by previous modules; we begin by using six modules namely an English Tokenizer, a sentence Splitter, a Gazetteer for location, a Part of the Speech Tagger and finally a Named Entity Transducer.

Figure 2. CFP Manager Overview

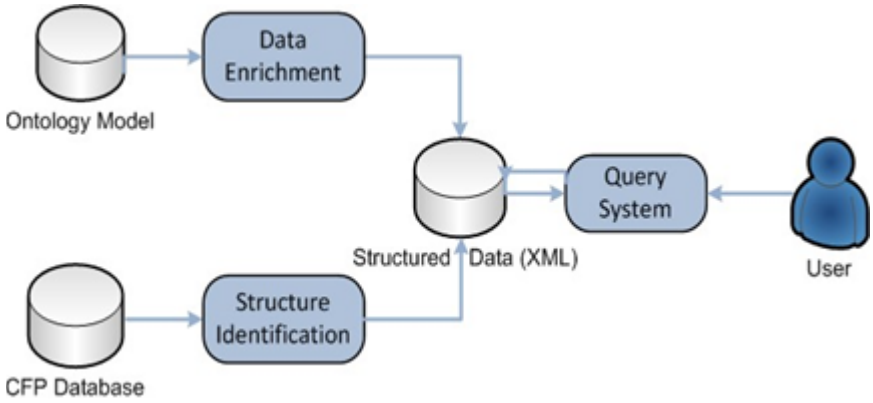
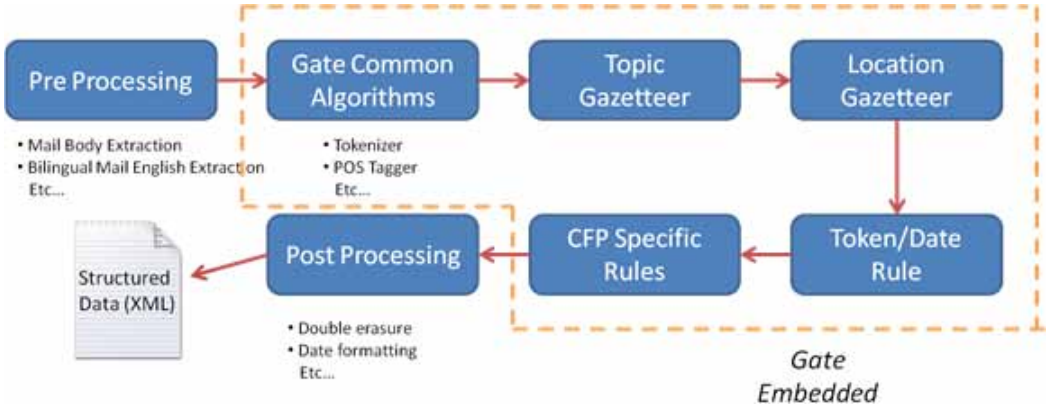


Figure 3. Information Extraction processing

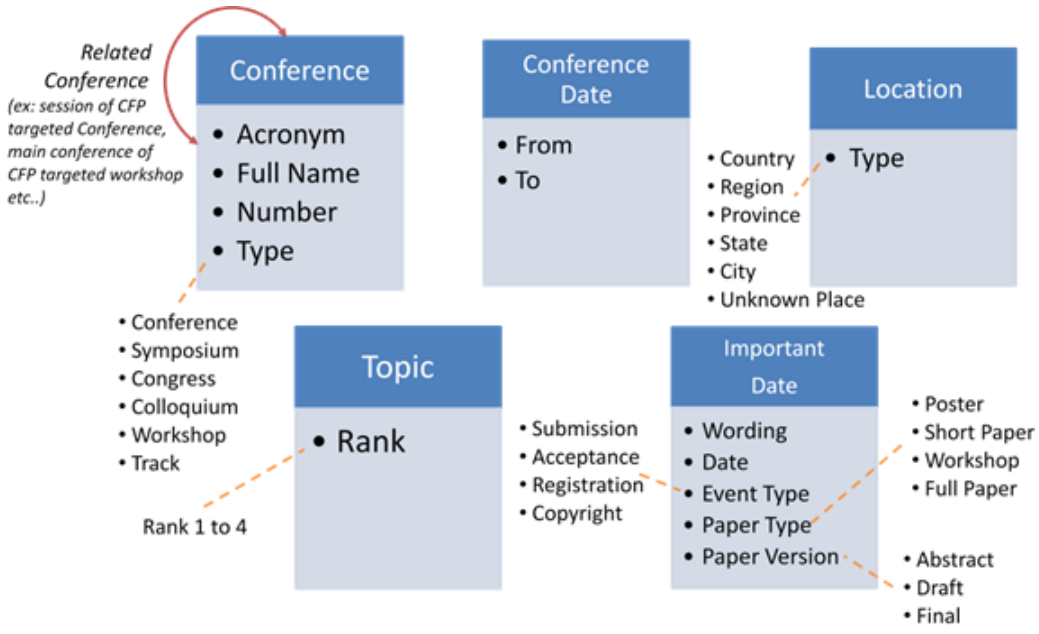


EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

This phase is followed by two gazetteers tagging respectively computer science related topics with a dictionary of 1085 terms divided in four different ranks based on ACM computing classification system (2012) and location forgot in the common algorithm phase (mainly USA states). Even if the basic GATE algorithms give a good base of useful annotations, we notice that some date and token needed for the following phases were not correctly annotated. Thus we add a phase with a set of 20 rules tagging date and token forgotten by the common algorithms. Then we finish the GATE phase with a set of rules specific to CFP in order to extract relevant data to our case. This phase is itself divided in 37 sub-phases processing 120 rules. All the data collected by the algorithm such as conference location, date, event (as shown on Figure 4) are then post processed. Data are ordered (put sub data with its affiliated one, like Date of a particular event), all extracted double are deleted, and if data already exists (in case of the same conference but with an extended deadline for example) we merge the two. Finally the result is stored in a structured data XML file.

As shown on Figure 4, we want to extract five types of data from CFP. Firstly the conference that can be divided in two labels *Acronym* and *Full Name*; from the *Full Name* we get the number and the type (symposium, workshop, etc…). Note that we are also interested in extracting related conferences; mostly conference holding the workshop advertised in the CFP, but it can also be a poster session just mentioned in the CFP conference or more rarely two conferences held together. From this we have the classic conference date and location, with location divided in five groups: country, region, province, state, city and unknown place. *Unknown Place* standing for place identified by rules but not present in our gazetteer like university, hotel or city to small to be in our gazetteer. *Important Date* represents every deadline link to the conference like submission or notification. We extract for this one the event and its date and make distinction between the different types of event and in the case of a paper, which type of paper and which version is needed. Finally topics are divided in four ranks as explained in the next part.

Figure 4. Extracted Data



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ontology Enrichment

Input data can be further improved via external data linked to the extracted one. We enrich our model via external ontology model. We use two of them for location and topic data, searching for linked information in the ontology and adding it in the structured data.

Location Ontology

Based on the FAO (Food and Agriculture Organization of the United Nation) geopolitical ontology (Iglesias-Sucasas, Kim, & Viollier, 2013), it provides information about geographical region and especially country upperlocations in the form of the five continents and twenty two continent subdivisions (Western Europe, Eastern Asia, etc.). It provides upper location information from country extracted in CFP.

Topic Ontology

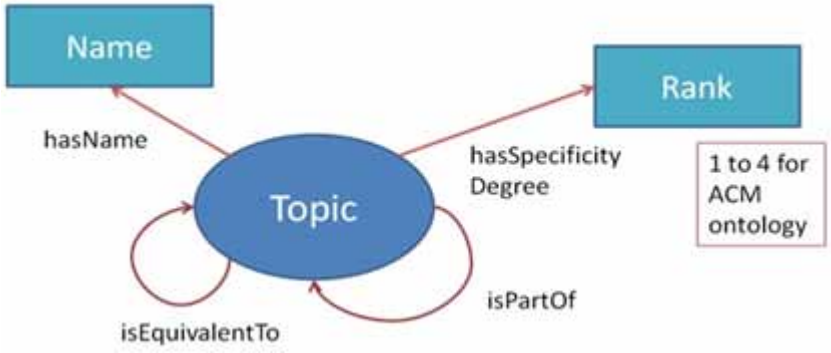
Based on the ACM computing classification system (2012), we use it to classify research topics with different layers. From this classification system we create an ontology model with four topic layers becoming more and more specific as shown on Figure 5. Thus we create relation between a given topic from the CFP and an equivalent or upper level one and so forth to the top level.

This improvement helps to link specific topics to more general and understandable ones. Even if, a given conference has only very specific topics tagged in its structured data file, this improvement allows us to access and classify it easily.

User Interface

The user interface allows the user to perform complex queries on his dataset and display it in various forms (CFP list, graphs, etc…), however we will focus only on the query part. As shown on Figure 6, the user can choose a date span for time sensitive data (like conference venue or deadline) or list plain text criteria (like topic or location) because more than looking for a conference name, the user tends to look for a conference regarding these constraints.

Figure 5. Topic Ontology Model



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Example of GUI Query



The user thus can apply to selected criteria the three basics Boolean operations AND, OR, NEG. The set of color represents the AND and OR operations: criteria with the same color are linked by a AND relation, thereby different colors stand for OR relation. The NEG is represented by a text in italic. Thus the query shown in Figure 6 should be read: (With First Draft From the 6th March 2014 to the 29th March 2014) AND ((Japan AND Text Mining AND NOT Artificial Intelligence) OR (Eastern Asia AND Data Mining)).

EXPERIMENTS

In order to validate our approach we set two experiments. The first one validates data extraction via precision and recall measures. The second measures the superiority of our approach in term of query results compared to other systems.

Text Mining Validation

The information extraction field has really strong links with the MUC (Message Understanding Conference) series (Chinchor & Sundheim, 1993), thus we use the well-known metrics introduced by it, Precision, Recall and F-Measure on two different criteria, strict and lenient, showing the utilization or not of partial result, defined as:

Strict:

$$Precision = \frac{Correct}{Correct + Spurious + Partial} \tag{1}$$

$$Recall = \frac{Correct}{Correct + Missing + Partial} \tag{2}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lenient:

$$Precision = \frac{Correct + Partial}{Correct + Spurious + Partial} \tag{3}$$

$$Recall = \frac{Correct + Partial}{Correct + Missing + Partial} \tag{4}$$

$$F - measure = \frac{(\beta^2 + 1) Precision * Recall}{\beta^2 Precision + Recall} \tag{5}$$

Where:

- Partial stands for annotation sharing a same span of text without being exactly the same
- Spurious indicates when a span of text is tagged in the gold standard and not in the tested one
- Missing indicates when a span of text is tagged in the tested model and not in the gold standard
- β is the weighting between Precision and Recall

We use these metrics and compare the tagging results of our system with a gold standard manually tagged on 100 CFP as shown on Table 2 showing the number of correct missing, spurious and partial data along with the recall (Equation 2,4), precision (Equation 1,3) and Fmeasure (Equation 5) on strict (s) and lenient (l) settings. Since this paper focus particularly on WikiCFP and IEEE on the next experiment, we repeat the same experiment on two 100 CFP datasets built exclusively with CFP from WikiCFP and IEEE respectively in Table 3 and Table 4. Note that we set the F-measure (Equation 5) weighting to 1 and the bottom line for summary calculates the precision, recall and F-measure for each document and then calculate the average.

We can divide the different extracted annotations in three different groups. The first one is crucial information used in query colored in white on Table 2 like *Conference Location*. For them we are expecting a score of 0.95 in precision and 0.90 in recall, the non apparition of the data being somehow preferable to the false one, the final user being able to access the original file anytime he wants. The second one in yellow represents information already inside another annotation like *Important Event Date* being a part of *Important Event*. For these annotations patterns are easily findable and the big part of the work has already been done. For these reasons we require a precision and recall rate near to the encompassing annotation, taking account of case of typo or unusual presentation. The last group in green represents data we don't really use at present but could be useful especially for linking conferences/CFP together and create an ontology from CFP. As we did not focus as much on this information for the moment, we conclude that a reasonable threshold should be set at 0.7 for both metrics.

As shown on Table 2, we successfully extract 1698 important data from our dataset with an F-measure of 0.92 and 0.94 respectively on strict and lenient settings. The reason why there is so many extracted data is due to the extraction of multiple instances of the same data in each document allowing us to discover more information (a conference acronym can have multiple spelling, a suit of location can give

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

more information than a previous one, etc…) As we can see most of the extracted data give satisfactory results within the threshold we previously fixed, especially in term of precision. In this case partial data counted as 18% of all the errors making a difference of 0.01 to 0.02 in summary result. However they are mostly shared between *Important*

Event and *Conference Full Name*, both of them exceeding our threshold in strict setting. *Related Conference Full Name* and *Conf Unknown Place* show also bad results. The former shows especially bad results in term of recall and it is intrinsically linked to the way it appears in rules. Its related rules are only triggered in combination with conference acronym leaving single full name in the text untagged and the same apply for *Conference Full Name*. Incidentally, the improvement of *Related Conference Acronym* retrieval should also improve its score. The latter represents every places tagged by our system only by rules and without a Gazetteer thus is make it more difficult to properly tag it and to know where it begin and end has shown by it missing and partial score. We can thus confirm the rules' tendency to be strict with low case of spurious and partial data and a majority of errors (48%) coming from missing data.

The WikiCFP dataset give in summary a little (0.01 to 0.02) worst result on recall and a slightly better result on precision (0.02 to 0.03). The repartition of error stay approximately the same for partial's one but increased for Missing going to 67%. Beside this, the results of the different extracted data remain close to the previous ones. We can guess that *Conf City*, *Conf Country* and *Important Event* results have slightly increased thanks to the WikiCFP template making these data more accessible.

Table 2. Precision, Recall, F-measure on extracted data

Annotation	Correct	Missing	Spurious	Partial	Rec. s	Prec. s	F1. s	Rec. l	Prec. l	F1. l
Conference Acronym	345	10	15	5	0.96	0.95	0.95	0.97	0.96	0.97
Conference FullName	72	31	12	14	0.62	0.73	0.67	0.74	0.88	0.8
ConfCountry	111	7	4	0	0.94	0.97	0.95	0.94	0.97	0.95
ConfCity	94	11	1	0	0.9	0.99	0.94	0.9	0.99	0.94
ConfProvince	2	1	3	0	0.67	0.4	0.5	0.67	0.4	0.5
ConfRegion	2	0	0	0	1	1	1	1	1	1
ConfState	8	1	0	0	0.89	1	0.94	0.89	1	0.94
ConfUnknown Place	24	5	0	4	0.73	0.86	0.79	0.85	1	0.92
ConferenceDate	127	9	4	1	0.93	0.96	0.94	0.93	0.97	0.95
ImportantEvent	278	32	9	12	0.86	0.93	0.9	0.9	0.97	0.93
ImportantEvent Date	290	5	10	0	0.98	0.97	0.97	0.98	0.97	0.97
ImportantEvent Wording	284	5	16	6	0.96	0.93	0.95	0.98	0.95	0.97
RelatedConference Acronym	40	5	8	1	0.87	0.82	0.84	0.89	0.84	0.86
RelatedConference FullName	21	4	7	4	0.72	0.66	0.69	0.86	0.78	0.82
Summary	**1698**	**126**	**89**	**47**	**0.91**	**0.93**	**0.92**	**0.93**	**0.95**	**0.94**

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Precision, Recall, F-measure on extracted data on the WikiCFP dataset

Annotation	Correct	Missing	Spurious	Partial	Rec. s	Prec. s	F1. s	Rec. l	Prec. l	F1. l
Conference Acronym	165	17	5	1	0.9	0.96	0.93	0.91	0.97	0.94
Conference FullName	73	37	9	6	0.63	0.83	0.72	0.68	0.9	0.77
ConfCountry	71	2	1	1	0.96	0.97	0.97	0.97	0.99	0.98
ConfCity	75	4	0	0	0.95	1	0.97	0.95	1	0.97
ConfProvince	15	2	0	0	0.88	1	0.94	0.88	1	0.94
ConfState	10	0	0	0	1	1	1	1	1	1
ConfUnknown Place	26	8	2	2	0.72	0.87	0.79	0.78	0.93	0.85
ConferenceDate	136	1	1	0	0.99	0.99	0.99	0.99	0.99	0.99
ImportantEvent	304	23	4	11	0.9	0.95	0.93	0.93	0.99	0.96
ImportantEvent Date	308	16	4	7	0.93	0.97	0.95	0.95	0.99	0.97
ImportantEvent Wording	306	18	4	9	0.92	0.96	0.94	0.95	0.99	0.97
RelatedConference Acronym	17	6	1	0	0.74	0.94	0.83	0.74	0.94	0.83
RelatedConference FullName	6	7	1	0	0.46	0.86	0.6	0.46	0.86	0.6
Summary	**1512**	**141**	**32**	**37**	**0.89**	**0.96**	**0.92**	**0.92**	**0.98**	**0.95**

Table 4. Precision, Recall, F-measure on extracted data on the IEEE dataset

Annotation	Correct	Missing	Spurious	Partial	Rec. s	Prec. s	F1. s	Rec. l	Prec. l	F1. l
Conference Acronym	349	10	15	2	0.97	0.95	0.96	0.97	0.96	0.97
Conference FullName	85	21	12	11	0.73	0.79	0.76	0.82	0.89	0.85
ConfCountry	99	6	2	0	0.94	0.98	0.96	0.94	0.98	0.96
ConfCity	78	6	3	0	0.93	0.96	0.95	0.93	0.96	0.95
ConfProvince	11	2	3	0	0.85	0.79	0.81	0.85	0.79	0.81
ConfRegion	3	0	0	0	1	1	1	1	1	1
ConfState	3	0	0	0	1	1	1	1	1	1
ConfUnknown Place	27	13	0	2	0.64	0.93	0.76	0.69	1	0.82
ConferenceDate	114	5	1	3	0.93	0.97	0.95	0.96	0.99	0.97
ImportantEvent	301	18	8	19	0.89	0.92	0.9	0.95	0.98	0.96
ImportantEvent Date	322	13	10	0	0.96	0.97	0.97	0.96	0.97	0.97
ImportantEvent Wording	317	17	12	12	0.92	0.93	0.92	0.95	0.96	0.96
RelatedConference Acronym	30	8	6	2	0.75	0.79	0.77	0.8	0.84	0.82
RelatedConference FullName	14	8	13	0	0.64	0.52	0.57	0.64	0.52	0.57
Summary	**1753**	**127**	**85**	**51**	**0.91**	**0.93**	**0.92**	**0.93**	**0.96**	**0.94**

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The IEEE dataset gives a result roughly identical to the first one. The type of error proportion is the same and the metrics for the summary give almost the same result with plus 0.01 on lenient precision. Finally, the extracted data result pattern and weak points are the same. As we shown on Table 3 and Table 4 the differences between the three datasets are really subtle and the same pattern can be found in the three experiments, we can infer that our system can handle the heterogeneity of CFP data sources, furthermore we can compare our system to the WikiCFP and IEE one introduced in the next part.

Comparison with Other Systems

After we confirmed our system extracts a reasonable amount of data, we will focus on the relevance of its query system. As a final user who wants to find a conference according to his preferences, do our query system and our system as a whole provide a suitable experience? To answer that we compare our CFP Manager with other systems we described previously. We choose the 3 systems: basic email searching (text searching), WikiCFP and IEEE Conference Search. The first two are representative to the way information on a broader type of CFP is currently provided. The later, by its various functionalities, represent the goal our system tends forward in terms of query.

We compare the response to each system to a series of queries done on a 100 CFP dataset. However, since using these systems per se is quite difficult we decide to simulate their working. For example, as shown on Table 1, WikiCFP cannot handle queries on deadline so this part of the query will be simply ignored, however, it use its own set of topics to classify CFP so we use it instead of extracted topics. Because of these variations, we will use different dataset for this case, with only CFP from WikiCFP. For the two others a combination of various CFP can be used. The tables will also present results with and without the help of the ontologies.

WikiCFP

As shown on Table 5, WikiCFP shows some weak points compared to our system. First of all on topics, it does not use CFP body text and thus misses some conferences (cf. *data mining* and *artificial intelligence* queries). It also misses seldom topics like *network protocols* or *wireless communication*. The impossibility to use Boolean operator leads to false positive as in the *China and artificial intelligence* request and, as said before, the impossibility to use deadline also leads to false positive. However our system using only topics from the topic ontology creates some problems when we want to use simple terms as *wireless*.

Basic Searching (Pseudo Email)

The basic text searching that stand for email shows the weakness we could predict from this kind of searching: the impossibility to understand the context. Thus it creates double in *data mining* for example as it don't understand the difference between a CFP and one of its updated version (for example extended deadline) and incidentally also creates false positives for deadline. Same for location found in CFP emails that aren't the conference location and thus creates false positive for *artificial intelligence and China* query. Finally without structure we cannot use date type so use a simple text query instead (*deadline march 2012)* however it miss the CFP because of the different ways the information is writing.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. CFP Manager / WikiCFP Comparison

Query \ Result by System Positive (False Positive)	CFP Manager			WikiCFP
	TM	Onto	Total	
Data Mining	16	0	16	9
China	7	0	7	7
Artificial Intelligence	11	28	39	6
Artificial Intelligence in China	2	2	4	13(5)
Artificial Intelligence With Final Version in May 2012	1	2	3	4(3)
Wireless communication	2	0	2	0
Wireless	X	X	X	7
Network protocols	1	0	1	0

Table 6. CFP Manager / Email Comparison

Query \ Result by System Positive (False Positive)	CFP Manager			Basic Searching
	TM	Onto	Total	
Data Mining	18	0	18	21
China	9	0	9	23(11)
Artificial Intelligence	9	30	39	10
Artificial Intelligence in China	3	2	5	3(1)
Data in march 2012 (deadline march 2012)	1	0	1	0(1)
Wireless communication	1	0	1	1
Wireless	X	X	X	9

IEEE Conference Search

As seen previously, the IEEE system already possesses many features to find a conference and thus give reasonable results. On the downside, it can only use date for conferences or final version but can't find notification, or draft types (cf. *Artificial Intelligence With Draft in May 2012*) contrary to our system.

CONCLUSION

This paper has introduced a Call For Papers manager system aimed at helping researcher to handle the daily flow of CFP and ease their research of a suitable conference. We have introduced the different systems currently available online to access CFP and pointed their main flaw as the lack of usable criteria to perform a request. Therefore, we have presented the main characteristics of our system: the extraction

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 7. CFP Manager / IEEE Conference Search Comparison

Query \ Result by System Positive (False Positive)	CFP Manager			IEEE
	TM	Onto	Total	
Data Mining	27	0	27	27
China	9	0	9	9
Artificial Intelligence	20	41	61	20
Artificial Intelligence in China	2	4	6	2
Artificial Intelligence With Final Version in May 2012	1	0	1	1
Artificial Intelligence With Draft in May 2012	2	0	2	X
Wireless communication	2	0	2	2
Network protocols	2	0	2	2

of relevant data from CFP dataset using text mining algorithms, mainly rule based ones; the enrichment of these data via ontology models for conference location and topics; and finally the utilization of these data by the final user to find a suitable conference via interface allowing complex queries.

In order to validate our approach, we have set two experiences. The first one has measured via precision and recall, the efficiency of information extraction offering satisfactory results with 0.95 for precision and 0.91 for recall on average on our three dataset with strict and lenient setting. The second one has compared the number of results obtain on different queries with four systems: WikiCFP, IEEE Conference Search, basic text searching and our system with and without the ontology enrichment. We found that our system has surpassed WikiCFP and basic text searching by eliminating the false positive they returning.

As future work, the system still need improvement like the amelioration of information extraction especially for *Related Conference Full Name* and *Related Conference Acronym* or also the extraction of new type of data like conference website for example. The contribution of the ontology model should also be reviewed and the use of bigger ontology could be considered. From this point we could consider different application, for instance we could continue our research concerning data providing and implement a calendar to the system. We could also look at the other side of the problem and implement the extraction part as a tool to upload contentaware CFP on existing platform. Even if the result gives a good glimpse of the system capacity to handle heterogeneous sources of CFP, additional tests on bigger dataset should be performed.

ACKNOWLEDGMENT

This research was supported by a Grant-in-Aid for Scientific Research (A), No. 13370017, Japan.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

ACM Classification System. Retrieved July 20, 2014 from http://www.acm.org/about/class/ccs98-intro

ACM Conference Calendar. Retrieved February 27, 2014 from http://www.acm.org/calendar-of-events

Chinchor, N., & Sundheim, B. (1993). MUC-5 evaluation metrics, *Proceedings of the 5th Conference on Message Understanding*, 69-78. 10.3115/1072017.1072026

Conal (Conference Alerts). Retrieved February 27, 2014 from http://www.conferencealerts.com/index

Conference Partner. Retrieved February 27, 2014 from http://myhuiban.com/conferences

Cox, C., Nicolson, J., Finkel, J. R., Manning, C., & Langley, P. (2005). Template Sampling for Leveraging Domain Knowledge in Information Extraction, *Proceedings of the PASCAL Challenges Workshop*, 1-4.

Cunningham, H., Maynard, D., Bontcheva, K., & Tablan, V. (2002). GATE: an Architecture for Development of Robust HLT Applications, *Recent Advanced in Language Processing,* 168-175.

Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. Cambridge University Press.

IEEE Conference Search. Retrieved February 27, 2014 from http://www.ieee.org/conferences_events/conferences/search/index.html

Iglesias-Sucasas, M., Kim, S., & Viollier, V. (2013). The FAO Geopolitical Ontology: A reference for country-based information. *Journal of Agricultural & Food Information*, *14*(1), 50–65. doi:10.1080/10496505.2013.747193

Ireson, N., Ciravegna, F., Califf, M. E., Freitag, D., Kushmerick, N., & Lavelli, A. (2005). Evaluating Machine Learning for Information Extraction, *Proceedings of the 22nd international conference on Machine learning*, 345-352.

Issertial, L., Saga, R., & Tsuji, H. (2012). Visualized comparison for CFP datasets by structure identification and ontology. *Proceedings of the 6th International Conference on New Trends in Information Science, Service Science and Data Mining*, 365-370.

Issertial, L., & Tsuji, H. (2011). Information extraction and ontology model for a 'call for paper' manager. *Proceedings of the 13th International Conference on information integration and web-based applications & services*, 539-542

Issertial, L., & Tsuji, H. (2013). Data management and user interface for a Call For Paper Manager. *Proceeding of the 37th Annual International Computer Software & Applications Conference*, 463-466. 10.1109/COMPSAC.2013.78

Jacso, P. (2011). The pros and cons of Microsoft Academic Search from a bibliometric perspective. *Online Information Review*, *35*(6), 983–997. doi:10.1108/14684521111210788

Jaimez-Gonzalez, C. R., & Luna-Ramirez, W. A. (2013). Towards an Automatic Multi-Agent System for Joining Information from Specialized Web Sites, *Proceedings of the IEEE Eleventh International Symposium on Autonomous Decentralized Systems*, 1-6. 10.1109/ISADS.2013.6513404

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Jeong, S., & Kim, H-G. (2010). SEDE: The Scholarly Event Description Ontology, *Journal of Information Science, 36, 2*, Sage Publications, 209-227.

Lazarinis, F. (1998). Combining Information Retrieval with Information Extraction for Efficient Retrieval of Calls for Papers, *Proceedings of the 20th Annual BCS-IRSG conference on Information Retrieval Research*, *British Computer Society,* 620.

Li, X., Rastan, R., Shepherd, J., & Paik, H. Y. (2013). Automatic Affiliation Extraction from Calls-For-Papers, *Proceedings of the 2013 workshop on Automated knowledge base construction*, 97-102. 10.1145/2509558.2509575

Sacco, G. M. (2005). DBWorld Xtended: Semantic Dissemination of Information through Dynamic Taxonomies. *Proceedings of I-KNOW*, *05*, 128–135.

Scnheider, K-M. (2006). Information extraction from calls for papers with conditional random fields and layout features, *Artificial Intelligence Review,* 25, 1-2, Springer, 67-77.

Sure, Y., Bloehdorn, S., Haase, P., Hartmann, J., & Oberle, D. (2005). The SWRC Ontology - Semantic Web for Research Communities, *Proceedings of the 12th Portuguese conference on Progress in Artificial Intelligence*, 218-231. 10.1007/11595014_22

WikiCFP. Retrieved February 27, 2014 from http://www.wikicfp.com

Xin, X., Li, J., & Tang, J. (2007). Enhancing Semantic Web by Semantic Annotation: Experiences in Building an Automatic Conference Calendar, *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*, 439-442.

Yang, S.-Y., & Hsu, C.-L. (2009). Ontology supported web crawler for information integration on Call For Papers, *Proceedings of the 8th International Conference on Machine Learning and Cybernetics*, 3354-3360. 10.1109/ICMLC.2009.5212759

Yuan, P., Jin, H., Li, Y., Chang, B., Ning, X., & Huang, L. (2013). SemreX: A Semantic Association-Based Scientific Literature Sharing System, Springer Proceedings in Complexity, Semantic Web and Web Science, Springer, 161-167.

This research was previously published in the International Journal of Knowledge and Systems Science (IJKSS), 6(4); edited by W.B. Lee; pages 35-49; copyright year 2015 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 21
Word Sense Based Hindi-Tamil Statistical Machine Translation

Vimal Kumar K.
Jaypee Institute of Information Technology, India

Divakar Yadav
Jaypee Institute of Information Technology, India

ABSTRACT

Corpus based natural language processing has emerged with great success in recent years. It is not only used for languages like English, French, Spanish, and Hindi but also is widely used for languages like Tamil, Telugu etc. This paper focuses to increase the accuracy of machine translation from Hindi to Tamil by considering the word's sense as well as its part-of-speech. This system works on word by word translation from Hindi to Tamil language which makes use of additional information such as the preceding words, the current word's part of speech and the word's sense itself. For such a translation system, the frequency of words occurring in the corpus, the tagging of the input words and the probability of the preceding word of the tagged words are required. Wordnet is used to identify various synonym for the words specified in the source language. Among these words, the one which is more relevant to the word specified in source language is considered for the translation to target language. The introduction of the additional information such as part-of-speech tag, preceding word information and semantic analysis has greatly improved the accuracy of the system.

INTRODUCTION

Natural language processing (NLP) is a part of artificial intelligence which interacts with the systems (computer) through natural languages to perform desired actions. It deals with understanding and analyzing human languages in order to perform various functionalities which can enhance the interaction between the machine and the humans. There are various widely used algorithms under NLP, especially statistical natural language processing but each algorithm has its own bottleneck. These algorithms are usually based upon the analysis of large textual corpora and then calculating probabilities in order to

DOI: 10.4018/978-1-7998-0951-7.ch021

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

achieve the desired results. According to linguistics, corpus refers to large structured texts consisting of numerous words which are used for statistical analysis of the text. Generally, the corpus should be annotated to provide an efficient statistical analysis. The Corpus consists of each and every word in every sentence used for the language analysis. These words are added to the corpus along with the information about its part of speech such as: verbs, adjectives, nouns, and adverbs etc., which are called as POS tags. Corpus based NLP techniques have emerged with great success in the recent years. It is not only used for languages like English, French, Spanish, and Hindi but also is widely used for languages like Tamil, Vietnamese etc.

Numerous algorithms have been introduced in statistical machine translation to provide various intelligent functionalities for human-computer interaction. All these algorithms parse the sentences and then group the words before the translation process. Parsing of free word order languages such as Indian is also a bottleneck in these methods (Bharati et al., 2009; Bharati & Sangal, 1993). Local word grouping (LWG) is basically used in Indian languages since there is a need for grouping the words based on the context in which it is used and the meaning of those words will be clear only when it is grouped together (Bharati et. al, 1991; Ray et. al, 2003; Balaji et al., 2014). In these existing technologies related to corpus based NLP, the statistical analysis for machine translation makes use of a parallel corpus. In a parallel corpus, each word in the source language is mapped parallel with its corresponding word in the target language. In addition to parallel corpus being used for translation, the part-of-speech of the words is also considered for machine translation. Also, in statistical machine translation, the target texts are generated on the basis of statistical models and these models are derived from the analysis of the text corpus of the two languages. Generally, a document is translated to a probable sentence in the target language according to the probability distribution $P(t|h)$ which refers to the probability of string t in the target language (for example, Tamil) given the string h in the source language (for example, Hindi).

Naive Bayes algorithm is one of the existing algorithms which are based on statistical analysis of the existing bilingual corpus. The algorithm uses probability of occurrence of words for translation from one language to another. For a particular word, its probability is calculated based upon the frequency of occurrence of the word in the corpus. The meaning which has maximum occurrence in the target language will be the probable translation for the input word. The mathematical representation of this algorithm is:

$$\mathrm{P}\left(\frac{x}{y}\right) = \mathrm{P}\left(\frac{y}{x}\right) * \mathrm{P}(x) / \mathrm{P}(y)$$

Since P(y) will not affect the result, the equation is equated as shown below:

$$\mathrm{P}\left(\frac{x}{y}\right) \cong \mathrm{P}\left(\frac{y}{x}\right) * \mathrm{P}(x)$$

This mathematical representation signifies that there is translation model and language model being used to perform the translation process. Translation model analyzes the translation between source and target language whereas the language model analyzes the target language being used. These statistical algorithms can further be improved by considering the tag information of the word under consideration but still there will be lack of efficiency and accuracy. Also, ambiguity is one of the concerns of these

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

algorithms. In order to overcome this ambiguity issue, an analysis on these languages has been made and found that the sense of word can be used to eradicate this ambiguity issue. Certain word has different translations based on its part-of-speech which has been included in this Naïve Bayes model. There are words that have multiple translations based on the context in which it is used and this can be identified using word sense disambiguation over the sentence. Thus, the objective of this system is to improve the efficiency of the statistical machine translation by providing additional information such as the part-of-speech of the word to be translated as well as the information about the preceding words and also to make use of the word's sense in the translation process. Since this translation process takes in to account the word's part-of-speech (syntactic information) as well as the word's sense (semantic information) during the machine translation, it improves the efficiency and accuracy of the translation.

Rest of the contents of this paper is arranged as follows: Section 2 discusses the existing works, researches and technologies which have been prevalent in the field of machine translation. It also discusses about challenges faced and solutions proposed in the respective field. Section 3 consists of a detailed description of the algorithm introduced, theory, implementation and mathematical interpretation. The results and its analysis are described in section 4. Finally, in section 5, it concludes the work along with some future directions.

RELATED WORK

Various techniques have been proposed in the sub field of inter language translation under the umbrella of natural language processing (NLP). Cross linkage of words from two different languages have empowered NLP facilitating bilingual dictionaries and other practical applications. This can be a challenging task as the linkage needs to be balanced and two languages can be varied on grounds of culture, meanings etc. Hindi to English linkage suffers from kinship relations (uncle in English can be mama or chacha in Hindi), musical instruments, grains, kitchen utensils etc. Direct and hypernamy linkage can cater to such challenges. Also, manual linking of two languages can be facilitated using WordNet SynsetMatcher tool (Saraswati et al., 2010; Alguliyev, et al., 2017).

Annotation in Hindi and English varies highly as both the languages differ in terms of grammar. The two techniques for tagging can be: linking all possible words in both source and target language, or, linking least possible words in both source and the target language. Since exact correspondence is hard to find and word to word translation is possible for very few words, thus remaining words are translated based on the combination of words. Thus, annotation can follow certain guidelines such as following fuzzy, regular and null links in punctuation which can lead to desired results (Yadav & Gupta, 2010; Hema & Geetha, 2016).

A system, ANUBHARTI which follows an approach of machine aided translation having the combination of example based and corpus based approaches with some elementary grammatical analysis is been developed for corpus generation (Bharati et al., 1997). In ANUBHARTI the traditional EBMT approach has been modified to reduce the requirement of a large example based corpus. ANUBHARTI-II in 2004 uses Hindi as a source language for translation to other Indian language.

A machine translation system which translates pure (standard) Hindi to pure English was implemented in 2004 by incorporating additional level to AnglaBharti-II and AnuBharti-II. The system gives satisfactory results in more than 90% cases (Sinha & Thakur, 2005; Sinha, 2004).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Vishal Goyal and G.S. Lehal (2010) proposed a machine translator from Hindi to Punjabi. The methodology used for the translation was direct translation and later on improving the language learning modules for the enhancement of the quality of the system. The accuracy of the translation is approximately 95%. The web tool developed has many application areas like over internet for the particular community of peoples, for the newspapers and sending e-mails from Hindi to Punjabi or vice versa (Goyal, 2010; Goyal & Lehal, 2010).

The performance of the multilingual information retrieval employing dictionary based strategy for query translations have been measured and analyzed by several articles reported in the literature (Korra et al., 2011; Abbache, 2016). The experimental result shows that the performance of the multilingual information retrieval system enhances over the monolingual information retrieval system by 31.4%.

For English to Hindi statistical machine translation, the key challenge is that the Hindi language is richer in morphology than the English language. Reordering of English source sentences in accordance with Hindi language and/ or making use of suffixes of Hindi words are the two strategies that can facilitate reasonable performance. Since Indian languages and English differ in word order, morphologically rich Indian languages and unavailability of huge parallel corpora for two languages make the above two strategies more challenging to derive the desired results with reasonable performance (Ananthakrishnan et al., 2009).

PROPOSED METHOD

This proposed method performs translation at the word level and it follows a step by step procedure. The paper focuses on a system which makes use of a translation model and a language model. The translation model is developed to provide the translation from one language to another, based on the analysis of both the languages. Since the translation process depends on the grammatical sequence of words in the source language, an analysis of grammatical structure of the source language is performed and a language model is developed for the same. It is identified for the languages under consideration that the word's sense also matter during the translation process, the word's sense is identified using the semantic analysis. The system architecture is as shown in Figure 1.

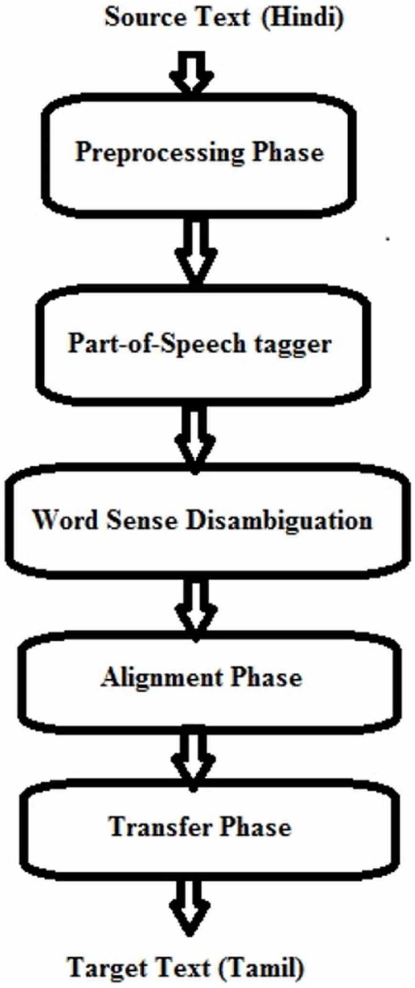
Following are the detailed description about the various phases of this proposed translation system.

Preliminary Phase: Corpus Generation

The text corpus is a large set of structured text electronically stored in the database. The corpus is used for the statistical analysis within the language territory. In this algorithm, a multi-lingual database is built for the language model and translation model by traversing over a big textual corpus in both the languages provided by Technology Development for Indian Languages (TDIL) Programme, Department of Information Technology (DIT), MC&IT, Govt of India. The database is made in accordance to the part-of-speech tagging for individual words in the source language (Hindi) along with semantically equivalent of the target word (Tamil). Additionally, words' frequency of occurrence in the text corpus is also stored. This is one part of corpus generation which is required for the translation model. Another database is built for Tamil text by traversing a Tamil corpus and storing values in the database for preceding words, current word, and their frequencies. This database signifies the language model required for the statistical system

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. System architecture



Pre-Processing Phase

Since this system is a word level translation system, the preprocessing phase is to tokenize the input text in to individual sentences. These sentences are further tokenized to extract the words and these words are passed on to the next phase of the translation system which is the part-of-speech tagging phase.

POS Tagging Phase

In machine translation, the part-of-speech tagging has to be performed that refers to the marking of the words for information based upon the context of the word and also the relationship between the word and the adjacent words. It refers to the identification of words as nouns, verbs, adjectives, adverbs etc. Part-of-speech tagging is generally of three types namely: rule based part-of-speech tagging, transformation based part-of-speech tagging and probabilistic based part-of-speech tagging. Ambiguity is one of the major problems that come across while performing part-of-speech tagging. This system uses the well-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

known HMM based POS Tagging. The tagger considers the preceding word information to improve the tagging process that reduces the ambiguity. As the proposed system considers the tagging information during translation, the accuracy of the system depends more on this phase of the translation process. In case the tag assigned to a particular word is wrong, then the system's translation won't be proper.

Word Sense Disambiguation using Hindi Wordnet

This phase uses the Hindi Wordnet dictionary to identify the various senses for the particular word with its identified part-of-speech. Semantic analysis between the words under consideration over the various senses of the word provides a score of its relevance. The high scored senses are used further in the translation. For the sense disambiguation, the latent semantic analysis (LSA) technique is used. LSA technique is a singular value decomposition method which decomposes the term-frequency matrix in to three different matrices namely left singular matrix (L), right singular matrix (R) and singular diagonal matrix (D). Left singular matrix is a matrix which has the Eigen vectors of the term matrix and its transpose and this matrix is also called as term matrix. Right singular matrix is one which contains the singular values of the term frequency matrix as the diagonal elements and it is also called as concept matrix. Whereas the singular diagonal matrix has the Eigen vectors of the transpose of term-frequency matrix and this matrix is known as concept by document matrix.

$$\begin{bmatrix} f_{t1}^{1} & f_{t2}^{1} & \ldots\ldots & f_{tn}^{1} \\ f_{t1}^{2} & f_{t2}^{2} & \ldots\ldots & f_{tn}^{2} \\ \ldots & \ldots & \ldots & \ldots \\ f_{t1}^{m} & f_{t2}^{m} & \ldots\ldots & f_{tn}^{m} \end{bmatrix} = L * R * S$$

where, f_{tn}^{m} - indicates the frequency of n^{th}-term in m^{th}-sentence

The term frequency of words is arranged in the form of a matrix with column as different terms (words) used in the text and row of the matrix as the sentence identification numbers. After the decomposition of this matrix, the product of the right singular matrix and singular diagonal matrix gives the frequency of semantically similar words in different sentences and the similarity between two sentences is found using cosine similarity over the product of right singular vector and singular diagonal vector of the sentence considered (Kumar et al., 2015). The cosine similarity is calculated using the below mentioned in the equation:

$$\cos\theta = \frac{\sum_{i=1}^{n} x_i * y_i}{\left(\sqrt{\sum_{i=1}^{n} x_i^2} * \sqrt{\sum_{i=1}^{n} y_i^2}\right)}$$

At this phase, the system has identified the relevant senses for the word, now the system has to rearrange the sentence since the naïve Bayes method requires the proper sequence in the target language.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Alignment Phase (Syntactic Restructuring)

Based on the analysis, it is found that Hindi language follows Subject Object Verb (SOV) sequence that is same as in Tamil. The transfer phase makes use of the target language model and considers the preceding word information in the target language level. Since, the text which is in source language follows the grammatical correctness of the words in target language; there is no need for rearranging the words based on the target language grammar under certain exceptional cases such as compound sentences, which gives emphasis on the restructuring of sentences in the target language Tamil. In Tamil language, there are various affixes being used in a sentence which doesn't exist in Hindi language. Accordingly, the grammatical restructuring of the words in source language happens such that the source language sentence is aligned with the grammar as in the target language. This is achieved by applying the statistical based alignment algorithm. Basically, this system makes use of GIZA++ (a tool for word alignment). This tool generates the alignment table which is required for the statistical method. The sample alignment table is as shown in Table 1 which has the position in target language, source word position, length of Hindi sentence and length of its corresponding Tamil word.

Naïve Bayes model is applied over this alignment table to identify the probable alignment in target language for the given sentence in source language. The one which has the higher probability will be used for aligning these source language words.

Transfer Phase

The basic model of language translation, also called as parallel translation, finds every individual target language word of the source language word in a list and substitutes this word with the equivalent target language word. But this method does not ensure accurate results as the relationship followed between words in source language and in target language are often one-to-many, instead of one-to-one. This is due to that a word in source language can give different equivalent translations in the target language based on the context where it is being used. Solutions for increasing the accuracy of the results are to distinguish the word on the basis of its part-of-speech and also on the basis of word's sense. In this proposed system, the part-of-speech of the word is used to identify the category of words and in the particular category a word can have multiple senses which are further identified using LSA and Wordnet tool. The sense of the word is used in our statistical translation model.

Probability serves a major role to analyze any complex data. This system follows analysis for translation by taking it into account the probability of target word given the source word and its corresponding part-of-speech. The probability can be represented as mentioned below:

Table 1. Sample alignment table

S.No	J (Word's Position in target language)	I (Word's Position in Source language)	L (Source Sentence Length)	M (Target Sentence Length)	Frequency
1	4	3	8	10	252
2	7	7	10	10	231

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$P\left(W_t \mid W_s, W_{tag}\right) = \arg\max_{j \to 0 to N} P\left(W_j, W_{tag} \mid W_t\right) * P\left(W_t\right)$$

Where, N indicates the total number of senses identified for word W_s

$$P\left(W_s, W_{tag} \mid W_t\right) = P\left(W_s \mid W_t\right) * P\left(W_{tag} \mid W_t\right)$$

$$P\left(W_s \mid W_t\right) = \frac{count\left(W_s, W_t\right)}{count\left(W_t\right)}$$

$$P\left(W_{tag} \mid W_t\right) = \frac{count\left(W_{tag}, W_t\right)}{count\left(W_t\right)}$$

$$P\left(W_t\right) = P\left(W_t \mid W_{prev_word}\right)$$

$$P\left(W_t\right) = \frac{count\left(W_t, W_{prev_word}\right)}{count\left(W_t\right)}$$

The above-mentioned formulae are used for finding the probable translation for one sense of the word in the source language. To find the most probable translation for that context in which it is being used, the above equation is applied over all the various senses of the word and the one which has the maximum probability value is selected as the best translation for the word. Sample tagged transfer database is as shown in Table 2. This database has Hindi words, its corresponding Tamil words based on the POS tag of the Hindi word along with the frequency of occurrence. In Table 2, the word अच्छी (*AACHCHI*) appears in 89 sentences out of which it is converted to நல்ல (*NALLA*) in 49 sentences and the rest of the sentences it is converted as நன்றாக (*NANDRAGA*). From this it is evident that the words' part of speech can be used to distinguish the target word translation. The same word appearing as adjective, has two different equivalent words in Tamil in *60* cases and in such cases, it can be narrowed to one particular word by considering its preceding word information as well as the word's sense.

Table 2. Sample transfer database

Hindi word	Tamil word	Tag	Frequency
अच्छा	நல்ல	JJ	53
अच्छी	நன்றாக	RB	29
अच्छी	நன்றாக	JJ	11
अच्छी	நல்ல	JJ	49
उत्पादक	உற்பத்தியாளர்	N_NN	24

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Sample precedence database

Precedence word	word	frequency
^	நல்ல	18
மிகவும்	நல்ல	25
^	நன்றாக	13
இது	நன்றாக	15
^	உற்பத்தியாளர்	4
நல்ல	உற்பத்தியாளர்	12

There comes the need for a precedence database whose sample is as shown in Table 3 and this database has the details about the preceding word for any given Tamil word along with its frequency. The use of ^ symbol for precedence word denotes that the word occurs in the beginning of any sentence.

For Example

Consider the input text, **"**ताजा साँसें और चमचमाते दाँत आपके व्यक्तित्व को निखारते हैं।**"**

After preprocessing the text, individual words are identified as unique tokens from the input text. This tokenized text is subjected to POS Tagging. The tagged sentence output is ताजा\JJ साँसें\N_NN और\CC_CCD चमचमाते\JJ दाँत\N_NN आपके\PR_PRP व्यक्तित्व\N_NN को\PSP निखारते\V_VM हैं\V_VAUX ।\RD_PUNC. This tagged text is considered for syntactic restructuring but in this case there won't be any change in the order of the sentence so the same output is fed to statistical transfer phase and is described as follows,

Considering the word *"*ताजा\JJ*"*,

From the wordnet the various senses that are found for (ताजा as Adjective) are - ताज़ा, ताजा, अम्लान, अशुष्क, आला, गरमागरम, टटका. These words are further subjected to word sense disambiguation using LSA. For the latent semantic analysis, the sentences formed using these various senses and the given word is converted in to term-frequency matrix which is used in the latent semantic analysis. The LSA algorithm identifies that ताजा and ताज़ा are the relevant words for the given input word.

Case-1: When source word W_s is "ताजा" and Tag W_{tag} is "JJ" and target word W_t is "புத்தாணர்ச்சியான" (Puthunarchiyana means Fresh), then $P(W_s, W_{tag}|W_t) = 30/34$

Also, when the target word W_t is "புத்தாணர்ச்சியான" and precedence is nothing, then $P(W_t) = 22/30$

So combined probability $P(W_t|W_s, W_{tag}) = (30/34)*(22/30) = 0.647$

Case-2: When source word W_s is "ताज़ा" and Tag W_{tag} is "JJ" and target word W_t is "புதுப்பித்த" (Puthupithu means New), then $P(W_s, W_{tag}|W_t) = 8/10$

If the target word W_t is புதுப்பித்த and precedence is nothing, then $P(W_t) = 1/10$

So combined probability $P(W_t|W_s, W_{tag}) = (8/10)*(1/10) = 0.08$

Out of these probabilities first one is highest so the target text is "புத்தாணர்ச்சியான".

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Similarly, the other words are translated. The final translated output will be பாத்தாணர்ச்சியான\JJ சாவாசம்\N_NN மற்றாம்\CC_CSD பளபளப்பான\JJ பற்கள்\N_NN தங்களின்\PR_PRP தோற்றத்தை\N_NN நிர்ணயிக்கிறதா\V_VM_VF .\RD_PUNC

RESULTS AND EVALUATION

This system has been evaluated using various sentences (restricted to health domain) and the Bilingual Evaluation Understudy (BLEU) score is calculated to be 0.68. The precision and recall of the system are found to be 0.85 and 0.7 respectively. As shown in the Table 4, the system is found to give very good results by introducing word sense based translation which is compared with its counterpart without word sense disambiguation (WSD). The precision and recall gradually improves with increase in corpus size. This is due to inclusion of additional words to the corpus which makes it mature enough compared to the lesser corpus size.

CONCLUSION

The basic advantage is that the algorithm considers the sense of the word during translation process. In addition to that, the introduced algorithm also tends to increase accuracy of the translation as compared to parallel translation or any other existing methods for language translation. The algorithm calculates probability of occurrence for individual word, precedence word and tagging of words during the translation. The precision and recall of the algorithm are 87% and 86.5% respectively, which shows considerable improvement compared to the system without word sense disambiguation, whose precision and recall are found to be 83% and 79% respectively.

Although the algorithm tends to improve the results, but still there is scope for modifications that can further improve the accuracy of the output. One of such modifications can be introducing good part-of-speech tagging in the source language. Another modification can be making the algorithm machine intelligent so that if the input word is not contained in the corpus, the algorithm intelligently derives proper translation. These further modifications, if implemented in future can improve results in machine translation from Hindi to Tamil.

Table 4. Comparison of the precision and recall

S. No.	Corpus Size (in number of words)	Statistical Machine Translation(with WSD)		Statistical Machine Translation(without WSD)	
		Precision (in %)	Recall (in %)	Precision (in %)	Recall (in %)
1	10000	65	63	58	57
2	20000	76	72.5	71	72
3	30000	87	86.5	83	79

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Abbache, A., Meziane, F., Belalem, G., & Belkredim, F. Z. (2016). Arabic Query Expansion Using WordNet and Association Rules. *International Journal of Intelligent Information Technologies*, *12*(3), 51–64. doi:10.4018/IJIIT.2016070104

Alguliyev, R. M., Aliguliyev, R. M., Isazade, N. R., Abdi, A., & Idris, N. (2017). A Model for Text Summarization. *International Journal of Intelligent Information Technologies*, *13*(1), 67–85. doi:10.4018/IJIIT.2017010104

Ananthakrishnan, R., Bhattacharyya, P., Hegde, J. J., Shah, R. M., & Sasikumar, M. (2008). Simple Syntactic and Morphological Processing Can Help English-Hindi Statistical Machine Translation. *Proceedings of International Joint Conference on Natural Language Processing*.

Balaji, J., Geetha, T. V., & Parthasarathi, R. (2014). A Graph Based Query Focused Multi-Document Summarization. *International Journal of Intelligent Information Technologies*, *10*(1), 16–41. doi:10.4018/ijiit.2014010102

Bharati, A., Husain, S., Misra, D., & Sangal, R. (2009). Two stage constraint based hybrid approach to free word order language dependency parsing. *Proceedings of the 11th International Conference on Parsing Technologies IWPT '09*. 10.3115/1697236.1697251

Bharati, A., & Sangal, R. (1993). Parsing free word order languages in the Paninian framework. *Proceedings of the 31st Annual Meeting on Association for Computational Linguistics*. 10.3115/981574.981589

Bharati, A., Vineet, C. P., Kulkarni, P. A., & Sangal, R. (1991). Local Word Grouping and Its Relevance to Indian Languages. In *Frontiers in Knowledge Based Computing (KBCS90) V. P. Bhatkar and K. M. Rege edn*. New Delhi: Narosa Publishing House.

Bharati, A., Vineet, C., P Kulkarni, P.A., Sangal, R. (1997). Anusaaraka, Machine translation in Stages. *A Quarterly in Artificial Intelligence*, *10*(3), 22-25.

Darbari, H. (1999). *Computer-assisted translation system – an Indian perspective. Proceedings of the Machine Translation Summit VII* (pp. 80–85). Singapore: Kent Ridge Digital Labs.

Goyal, V. (2010, October). Development of a Hindi to Punjabi Machine Translation System. Retrieved from http://www.languageinindia.com

Goyal, V., & Lehal, G. S. (2008). Hindi Morphological Analyzer and Generator. *Proceedings of the 2008 First International Conference on Emerging Trends in Engineering and Technology*. 10.1109/ICETET.2008.11

Goyal, V., & Lehal, G. S. (2009). A Machine Transliteration System for Machine Translation system: An Application on Hindi-Punjabi Language Pair. *Atti Della Fondazion Giorgio Ronchi (Italy)*, *LXIV*(1), 27–35.

Goyal, V., & Lehal, G. S. (2010). Web Based Hindi to Punjabi Machine Translation System. *Journal of Emerging Technologies in Web Intelligence*, *2*(2). doi:10.4304/jetwi.2.2.148-151

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hema, R., & Geetha, T. V. (2016). Recognition of Chemical Entities using Pattern Matching and Functional Group Classification. *International Journal of Intelligent Information Technologies*, *12*(4), 21–44. doi:10.4018/IJIIT.2016100102

Korra, R., Sujatha, P., Chetana, S., & Kumar, M. N. (2011). Performance evaluation of Multilingual Information Retrieval (MLIR) system over Information Retrieval (IR) system. *Proceedings of the 2011 International Conference on Recent Trends in Information Technology (ICRTIT)*. 10.1109/ICRTIT.2011.5972453

Kumar, K. V., Yadav, D., & Sharma, A. (2015). Graph Based Technique for Hindi Text Summarization. In *Advances in Intelligent Systems and Computing Information Systems Design and Intelligent Applications* (pp. 301-310). doi:10.1007/978-81-322-2250-7_29

Kumar, P., & Goyal, V. (2010). Development of Hindi-Punjabi Parallel Corpus Using Existing Hindi-Punjabi Machine Translation System and Using Sentence Alignments. *International Journal of Computers and Applications*, *5*(9), 15–19. doi:10.5120/941-1319

Pandian, S. L., & Geetha, T. (2008). Morpheme based Language Model for Part-of-Speech Tagging. *Polibits*, *38*, 19–25. doi:10.17562/PB-38-2

Ray, P. R., Harish, V., Basu, A., & Sarkar, S. (2003). Part of speech tagging and local word grouping techniques for natural language parsing in Hindi. *Proceedings of the International Conference on Natural Language Processing*.

Saraswati, J., Shukla, R., Goyal, R. P., & Bhattacharyya, P. (2010). Hindi to English Wordnet Linkage: Challenges and Solutions. *Proceedings of the 8th International Conference on Natural Language Processing*.

Saravanan, K., Parthasarathi, R., & Geetha, T. V. (2003). Syntactic Parser for Tamil. *Proceedings of the Tamil Internet Conference*, Chennai, Tamilnadu, India (pp. 28-37).

Sinha, R. M. K. (2004). An Engineering Perspective of Machine Translation: AnglaBharti-II and AnuBharti-II architectures. *Proceedings of International Symposium on Machine Translation, NLP and Translation Support System (iSTRANS)* (pp. 1-9).

Sinha, R. M. K., & Thakur, A. (2005). Machine Translation of bi-lingual Hindi-English (Hinglish) tex. Proceedings of the tenth Machine Translation Summit, MT Summit X, Phuket, Thailand (pp. 149-156).

Yadav, R. K., & Gupta, D. (2010). Annotation Guidelines for Hindi-English Word Alignment. *Proceedings of the 2010 International Conference on Asian Language Processing*. 10.1109/IALP.2010.58

This research was previously published in the International Journal of Intelligent Information Technologies (IJIIT), 14(1); edited by Vijayan Sugumaran; pages 17-27; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 22
Exploiting Chi Square Method for Sentiment Analysis of Product Reviews

Nilesh M Shelke
Priyadarshini Indira Gandhi College of Engineering, India

Shrinivas P Deshpande
DCPE, HVPM, India

ABSTRACT

Sentiment analysis is an extension of data mining which employs natural language processing and information extraction task to recognize people's opinion towards entities such as products, services, issues, organizations, individuals, events, topics, and their attributes. It gives the summarized opinion of a writer or speaker. It has received lot of attention due to increasing number of posts/tweets on social sites. The proposed system is meant to classify a given text of review into positive, negative, or the neutral category. Primary objective of this article is to provide a method of exploiting permutation and combination and chi values for sentiment analysis of product reviews. Publicly available freely dictionary SentiWordNet 3.0 has been used for review classification. The proposed system is domain independent and context aware. Another objective of the proposed system is to identify the feature specific intensity with which reviewer has expressed his opinion. Effectiveness of the proposed system has been verified through performance matrix and compared with other research work.

INTRODUCTION

As the publication of posts/opinion/thoughts/tweets on social sites/blogs/forums became easy, internet sites are flooding with product reviews. Reviews are useful to customers for taking decisions for purchasing, institutes/companies for branding, politicians for making image in the public, etc. This public generated review holds sentiments for various products/persons/movies etc. But their browsing is tedious as they are huge in number. Sentiment Analysis from text is multi-facetted research issue. Even though

DOI: 10.4018/978-1-7998-0951-7.ch022

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

there are tremendous challenges and many difficulties in sentiment analysis from text; its contribution and applications to industry is not ignorable. Because sentiment analysis bases its results on factors that are so inherently humane, it will become one the major drivers of many business decisions in future. Practical applications of sentiment analysis are huge. Sentiment analysis can be successfully applied to learn user preferences and interests from users' personal writings. These methods are often studied for consumer feedback analysis. Similarly, e-learning systems can benefit from affective tutoring approaches.

Sentiment Analysis is an extension of data mining which employs Natural Language Processing and Information Extraction task to recognize people's opinion towards entities such as products, services, issues, organizations, individuals, events, topics and their attributes. Sentiment Analysis analyze each text and detect which part contain opinionated word, which is being opinionated and who has written the opinion. It gives the summarized opinion of a writer or speaker. Sentient analysis can be done at word level, sentence level and document level. In a broader sense, sentiment analysis is focused to detect the attitude of a speaker or a writer with respect to some topic or about overall coverage of a document (Zhao et al., 2015). Sentiment can be defined as

Sentiment:= $<o_j, f_{jk}, so_{ijkl}, h_i>$,

o_j is the target: what/whom the sentiment is expressed to.

f_{jk} is a feature of the object o_j,

so_{ijk} is the sentiment value of the opinion of the opinion holder h_i *on feature* f_{jk} *of object* o_j*;* so_{ijkl} *is +ve,-ve, or neutral, or a more granular rating.*

*hi is h*older: who expresses the sentiment (Pawar, 2016).

Identification of sentiments or moods embedded in the (written) natural language is critical and a challenging issue. Researchers are attempting to apply different approaches and techniques and extract useful information from this data.

As such demand of sentiment analysis applications is increasing rapidly. Primary tasks in sentiment analysis is finding the "aspects" or "features", finding their corresponding sentiments, identifying the intensity of the sentiments and finally calculating the overall opinion of the reviewer. Consider some actual reviews from the dataset.

1. The bread is top *notch* as well.
2. In fact, this was not a Nicosia salad and was *barely eatable*.
3. I have to say they have one of the *fastest* delivery times in the city.
4. The food is always *fresh* and *hot* - ready to eat!
5. Did I mention that the coffee is *outstanding*?
6. Certainly not the *best* sushi in New York, however, it is always *fresh* food, and the place is *very clean, sterile*.

In the above example reviews, underlined terms are the product features and italic terms indicates the sentiments about those features termed as sentiment features.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

This paper is intended for exploiting chi square method for sentiment analysis. Major contributory research work on chi square method for exploiting feature selection from text do not occur. Researchers have used chi square method mainly in statistics for verifying the interdependence of two events. In particular, chi square can be exploited to test the dependency of some special terms with specific classes. In sentiment analysis, researcher has employed it for the classification of any review to a particular class like positive or negative or to study the interdependence between feature and a class as by (Janaki Meena et al., 2010). In the proposed system chi square method is employed for feature selection. Chi square method has been exploited for fine feature selection.

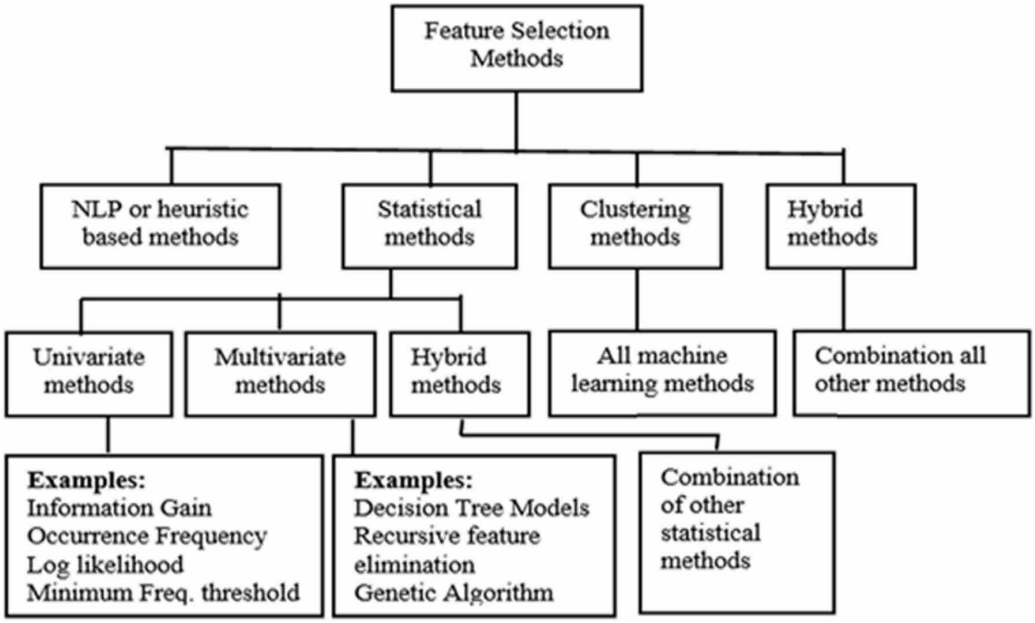
This paper is organized as follows: Section II details the work done on sentiment analysis by various researchers; Section III introduces proposed system which includes algorithms for feature extraction, the aspect-oriented sentiment analysis; Section IV evaluates the proposed work with existing works and finally conclusion has been provided.

LITERATURE SURVEY

Figure 1 is the summary of feature selection methods researchers worked upon. Natural language processing techniques includes locating the terms near subjective expression which acts as features and take the help of POS Taggers like noun and noun phrases which are feature rich (Heeryon et al., 2016).

Statistical methods are highly popular for feature extraction. Univariate works one at a time on attributes. These techniques are computationally efficient. Demerit of these techniques is that they do not work upon interaction amongst the attributes. Univariate methods like Information Gain for feature extraction are implemented by (Gao, 2014; Ghosal, 2016; Ferruh, 2014; Jiamin, 2015). Study has coined a new term, Maximum Discrimination (MD) a statistical method for implementation of feature selection (Tang, 2016).

Figure 1. Summary of features selection techniques



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Statistical feature selection multivariate methods like Decision Tree Model has been used (Hwang, 2002). Statistical binary bat algorithm has been implemented (Rani and Rajlakshmi, 2015). Recursive feature elimination method has been implemented by (Abualigah, 2016). Previous studies have employed genetic algorithm for feature selection. Only drawback of multivariate methods is that they are computationally expensive unlike univariate methods. Multivariate methods word on group of attributes. Hybrid techniques use univariate, multi-variate and other methods for high accuracy (Zhao and Yafei Wang, 2010; Ahmad, 2015; Singh and Kakkar, 2015).

Feature clustering for feature extraction using PMI has been proposed (Jeyadurga, 2004). Depending upon frequency, a head word of the cluster is selected. This head word is then compared with other word for PMI similarity and thus features are decided. However, drawback of such technique is that feature clustering loses minor features and can collect major features only (Zhang, 2011).

Many researchers achieved this task of feature selection by machine learning Approaches (Su, 2018; Zhang, 2011; Masashi, 2016; Zhang, 2012; Lu, 2009; Zhai, 2011; Kharde, 2016). Conditional Random Fields (CRF) also can be applied for feature extraction (Batcha and Azi, 2013). However, demerit of machine learning approaches is that they require huge data and techniques which are data dependent.

(Wu et al., 2014) found better results with statistical machine learning than semantic approach and accuracy increases with increase in training dataset. Work was done for financial website Sina Finance. Study could establish the relationship between stock price volatility and sentiments expressed by the stakeholder. (Xinjie Zhou et al., 2016) tried experiments on Chinese CMSAE dataset. Proposed system called CMiner is used for aspect orientation and opinion summarization. However, methodology proposed is data dependent and work has been done for Chinese reviews only.

The mutual information measure (PMI) method has been employed by the researchers for identifying the correspondence between the features and the classes. The point-wise mutual information (PMI) gives the co-occurrence strength. This is the basic PMI. Another mostly used feature selection technique is Latent Semantic Indexing, accompanied by Principal Component Analysis wherein, reduced dimensionality technique is used by (Walaa Medhat, 2014). (Baumgarten et al, 2013) attempted sentiment mining for twitter using keywords. But this attempt carries with an inherent problem of keywords like *Ambiguity in Keyword Definitions, Incapability of Recognizing Sentences without Keywords*, *Lack of Linguistic Information and* Poor recognition of affect when negation is involved.

PROPOSED SYSTEM

Figure 2 depicts architecture of the proposed system. Reviews collected from amazon.com is of varying domains like restaurants, camera, computer, antivirus, vehicles. Reviews are un-structured data with slangs, abbreviations like TTYL, B4N! …, Gr8.., CU, … and noise like smiley, special Symbols (!, ?, ….). This necessitates slang conversion and care of symbols in the beginning. After pre-processing input comment is given to feature selection method which outputs features of products about which author has expressed his/her reviews. Feature selection is done by term frequency parameters. Figure 3 depicts the flowchart for feature selection by Term Frequency. Term frequency for every term is calculated by the formula

$$tf_{scorei} = \text{count_of}(x_i)/\text{size_of}(x)$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Architecture of the proposed system

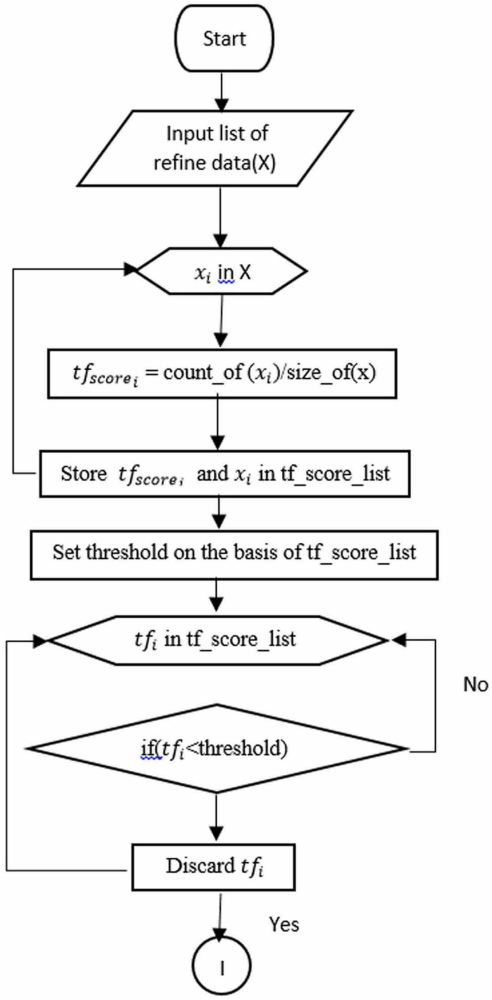
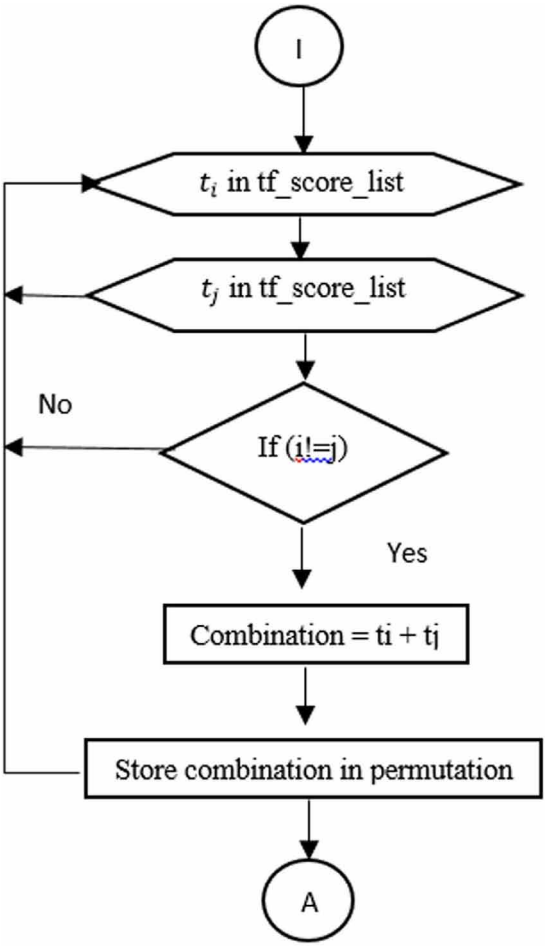


Figure 3. Feature Selection using TF-IDF score

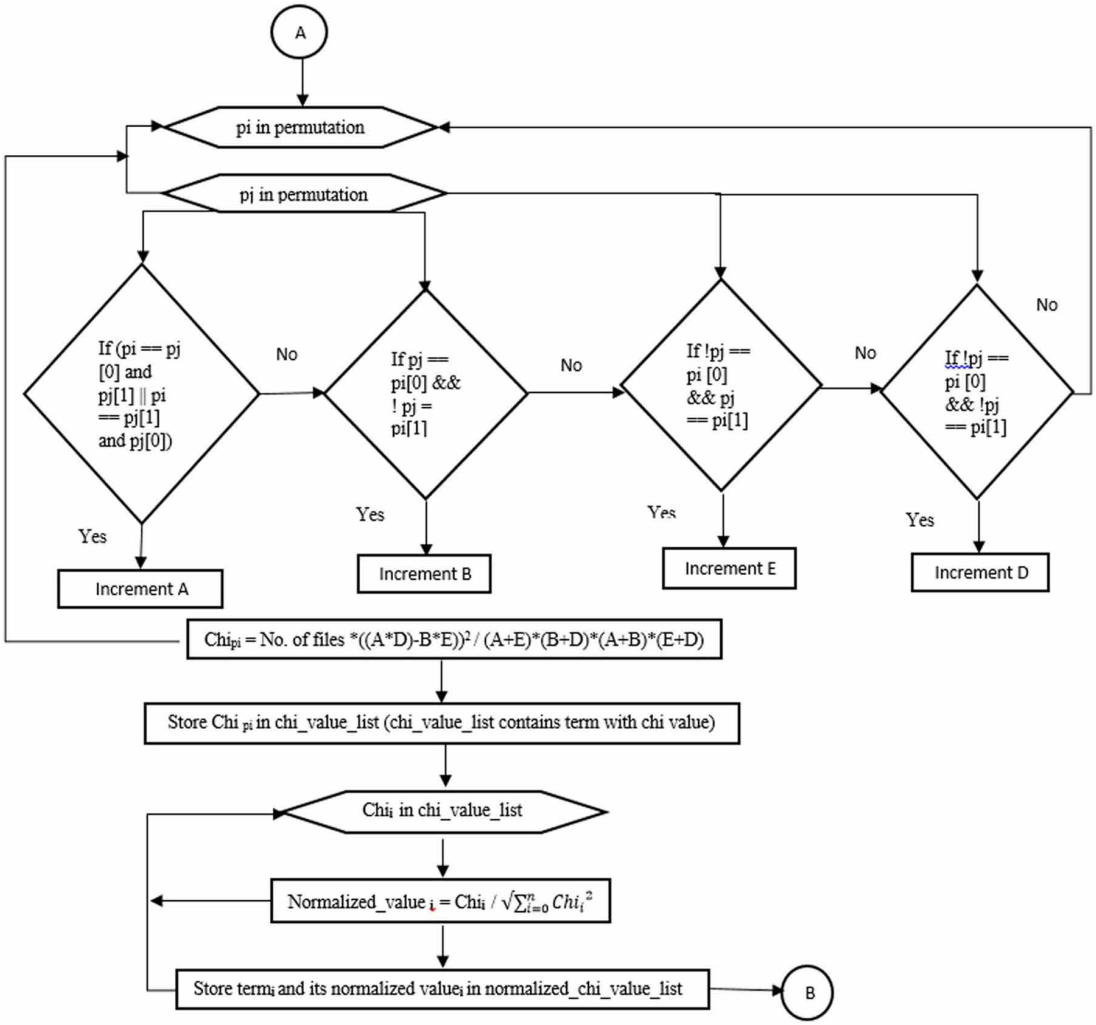


Term frequency score of every term is stored in tf_score_list. A threshold for term frequency score is decided. Terms with frequency score below this threshold are removed. This helped to remove noisy features. Further pruning of the features is done by chi square method. Chi square method has been applied in a novel way for fine pruning of the features. Figure 4 shows permutation and combination of the probable features obtained from Figure 3. Figure 4 and Figure 5 shows the mechanism of chi square method. Figure 6 depicts the flowchart for normalizing chi score values. Chi square values are fed for calculation of sentiment scores of all the features. SenitiWordNet 3.0 has been used for calculation of sentiment score. Figure 7 provides flowchart for calculation of sentiment score for the features of product reviews with proper polarity and intensity.

Mechanism of feature selection through chi square method is explained here with actual reviews from the dataset. Intentionally short reviews are taken for the sake convenience of explanation. Actual reviews contain 4-5 statements also. Feature with a character less than 3, could not observe in the dataset. Hence those words are omitted along with adjectives and adverbs from the candidate features list.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Making combination in permutation for chi square method



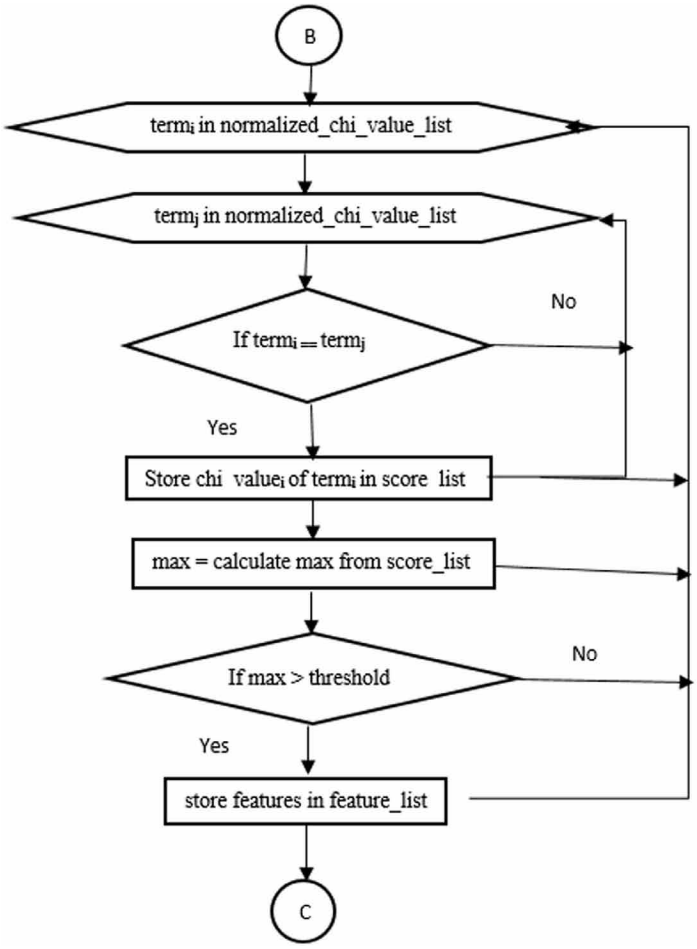
1. The bread is top notch as well.
2. In fact, this was not a Nicosia salad and was barely eatable.
3. The food is always fresh and hot - ready to eat!
4. Did I mention that the coffee is OUTSTANDING?

Permutations Starts

File line = bread notch
Permutations = bread bread, bread notch, notch bread, notch notch
----------`--------
File line = fact Nicosia salad
Permutations = fact fact, fact Nicosia, fact salad, nicoise fact, nicoise nicoise, nicoise salad, salad fact, salad nicoise, salad salad.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Fine pruning of feature selection using chi square method



File line = food fresh hot ready eat
Permutations = food food

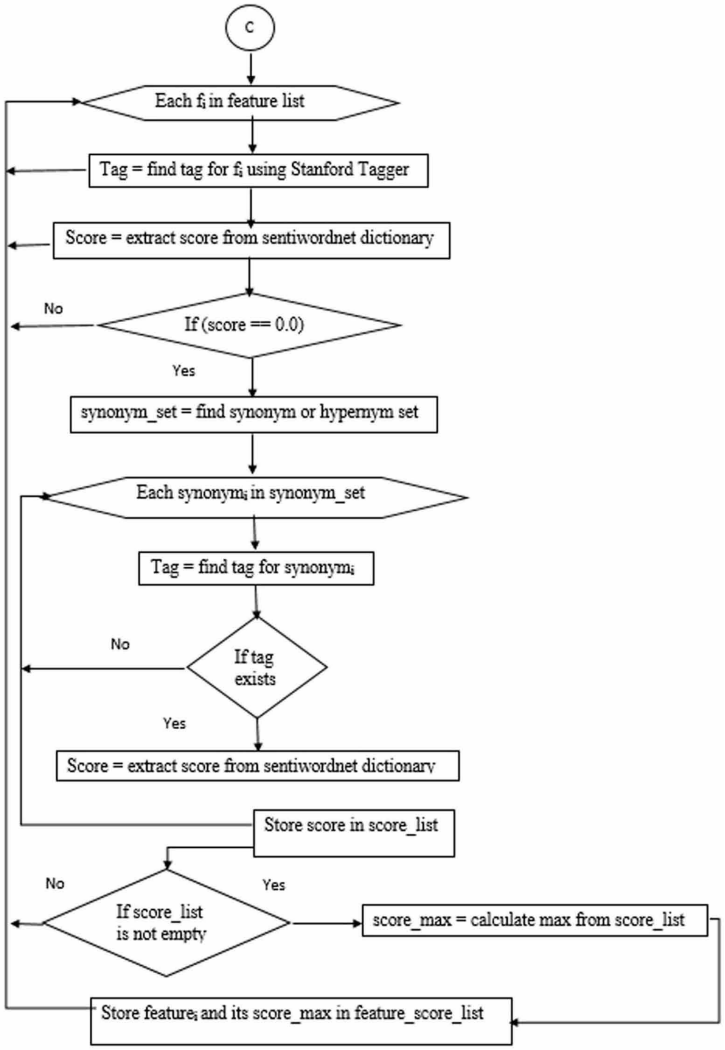
File line = mention coffee outstanding
Permutations = mention mention, mention coffee, coffee mention, permutations = coffee coffee
---

Permutations Ends

- For first term in permutation, calculate chi-value
- A = no of times both terms found in permutations
- B = no of times first term found in permutations but not second term
- E = no of times second term found in permutations but not first term
- D = no of times both terms not found in permutations

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Normalization and storing of chi square values



$$\text{Chi value} = \frac{\text{noOfFiles} \times \left(\left(\text{A} \times \text{D}\right) - \left(\text{B} \times \text{E}\right)\right) \times \left(\left(\text{A} \times \text{D}\right) - \left(\text{B} \times \text{E}\right)\right)}{\left(A+E\right) \times \left(B+D\right) \times \left(A+B\right) \times \left(E+D\right)}$$

==========

Permutation = bread notch

A = 2, B = 8, E = 1, D = 104, feature = bread, chi value = 1.1337868480725624

Permutation = notch bread

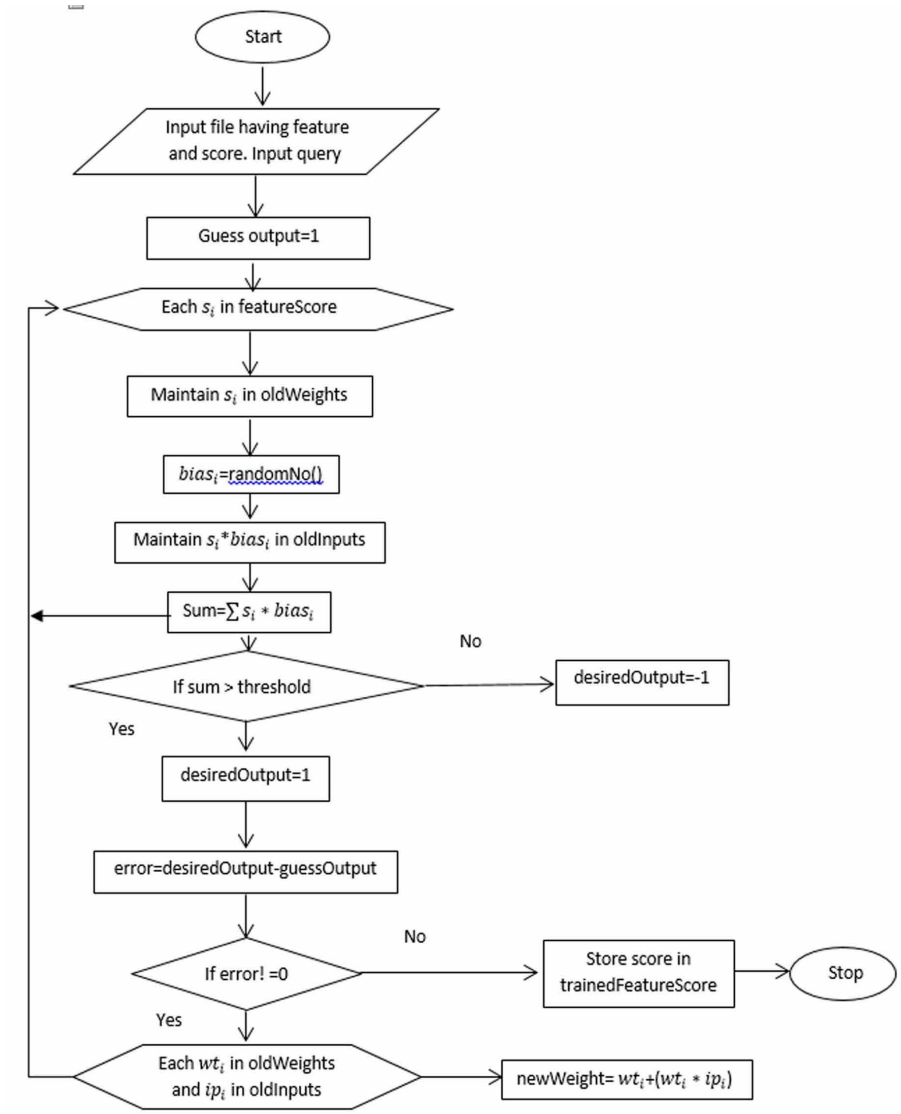
A = 2, B = 1, E = 8, D = 104, feature = notch, chi value = 1.1337868480725624

Permutation = fact nicoise

A = 2, B = 3, E = 3, D = 107, feature = fact, chi value = 1.3892561983471075

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Calculation of sentiment scores of features // Chi Test starts



Permutation = fact salad
A = 2, B = 3, E = 10, D = 100, feature = fact, chi value = 0.4251250367755222

Permutation = nicoise fact
A = 2, B = 3, E = 3, D = 107, feature = nicoise, chi value = 1.3892561983471075

Permutation = nicoise salad
A = 2, B = 3, E = 10, D = 100, feature = nicoise, chi value = 0.4251250367755222

Permutation = salad fact, A = 2, B = 10, E = 3, D = 100, feature = salad, chi value = 0.425

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Permutation = salad nicoise, A = 2, B = 10, E = 3, D = 100, feature = salad, chi value = 0.42512503

Permutation = mention coffee, A = 2, B = 1, E = 1, D = 111, feature = mention, chi value = 4.326

Permutation = coffee mention
A = 2, B = 1, E = 1, D = 111, feature = coffee, chi value = 4.326194019274377

- **Normalise all the chi values**
  feature = bread, normalized score = 0.14531439742822497
  feature = notch, normalized score = 0.14531439742822497
  feature = fact, normalized score = 0.1780572139105606
  feature = fact, normalized score = 0.05448712750170592
  feature = nicoise, normalized score = 0.1780572139105606
  feature = nicoise, normalized score = 0.05448712750170592
  feature = salad, normalized score = 0.05448712750170592
  feature = salad, normalized score = 0.05448712750170592
  feature = mention, normalized score = 0.55447660037437
  feature = coffee, normalized score = 0.55447660037437
- Finally, find maximum Chi value for a feature
  feature = bread, max chi value = 0.14531439742822497
  feature = notch, max chi value = 0.14531439742822497
  feature = fact, max chi value = 0.1780572139105606
  feature = nicoise, max chi value = 0.1780572139105606
  feature = salad, max chi value = 0.05448712750170592
  feature = mention, max chi value = 0.55447660037437
  feature = coffee. Max chi value = 0.55447660037437
  // Chi Test ends

Those terms which have more than one normalize chi value, their maximum score have taken into consideration. Thus this process outputs fine product features. Flow of this mechanism is depicted in Figure 4 and Figure 5.

For calculation of sentiment scores of features, SentiWordNet 3.0 (SWN) is exploited which publicly available freely. SWN contains a word with many senses as found by (Apoorva et al., 2016) Hence, only those senses are considered which POS tag matches with sentiment feature expressed in the review. If still multiple senses occur, then following formula is used for calculation of sentiment score.

Score= (1/2) (1st score from SWN)*(1/3) (2nd score from SWN)…….

This mechanism is depicted in Figure 6. In case sentiment feature expressed in review is not available in SWN, it's all synonyms are searched. In case of unavailability of synonym its hypernyms are searched. Synonym or hypernym with maximum score is selected. Flowchart 7 is used for polarity classification.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ALGORITHM: Determining the Polarity of Reviews
Input: Single Comment.
Output: Polarity, Intensity.

Step 1: Select Comment.
Step 2: Apply POS tagging using Stanford POS tagger.
Step 3: Get features from feature selection algorithm
Step 4: Take Noun phrase and verb Phrase using Stanford parser.
Step 5: If it is Noun phrase then for every Noun Phrase, check Noun and Adjective. If it is a Noun then calculate the polarity with the formula
PolarityN=$\Sigma^{i}_{k=0}$ $[(1-Ni)*Ni]$
and if it is adjective then calculate it with the fomula
PolarityA= $\Sigma ik^{=}_{0[}(1-Ai)*Ai]$
Setp 6:
If it is a Verb phrase, then for every Verb phrase, check for verb and adverb.
If it is a verb, calculate the polarity as
PolarityV=Σ $ik^{=}_{0[}(1-Vi)*Vi]$ and
if it is adjective, calculate it as
PolarityA=Σ $ik=0_{[(1}-Advi)*Advi]$
Step 7: Intensity= polarityN + polarityA;
Step 8: classification
If intensity>0
Positive;
Else
Negative;

ALG**ORITHM 5.2: Fe**ature Calculation and Score Calculation

Step 1: Get N,A,V,Adv as word
Step 2: find word in to SentiWordNet
If(found)
Score=(1/2)(1st score from SWN)*(1/3)(2nd score from SWN)…….
Return score;
Step 3: Else
Apply stemming and get set of synonyms from WSD
For each synonym
Calculate score for each word.
End for.
End else
Step 4: Return max (score);
Step 5: end

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

RESULT ANALYSIS

The results obtained from the implemented methods and their comparison with existing approaches are discussed here. The proposed approaches are implemented using NetBeans IDE 8.0.2. 200 reviews each of camera, computer, Norton Antivirus and more than 600 reviews of restaurant are tested. For clarity and uniformity, results of average of 100 reviews are shown here. Figures 8 and 9 show graphs with improvement of work and accuracy comparison. Tables 1-4 show information regarding restaurant domain.

The performance of a developed model is determined by how good the classifications (predictions) reflect the actual events. There are two possible predicted classes: "POSITIVE" review and "NEGATIVE" review (AlShahrani, 2017).

- **True Positives (TP):** These are cases in which proposed system predicted POSITIVE (they are positive reviews), and they are actually the POSITIVE reviews.
- **True Negatives (TN):** System predicted NEGATIVE, and they are actually NEGATIVE reviews.
- **False Positives (FP):** System predicted POSITIVE, but they are actually NEGATIVE reviews.
- **False Negatives (FN):** System predicted NEGATIVE, but they actually are POSITIVE reviews (Dey, 2016).

There is a list of rates that are often computed from a confusion matrix for a binary classifier: (Saba et al, 2016; Virmani et al, 2016; Nath et al, 2016).

- **Accuracy:** Overall, how often the classifier is correct can be calculated by

$$A = \frac{\left(TP + TN\right)}{\left(TP + TN + FP + FN\right)}$$

- **Recall**

$$R = \frac{\text{No: of sentences correctly classified as class } x}{\text{Actual No: of sentences of class } x}$$

- **Precision:** When it predicts POSITIVE, how often is it correct?

$$P = \frac{\text{No: of sentences correctly classified as class}}{\text{No of sentences classified as class } x}$$

- **F Score:** This is a weighted average of the true positive rate (recall) and precision.

$$F - score = \frac{2 * Precision * Recall}{\left(Precision + Recall\right)}$$

Proposed work is the extended work done in (Nilesh Shelke et al., 2016). Proposed method outperforms the previous work.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Evaluation parameters for restaurant domain

Methods	TP	FP	TN	FN
Previous work (Nilesh Shelke et al., 2016)	62	9	25	4
Proposed System	77	6	12	5

Table 2. Performance matrix for restaurant domain

Methods	Precision	Recall	FM	Accuracy
Previous work (Nilesh Shelke et al., 2016)	0.8732	0.7126	0.7847	0.87
Proposed System	0.9277	0.8651	0.8953	0.89

CONCLUSION

The major contribution of the research work is extraction of features from product reviews for which chi square method has been implemented. Another contribution is that the methodology applied is dataset independent i.e. same method is useful for all the datasets of product reviews. Feature selection task addressed through chi square method is context aware also. It has been observed that feature position is more important than feature frequency. The proposed system is an extended work done in (Nilesh Shelke et al, 2016). Spurious features could observe in the previous work. Hence, they have been pruned by TF-IDF measure and then by employing chi-square method. These methods could extract concrete features from the reviews.

Table 3. Comparison with (M. Govindarajan, 2014)

Algorithm	Accuracy
Naïve Bayes	85
Support Vector Machine	85.2
Genetic Algorithm	85.3
Previous work (Nilesh Shelke et al., 2016)	87
Proposed System	89

Table 4. Proposed work compared with (Aishwarya Mohan, 2014)

Methods	Precision	Recall	FM	Accuracy
Naïve Bayes	0.851	0.85	0.85	0.85
KNN	0.866	0.865	0.865	0.865
Previous work (Aishwarya Mohan, 2014)	0.8732	0.7126	0.7847	0.87
Proposed System	0.9277	0.8651	0.8953	0.89

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Graph showing improvement over the previous work done by (Nilesh Shelke et al., 2016)

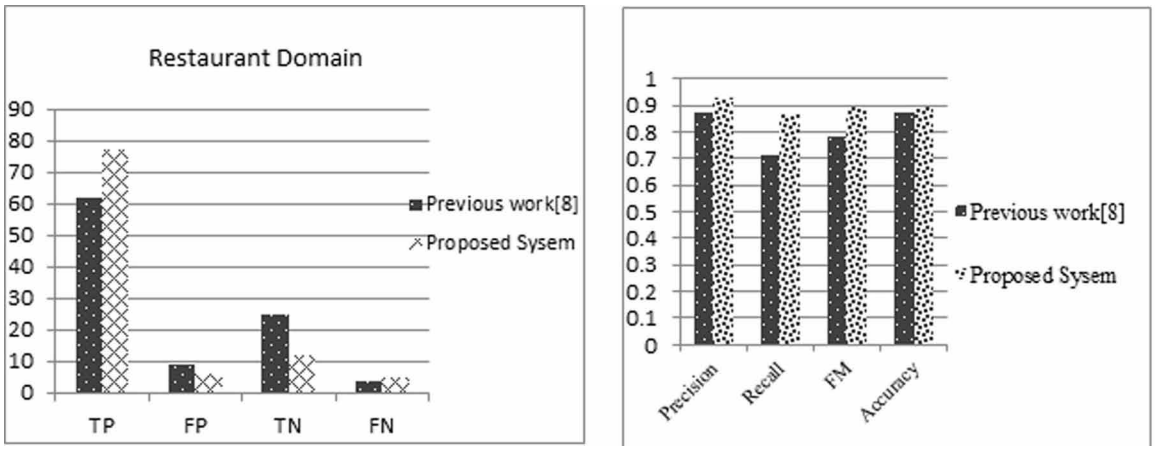
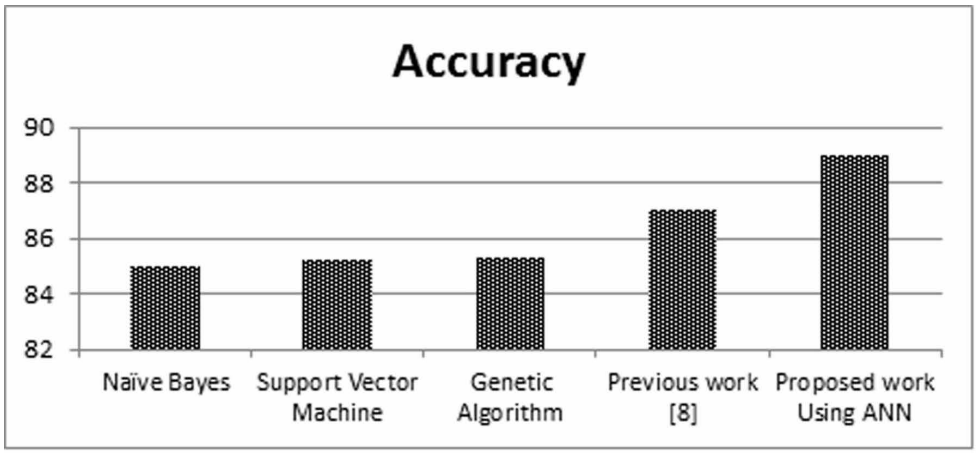


Figure 9. Graph showing Accuracy comparison with (M. Govindarajan, 2014)



Some of the limitation of the proposed system is that it cannot handle the implicit features. For example, 'this camera is very heavy', leads to the weight feature. Also work needs to be done on negative sentences like 'camera is too big'. Moreover, there are limitations in categorizing an opinion in comparative sentences. One has to be dependent on the order of words in comparative sentences to determine the opinion e. g. The review, "Restaurant Taj is better than restaurant Haldiram" depicts completely different opinion from "Restaurant Haldiram is better than restaurant Taj". Concept based methods needs to be incorporated for such types of feature extraction. There are enormous future perspectives to extend the proposed work. Building information retrieval models, designing Human-Computer Interfaces (HCI) affective interfaces, building educational dialog systems, developing expressive text-to-speech synthesis, conducting subjective natural language processing tasks, improving text summarization or machine translation, and including affective semantics for foreign language learning etc. can be possible with exploration of sentiment analysis and emotion extraction.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

M., Aishwarya,, Manisha, R., Vijayaa, B, & Naren, J. (2014). An approach to perform aspect level sentiment analysis on customer reviews using sentiscore algorithm and priority based classification (IJCSIT). *International Journal of Computer Science and Information Technologies*, *5*(3), 4145–4148.

AlShahrani, A. M., Al-Abadi, M. A., Al-Malki, A. S., Ashour, A. S., & Dey, N. (2017). Automated System for Crops Recognition and Classification. In N. Dey, A. Ashour, & S. Acharjee (Eds.), *Applied Video Processing in Surveillance and Monitoring Systems* (pp. 54–69). Hershey, PA: IGI Global; doi:10.4018/978-1-5225-1022-2.ch003

Singh, A., & Kakkar, V. (2015). A Study of Text Mining Techniques and Optimization Using Genetic Algorithm. *International Journal of Advance Foundation and Research in Science and Engineering*, *1*, 1–7.

Sharma, A., & Dey, S. (2012). Performance Investigation of Feature Selection Methods and Sentiment Lexicons for Sentiment Analysis. *International Journal of Computer Applications on Advanced Computing and Communication Technologies for HPC Applications - ACCTHPCA*, 15-20.

Agarwal, A., Sharma, V., Sikka, G., & Dhir, R. (2016). Opinion Mining of News Headlines using SentiWordNet, In *Symposium on Colossal Data Analysis and Networking*. IEEE Explore.

Baumgarten, Mulvenna and Runney. (2013). Keyword-Based Sentiment Mining using Twitter, *International Journal of Ambient Computing and Intelligence, 5*(2), 56-69.

Bo Tang, Steven Kay, Haibo He. (2016). Toward Optimal Feature Selection in Naïve Bayes for Text Categorization, published in IEEE Transaction on Data and Knowledge Engineering, 14-20.

Wu, D. D., Zheng, L., & Olson, D. L. (2014). A Decision Support Approach for Online Stock Forum Sentiment Analysis. *IEEE Transactions on Systems, Man, and Cybernetics. Systems*, *44*(8), 1077–1087. doi:10.1109/TSMC.2013.2295353

Dey, N. (Ed.). (2016). *Classification and clustering in biomedical signal processing*. IGI Global. doi:10.4018/978-1-5225-0140-4

Adnan, D., & Fei, S. (2012). Feature Selection for Sentiment Analysis Based on Content and Syntax Models. *Decision Support Systems*, *53*(4), 704–711. doi:10.1016/j.dss.2012.05.023

Li, F., & Yang, Y. (2005). Analysis of Recursive Feature Elimination Method. In *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 633-634). ACM. 10.1145/1076034.1076164

Yĭğit, F., & Baykan, Ö. K. (2014, November). A new feature selection method for text categorization based on information gain and particle swarm optimization. In *2014 IEEE 3rd International Conference on Cloud Computing and Intelligence Systems (CCIS)* (pp. 523-529). IEEE.

Govindarajan, M. (2014). Sentiment Analysis of Restaurant Reviews using Hybrid Classification Method. In *Proceedings of 2nd IRF International Conference*, Chennai, India (pp. 127-130).

Masashi, H., Shimada, K., & Endo, T. (2016). Aspect Identification of Sentiment Sentences using a Clustering Algorithm [in Japanese]. *Journal of Social and Behavioral Sciences*, *27*, 22–31.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Cho, H., & Lee, J. S. (2016, January). Data-driven feature word selection for clustering online news comments. *In 2016 International Conference on Big Data and Smart Computing (BigComp)* (pp. 494-497). IEEE.

Janaki Meena, M., Chandran, K. R., Mary Brinda, J., & Sindhu, P. R. (2010). Enhancing Feature Selection using Statistical Data with Unigrams and Bigrams. *International Journal of Computers and Applications*, *1*(11), 7–11. doi:10.5120/250-407

Jeyadurga, P., Vijaya Lakshmi, P., & Kanchana, J. (2004). PMI Based Clustering Algorithm for Feature Reduction in Text Classification. *International Journal of Innovative Research in Science and Engineering and Technology*, *3*(3), 1733–1740.

Xu, J., & Jiang, H. (2015). An Improved Information Gain Feature Selection Algorithm for SVM Text Classifier. In *International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)* (pp. 273–276.

Abualigah, L. M., Khader, A. T., & Al-Betar, M. A. (2016). Unsupervised Feature Selection Technique based on Genetic Algorithm for Improving the Text Clustering. In *7th International Conference on Computer Science and Information Technology (CSIT)* (pp. 1–6). 10.1109/CSIT.2016.7549453

Lu, Y., Zhai, C., & Sundaresan, N. (2009). Rated Aspect Summarization of Short Comments. In *Proceedings of the 18th International Conference on World Wide Web* (pp. 131–140). ACM.

Cao, J., Zeng, K., Wang, H., Cheng, J., Qiao, F., Wen, D., & Gao, Y. (2014). Web-Based Traffic Sentiment Analysis: Methods and Applications. *IEEE Transactions on Intelligent Transportation Systems*, *15*(2), 844–853. doi:10.1109/TITS.2013.2291241

Liu, L., Wei, S., Wang, H., Chuchu, L. I., & Lu, J. (2014). A Novel Feature-based Method for Sentiment Analysis of Chinese Product Reviews. *China Communications*, 11(3), 154–164.

Nath, S. S., Mishra, G., Kar, J., Chakraborty, S., & Dey, N. (2014, July). A survey of image classification methods and techniques. In *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)* (pp. 554-557). IEEE. 10.1109/ICCICCT.2014.6993023

Shelke, N., Deshpande, S., & Thakare, V. (2017). Domain independent approach for aspect oriented sentiment analysis for product reviews. In *Proceedings of the 5th international conference on frontiers in intelligent computing: Theory and applications* (pp. 651-659). Springer Singapore.

Batcha, N. K., Aziz, N. A., & Shafie, S. I. (2013). CRF based feature extraction applied for supervised automatic text summarization. *Procedia Technology*, *11*, 426-436.

Pawar, A. Jawale, M. and Kyatanavar, D. (2016). Fundamentals of Sentiment Analysis: Concepts and Methodology. In *Environment of Computation Intelligence* (pp. 25-35). Springer International Publishing.

Kumbhar, P., & Mali, M. (2016). A survey on feature selection techniques and classification algorithms for efficient text classification. *International Journal of Science and Research*, *5*, 1267-1273.

Saba, L., Dey, N., Ashour, A. S., Samanta, S., Nath, S. S., Chakraborty, S., & Suri, J. S. (2016). Automated stratification of liver disease in ultrasound: An online accurate feature classification paradigm. *Computer Methods and Programs in Biomedicine*, *130*, 118–134. doi:10.1016/j.cmpb.2016.03.016 PMID:27208527

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ahmad, S. R., Bakar, A. A., & Yaakub, M. R. (2015, July). Metaheuristic algorithms for feature selection in sentiment analysis. In 2015 Science and Information Conference (SAI) (pp. 222-226). IEEE.

Su, Q., Xu, X., Guo, H., Guo, Z., Wu, X., Zhang, X., ... Su, Z. (2008, April). Hidden sentiment association in chinese web opinion mining. In *Proceedings of the 17th international conference on World Wide Web* (pp. 959-968). ACM.

Rani, A. S. S., & Rajalaxmi, R. R. (2015, February). Unsupervised feature selection using binary bat algorithm. In *2015 2nd International Conference on Electronics and Communication Systems (ICECS)* (pp. 451-456). IEEE.

Chen, T., Xu, R., He, Y., Xia, Y., & Wang, X. (2016). Learning User and Product Distributed Representations Using a Sequence Model for Sentiment Analysis. *IEEE Computational Intelligence Magazine*, *11*(3), 34–44. doi:10.1109/MCI.2016.2572539

Ghosal, T., Das, S. K., & Bhattacharjee, S. (2016). Sentiment analysis on Bengali Corpus. In *Annual IEEE India Conference*. IEEE.

Virmani, J., Dey, N., & Kumar, V. (2016). PCA-PNN and PCA-SVM based CAD systems for breast density classification. In *Applications of intelligent optimization in biology and medicine* (pp. 159–180). Springer International Publishing.

Kharde, V., & Sonawane, P. (2016). Sentiment Analysis of Twitter Data: A Survey of Techniques. *International Journal of Computers and Applications*, *139*(11), 5–15.

Medhat, W., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113.

Che, W., Zhao, Y., Guo, H., Su, Z., & Liu, T. (2015). Sentence Compression for Aspect-Based Sentiment Analysis, *IEEE/ACM Transactions on Audio*, *Speech, and Language Processing*, *23*(12), 2111–2123.

Zhao, W., & Wang, Y. (2010). An Improved Genetic Algorithm for Text Feature Selection. In *International Conference on Intelligent Computing and Cognitive Informatics (ICICCI)* (pp. 7-10). 10.1109/ICICCI.2010.129

Zhou, X., Wan, X., & Xiao, J. (2016). CMiner: Opinion Extraction and Summarization for Chinese Microblogs. *IEEE Transactions on Knowledge and Data Engineering*, *28*(7), 1650–1661. doi:10.1109/TKDE.2016.2541148

Xie, Y., Chen, Z., Zhang, K., Cheng, Y., Honbo, D. K., Agrawal, A., & Choudhary, A. N. (2014). MuSES: Multilingual Sentiment Elicitation System for Social Media Data. *IEEE Intelligent Systems*, *29*(4), 34–42. doi:10.1109/MIS.2013.52

Zhao, Y., Qin, B., & Liu, T. (2015). Creating a fine-grained corpus for Chinese sentiment analysis. *IEEE Intelligent Systems*, 30(1), 36–42.

Hwang, Y.-S., & Rim, H.-C. (2002). Decision Tree Decomposition-based Complex Feature Selection for Text Chunking. In *Proceedings of the 9th International Conference on Neural Information Processing* (Vol. 5, pp. 2217–2222). 10.1109/ICONIP.2002.1201887

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Zhai, Z., Xu, H., Kang, B., & Jia, P. (2011). Exploiting Effective Features for Chinese Sentiment Classification. *Expert Systems with Applications*, *38*(8), 9139–9146. doi:10.1016/j.eswa.2011.01.047

Zhang, H. Yu., & Xu, Z., M. and Shi, Y. (2011). Feature-level sentiment analysis for Chinese Product Reviews. In *3rd International Conference on Computer Research and Development* (pp. 135-140).

Zhang, S. (2011). Product Features Extraction and Categorization in Chinese Reviews. In ICCGI 2011: The Sixth International Multi-Conference on Computing in the Global Information Technology (pp. 38-42). IARIA publications.

Zhang, W., Xu, H., & Wan, W. (2012). Weakness finder: Find product weakness from Chinese reviews by using aspects based sentiment analysis. *Expert Systems with Applications*, *39*(11), 10283–10291. doi:10.1016/j.eswa.2012.02.166

Gao, Z., Xu, Y., Meng, F., Qi, F., & Lin, Z. (2014). Improved Information Gain-based Feature Selection for Text Categorization. In *4th International Conference on Information Theory and Aerospace and Electronics Systems (VITAE).* IEEE.

This research was previously published in the International Journal of Synthetic Emotions (IJSE), 9(2); edited by Amira S. Ashour and Nilanjan Dey; pages 76-93; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 23
Business Applications of Deep Learning

Armando Vieira
Redoctopus, UK

ABSTRACT

Deep Learning (DL) took Artificial Intelligence (AI) by storm and has infiltrated into business at an unprecedented rate. Access to vast amounts of data extensive computational power and a new wave of efficient learning algorithms, helped Artificial Neural Networks to achieve state-of-the-art results in almost all AI challenges. DL is the cornerstone technology behind products for image recognition and video annotation, voice recognition, personal assistants, automated translation and autonomous vehicles. DL works similarly to the brain by extracting high-level, complex abstractions from data in a hierarchical and discriminative or generative way. The implications of DL supported AI in business is tremendous, shaking to the foundations many industries. In this chapter, I present the most significant algorithms and applications, including Natural Language Processing (NLP), image and video processing and finance.

1. INTRODUCTION

Artificial Intelligence (AI) is a relatively new area of research – that started in the 50's – marked by some successes and many failures. The initial enthusiasm originated in the materialization of the first electronic computer, soon fade away with the realization that most problems that the brain do in a blink are in fact very hard to solve by machines. These problems include, locomotion in uncontrolled environments, language translation, or voice and image recognition. Despite many attempts, it also become clear that the traditional approach to solve complex mathematical equations was insufficient to solve the most basic situations that a 2 years old toddler had no difficulty – like understand basic language concepts.

This lead to the so-called long "AI winter", where many researchers simply gave up of creating machines with human level cognitive capabilities – despite some successes in between, like the IBM machine Deep Blue that become the best Chess player in the world, or the application of neural networks for hand writing digits recognition in late 80's.

DOI: 10.4018/978-1-7998-0951-7.ch023

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

But AI is today one of the most exciting research fields with plenty of practical applications, from autonomous vehicles to drug discovery, robotics, language translation and games. Challenges that seemed insurmountable just a decade ago have been solved – sometimes with super-human accuracy - and are now present in products and ubiquitous applications, like voice recognition, navigation systems, facial emotion detection and even in art creation, like music and painting. For the first time, AI is leaving the research labs materializing in products that may seems to have emerged from Science Fiction movies.

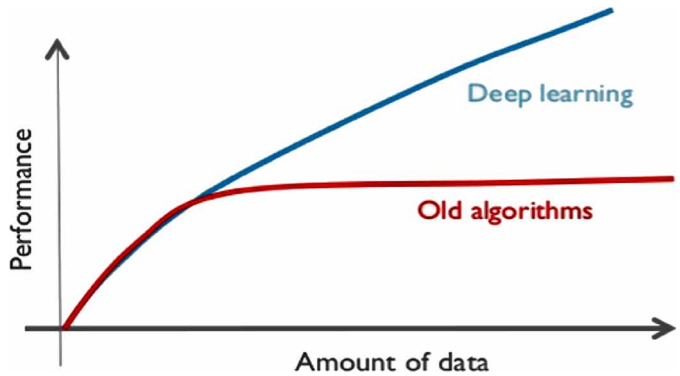
How this revolution became possible in such a short period of time? What changed in recent years that puts us closer to the General AI dream? The answer is more a gradual improvement of algorithms and hardware than a single breakthrough. But certainly on the top of the list is Deep Neural Networks, or commonly referred Deep Learning (DL) (I. Goodfellow 2006).

Artificial Neural Networks is hardly a new field. They were around for about 50 years and got some practical recognition after mid 80's with the introduction of a method (Backpropagation) that allowed training of multiple layers neural networks. However the true birth on Deep Learning may be traced to the year of 2006, when Geoffrey Hinton (R. S. G. E. Hinton 2006) presented an algorithm to efficiently train deep neural networks in an unsupervised way – data without labels. They called them Deep Belief Networks, and consisting of staked Restrictive Boltzmann Machines. They differ from previous networks since they were generative models capable to learn the statistical properties of data being presented without any supervision.

Inspired by the depth structure of the brain, deep learning architectures have revolutionized the approach to data analysis (R. S. G. E. Hinton 2006) (I. Goodfellow 2006) (Schmidhuber 2015, Schmidhuber 2015). Deep Learning networks have won a paramount number of hard machine learning contests, from voice recognition, image classification, Natural Language Processing (NLP) to time-series prediction – sometimes by a large margin. Traditionally AI relied on heavily handcrafted features, for instance, to have decent results in image classification, several pre-processing techniques have to be applied, like filters, edge detection, etc. The beauty of DL is that most, if not all, features can be learned automatically from the data – provide enough (sometimes million) training data examples are available.

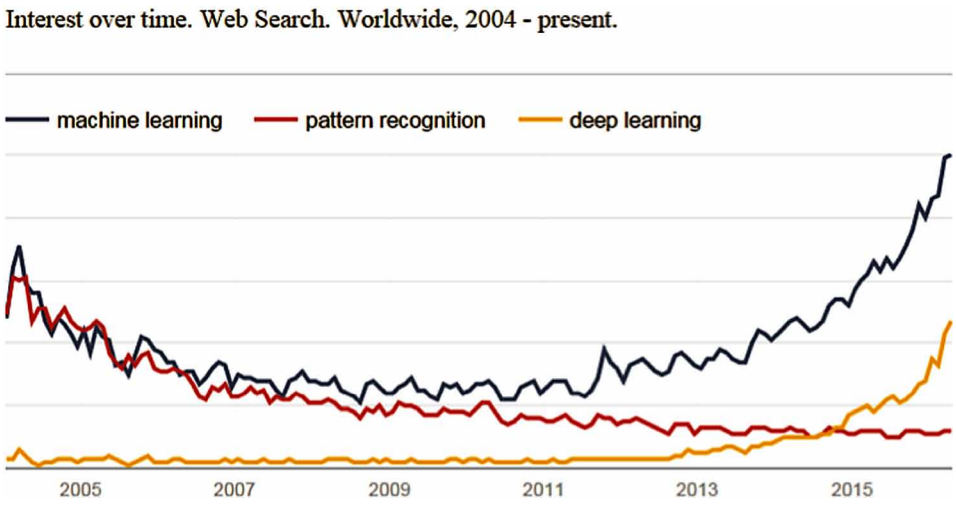
Being essentially non-supervised machines, deep neural architectures can be exponentially more efficient than shallow ones. The number of computational elements is only limited by the number of training samples – which can be of the order of billions. Deep models can be trained with hundreds of millions of weights and therefore tend to outperform shallow models such as SVMs. Moreover, theoreti-

Figure 1. DL has a higher capacity to learn from large amounts of data



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Evolution of interest on deep learning
Source: Google Trends



cal results suggest that deep architectures are fundamental to learn the kind of complex functions that represent high-level abstractions (e.g. vision, language, semantics), characterized by many factors of variation that interact in non-linear ways, making the learning process difficult.

1.1. From a Long Winter to a Blossom Spring

Today it's difficult to find any AI based technology that do not rely on Deep Learning. The implications of DL in technological applications of AI will be so profound that we may be on the verge of the biggest technological revolution of all times.

One of the remarkable features of DL neural networks is their (almost) unlimited capacity to accommodate information from large quantities of data without overfitting – as long as strong regularizers are applied. DL is much of a science as of an Art and, while it's very common to train models with billions of parameters, on millions of training examples, that is only be possible with carefully selection and fine-tuning of the learning machine.

The main characteristics that make DNN unique can be summarized in the following points:

1. High learning capacity (since DNN have millions of parameters, they don't saturate easily - the more data you have the more they learn).
2. No feature engineering required (learning can be performed end-to-end – being robotic control, language translation or image recognition).
3. Abstraction representation (DNN are capable of generating abstract concepts from data).
4. High Generative capability (DNN are much more than simple discriminative machine – they can generate unseen - but plausible - data based on latent representations).
5. Knowledge transfer (this is one of the most remarkable properties – we can teach a machine in one large set of data - images, music or biomedical data – and transfer the learning to a similar problem where less data is known. One of the most remarkable examples is DNN that capture and replicate artistic styles).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

6. Excellent unsupervised capabilities (as long as you have lots of data, DNN can learn hidden statistical representations without any labels required).
7. Multimodal learning (DNN can integrate seamlessly disparate sources of high-dimensional data, like text, images, video and audio to solve hard problems like automatic video captions generation and visual questions and answers).
8. They are relatively easy to compose and embed domain knowledge to handle uncertainty.

The less appealing aspects of DNN models are[1]:

1. They are hard to interpret - despite the ability to extract latent features from the data, DNN are black boxes that learn by associations and co-occurrences. They lack the transparency and interpretability of other methods, like decision trees.
2. They are only partially able to uncover complex causality relations or nested structural relationships, common in domains like biology.
3. They can be relatively complex and time consuming to train, with many hyper parameters that require careful fine-tuning.
4. They are sensitive to initialization and learning rate – its very easy the network to be unstable and not converge. This is particularly acute for recurrent neural networks.
5. A loss function has to be provided - sometimes is very hard to find a good one.
6. Knowledge is not incremental – each dataset has to be trained from scratch.
7. Knowledge transference is possible for certain models, but its specific for each problem.

1.2. Why DL Is Different?

Machine Learning (ML) is a vague but hardly new concept. It can be summarized in one simple sentence: finding patterns in data. These patterns can be anything from cycles in the stock market to distinguish images of cats from dogs. ML can also be described as the art of teaching machines how to make decisions. Yes, an art; perhaps more than a science.

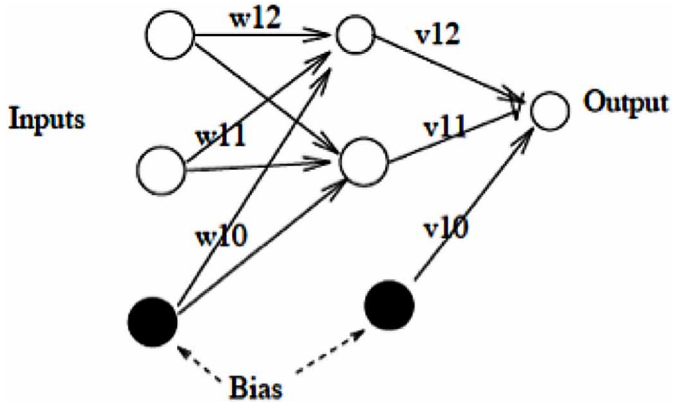
ANNs started with a work by McCullogh and Pitts who showed that sets of simple units (artificial neurons) could perform all possible logic operations - thus capable of universal computation. This work was concomitant to Von Neumann and Turing who first dealt with statistical aspects of the information processing of the brain and how to build a machine capable of reproduce them. Frank Rosembalt invented the Perceptron machine to perform simple pattern classification. However, this new learning machine was incapable of solving simple problems, like the logic XOR. In 1969 Minsky and Papert showed that Perceptrons had intrinsic limitations that could not be transcended, thus leading to a fading enthusiasm for ANNs during the 70's.

1.3. The Multilayer Preceptron

The Multilayer Preceptron (MLP) was proposed to solve problems that were not linearly separable, i.e., categories that cannot be separate by a set of straight lines. An example of a multilayered perceptron MLP is shown in next Figure:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. The multilayer preceptron



The ANN consists of a set of inputs, connected to a set of hidden unities through weights w. The hidden unities are connected to the output through weights v. Initially all the weights and the bias term are set to random numbers. The activity in the network is propagated forwards via weights from the input layer to the hidden layer where some function of the net activation is calculated, normally the transfer functions are sigmoid, *tanh* or, more recently, rectified linear unities (*ReLU*). Then the activity is propagated via more weights to the output neurons.

Two sets of weights must be updated: those between the hidden and output layers and those between the input and hidden layers. The error due to the first set of weights is calculable by Least Mean Square rule. To propagate backwards that part of the error due to the errors in the second set of weights (W) we used the backpropagation algorithm, that simply states that the errors should be proportional to the weight contribution. The algorithm has two main parameters: learning rate and momentum (to avoid traps in local minima). Also the number of unities in the hidden layer is an important input (more hidden unities will increase the computational power but it could also compromise the generalization capabilities).

The choice of the network parameters is normally performed by *n*fold cross validation: fixing *n*-1 parts of the training data for training and the remaining for testing, and then swapping these segments.

The key insight that makes this procedure relatively efficient for feedforward networks is that—simply by applying the chain rule from basic calculus—the derivatives of the error with respect to filter values in a given layer can be efficiently computed from those in the layer just above. Derivative computations thus start at the top layer and then propagate backwards through the network down to the first layers.

Another important technical innovation enabling large-scale backpropagation is stochastic gradient descent (SGD). SGD involves breaking training data into small, randomly chosen batches. Gradient descent is done on each batch in sequence until the training data are exhausted, at which point the procedure can begin again, usually on newly chosen random batches. SGD enables backpropagation on much larger data sets.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. DEEP NEURAL NETWORKS

So, why all the excitement about AI powered by Deep Learning? As mentioned before, the difference of DL is both quantitative (an improvement of 5% in voice recognition makes all the difference between a useful user interface or a useless one) and qualitative (how DL models are trained, the subtle relations they can extract from high-dimensional data and how these relations can be integrate into an unified perspective).

It was long known that ANN with more hidden layers (deeper) could have a higher computational power and be better suited to solve classification or regression problems (D. E. Rumelhart 1986). The challenge was how to train them, i.e., learn the weights or connections that link a layer of neurons to the others. The backpropagation algorithm worked fine for ANN with a single layer, but it strives to generalize for deeper architectures due to the so called vanish gradient problem – i.e., the correction signal from the output was dissipated as it travels to lower layers.

However, in 2006, Hinton et al. proposed an unsupervised learning algorithm using a method called Contrastive Divergence, that was successful in training deep generative models known as Deep Belief Networks (DBN) (R. S. G. E. Hinton 2006). A DBN is composed of a stack of Restricted Boltzmann Machines (RBMs) trained by a greedy, layer-by-layer learning algorithm (contrastive divergence). They are normally used for unsupervised tasks, but that could be fine-tuned to perform supervised learning by attaching a Softmax layer to the top layer.

There are many DL approaches and architectures, but most of the DNNs can be classified into five major categories:

1. Networks for unsupervised learning, designed to capture high-order correlation of data by capturing jointly statistical distributions with the associated classes - when available. Bayes rule can later be used to create a discriminative learning machine.
2. Networks for supervised learning. These networks are designed to provide maximum discriminative power in classification problem are trained only with labeled data - all the outputs should be tagged.
3. Hybrid or semi-supervised networks, where the objective is to classify data using the outputs of a generative (unsupervised) model. Normally, data is used to pre-train the network weights to speed up the learning process prior to the supervision stage.
4. Reinforcement learning - the agent interact and changes the environment and receives feedback only after a set of actions are completed. This type of learning is normally used in the field of robotics and games.
5. Generative Neural Networks - Deep generative models are a powerful approach to unsupervised and semi-supervised learning where the goal is to discover the hidden structure within data without relying on labels. Since they are generative, such models can form a rich imagery the world in which they are used: an imagination that can harnessed to explore variations in data, to reason about the structure and behavior of the world, and ultimately, for decision-making. A great advantage of these models is there is no need to supplement an external loss function as they learn the structure of the data autonomously.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Despite all the hype around deep learning, traditional models still play an important rule in solving machine learning problems, especially when the amount of data is not very large of the input features are relatively "clean". Also, if the number of variables is large compared with the number of training examples, SVMs or ensemble methods, like Random Forest or Extreme Gradient Boosting trees (XGBoost), may be simpler, faster and better options.

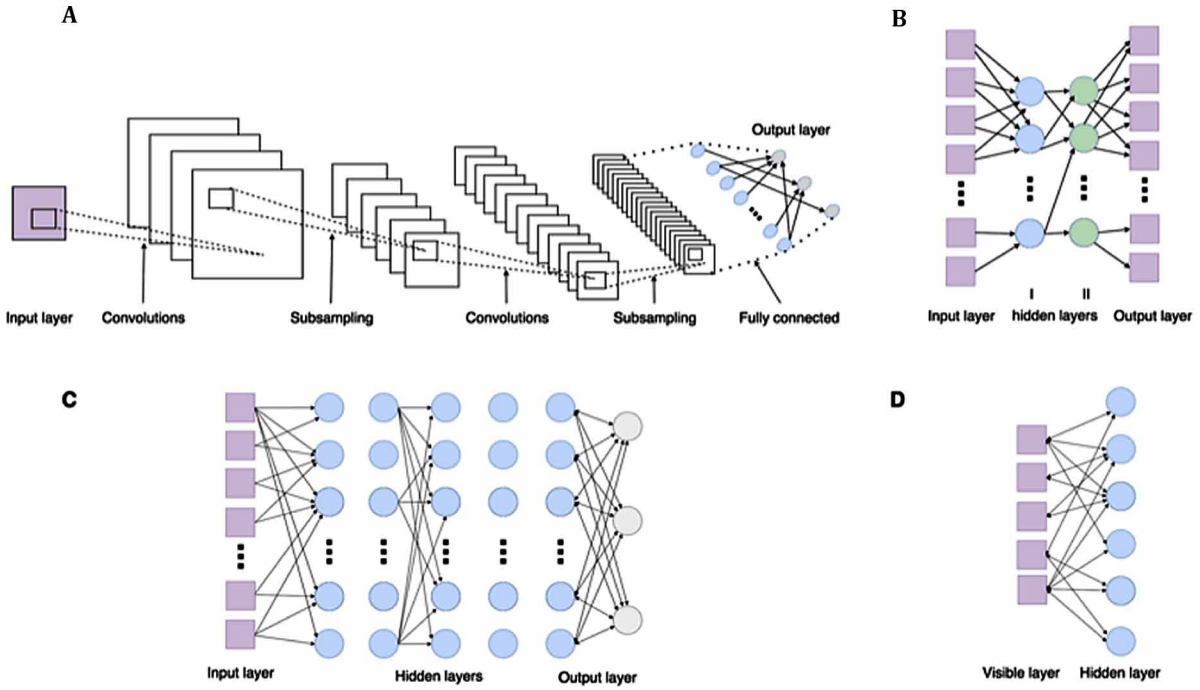
The most popular types of DNN architectures are: stacked denoising autoencoders (SdAE), Deep Belief Networks (DBN), Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN). Many advances in machine vision were achieved using CNNs - making this DNN type the standard for image processing. However, there are many flavors of DNNs that are applicable to the various business applications, depending on architecture, connectivity, initialization, training method and loss functions being used.

Figure 4 presents a brief summary of these popular DNN architectures. Following is a brief guideline to the terminology and the most popular types of deep neural networks.

2.1. Convolutional Neural Networks (CNN)

A CNN is composed by several blocks composed by different types of layers. Each module consists of a convolutional layer and a pooling layer. These modules are often stacked up with one on top of another, or with a softmax logistic layer on top of it, to form a deep model. The CNN uses several tricks that

Figure 4. Four of the most popular classes of deep learning architectures in biological data analysis: a) convolutional neural network (CNN) has several levels of convolutional and subsampling layers optionally followed by fully connected layers with deep architecture; b) stacked autoencoder consisting of multiple sparse autoencoders; c) deep belief network (DBN); d) restricted Boltzmann machine (RBM) architecture includes one visible layer and one layer of hidden units



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

makes them well suited for image processing like weight sharing, adaptive filters and pooling. Pooling takes subsamples of the convolutional layer to feed the next layer - acting as a powerful regularizer. Weight sharing and pooling schemes (most usually a max pooling), allow the CNN to generate conservation properties like translation invariance. CNNs are highly effective and have been commonly used in computer vision and image recognition.

CNNs operate on what should be considered a signal stream rather than a feature vector. That is, fully connected neural nets consist of activation units bound to all inputs of the feature vector. Every unit has a weight specific to each feature in the input. Convolutional layers, on the other hand, utilize weight sharing by sliding a small (trainable) filter of weights across the input vector (or 2D input map, as CNNs are often used on images) and convolving each overlayed region of input with the filter.

CNNs with max-pooling are powerful enough to mimic low-level stages of primates visual cortex having biologically plausible feature detectors, such as Gabor filters. However, once trained, the CNN acts a simple feedforward machine with frozen weights.

2.2. Deep Autoencoders

An autoencoder is a DNN having as outputs the input data input itself. If they are trained with some added noise, these architectures can act as generative models and are called denoising autoencoders. An autoencoder can be trained with a greedy layer-wise mode, much like the DBNs, to form a deep model.

Autoencoders can be stacked to form a deep network by forwarding the outputs of the autoencoder in the layer below as input to the layer above. The unsupervised pre-training is done a layer at a time and each layer is trained to minimizing the error in reconstructing of its input. After pre-trained, the network can be fine-tuned by adding a softmax layer and applying supervised backpropagation - as if they were multilayer perceptrons.

Stacked denoising AutoEncoders (SdAE) is a stochastic version of AE obtained by adding noise to the input in order to prevent learning of the identity map. They try to encode the input while undo the effect of a corruption capturing the statistical dependencies in the inputs.

Variational AutoEncoders (VAE) are composed by an autoencoder with extra constraints on the encoded representations being learned – it learns a latent stochastic model to generate its input data and a function to approximate sampling from latent variables – thus it's a generative model.

The parameters of the model are trained via two loss functions: a reconstruction loss forcing the decoded samples to match the initial inputs (just like a normal autoencoders), and the KL divergence between the learned latent distribution and the prior distribution, acting as a regularization term. You could actually get rid of this latter term entirely, although it does help in learning well-formed latent spaces and reducing overfitting to the training data. See this tutorial and some examples of code applied to the MNIST dataset on the Keras blog.

2.3. Deep Belief Networks

A Boltzmann Machine formed by two layers - one of visible units and other of hidden units with no lateral connections is called a Restricted Boltzmann Machine (RBM). A Deep Belief Network (DBN) is a probabilistic generative model formed by several stacked hidden layers (normally binary). The top two layers are symmetrically connected through indirect connections. Each layer receives a directed connection from the layer below. DBN are trained by greedy layerwise strategy using the Contrastive

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Divergence (CD) algorithm proposed by Hinton et al. Apart from providing good initialization points for the weights of the network, the DBN also have other interesting properties: 1) we can use all data, even unlabeled datasets; 2) it can be seen as a probabilistic generative model, very useful within the Bayesian framework; and 3) the over-fitting problem, can be effectively alleviated by the pre-training step and other strong regularizers, like dropout.

2.4. Recurrent Neural Networks

Traditional ML methods, like Support vector machines, logistic regression, and feedforward networks, have proved useful without explicitly modeling time in temporal process – by projecting time as space. This assumption, however, is incapable to model long-range dependencies and has limited usability in complex temporal patterns. Recurrent Neutral Networks (RNNs) are a rich family of models differentiable end-to-end, thus amenable to gradient-based training, latter regularized via standard techniques, like dropout or noise injection. Recurrence is key to solve hard problems, like language, as it seems to be present in most of the brain mechanisms.

The first structures of RNN were introduced by Jordan as feedforward networks with a single hidden layer extended with special units. Output node values were fed to the special units, which then feed these values to the hidden nodes at the following time step. If the output values are actions, the special units allow the network to remember actions taken at previous time steps. Additionally, the special units in a Jordan network are self-connected.

RNN are a class of unsupervised or supervised architectures to learn temporal, or sequential, patterns. A RNN can be used to predict the next data point in a sequence using the previous data samples. For instance in text, a sliding window over previous words is used to predict the next word or set of words in the sentence. RNNs are generally trained with Long Short Term Memory (LSTM) algorithm proposed by Schmidhuber (Schmidhuber 2015) or GRU (gated recurrent units). The flip side is that they are difficult to train in capturing long-term dependencies due to the well known gradient vanishing or gradient explosion problems and great care required in optimizing hyper parameters.

RNN have become recently very popular, especially with the introduction of several tricks, like bidirectional learning (forward and backward sequence prediction) and attentive mechanisms that allow the use of dynamic size sliding windows, especially useful to build language models.

2.5. Training DNNs

Almost all DL neural networks are trained evolving the optimizing a loss function, like likelihood maximization. The loss function may assume many forms depending on the problem. The likelihood of a set of parameter values, θ, given outcomes x, is the probability of those observed outcomes given those parameter values, $P(x|\theta)$. For a classification problem, it's typically defined as the log likelihood, given the observations of labels y, we want to generate a model with a probability distribution than maximizes the $\log(P(y|x))$. This is equivalent to minimizing the cross-entropy, a measure of similarity between two distributions P and Q: the sum of the entropy of P and the Kullback–Leibler divergence between Q and P (also known as the relative entropy of Q with respect to P):

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.5.1. Initialization, Overfitting, and Regularization

Deep Neural Networks are relatively large and complex machines - models with 100 million or even billions of parameters are not uncommon. Models with a large number of free parameters can describe very complex relations within data but extreme caution should be taken that they are able to generalize properly. Traditional ways to regularize DNN include cross validation, noise addition, L2 regularization and, more recently, dropout, i.e., removal of a randomly subset of connections - a very simple but effective technique.

Learning, i.e., adjusting the weights of the neuron connections, is normally performed with backpropagation using Stochastic Gradient Descendent (SGD) in batch mode, with a very small (and decaying) learning rate and a some momentum term. Getting a neural network learning properly is hard. A proper initialization of weights is important to guarantee convergence – normally we should start with a set of small random values in the interval [-1/a, 1/a] where *a* is the number of connections to the neuron.

The learning rate should be small - depending of the network architecture, complexity of the problem and size of the dataset – the higher the lowest the learning rate. A value between 0.01 and 0.001 is normally a safe choice. The number of epochs is typically in the order of thousands and the batch size not less than 100 – again highly dependent of the problem.

2.6. DL as a Service

2015 may well be considered the year of Deep Learning. Remarkable innovations were achieved, like Microsoft neural net of 152 layers (where 6 or 7 layers was the norm) that lead to the winning of the ImageNet competition - with superhuman accuracy.

At the same time all big players are creating their own DL platforms and open sourcing some of their core algorithms – the era of AI-as-a-service. Amazon, IBM, Google, Facebook, Twitter, Baidu, Yahoo, and Microsoft. Google is offering TensorFlow for free. Next table presents a summary of the principal services offered by these companies.

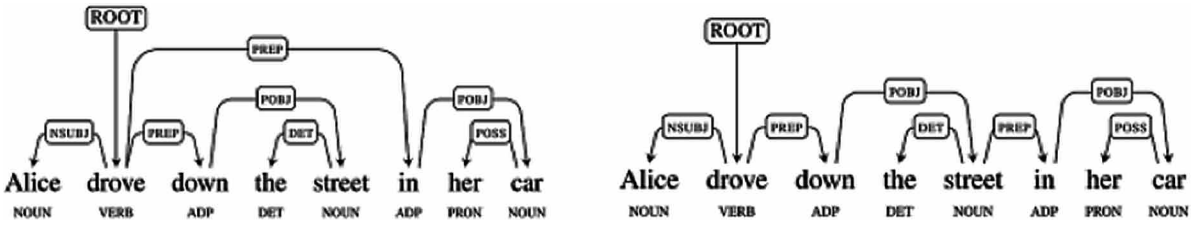
Deep Learning is moving to open source and the Cloud. Google, Facebook, IBM, Amazon and Microsoft are trying to establish ecosystems around AI services provided in the cloud. “This technology will be applied in pretty much every industry out there that has any kind of data—anything from genes to images to language”

Table 1. Main machine learning platforms

Company	Cloud-Based ML Platform	DL Technology (Open Sourced)
Amazon	Amazon Machine Learning	DSSTNE
Baidu	-	Deep Speech 2 PaddlePaddle
Facebook	-	TorchNet, FastText
Google	NEXT Cloud	TensorFlow
IBM	Watson	IBM system
Microsoft	Azur	CNTK
Twitter	-	Cortex

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Parsing a sentence: two possibilities



3. APPLICATIONS TO NATURAL LANGUAGE PROCESSING

Deep Learning is having a tremendous impact in Natural Language Processing (NLP). Language understanding is one of the oldest, and probably the hardest problem for machines. However, as large corpus of data become available on the Internet, DL is a natural option to solve innumerous problems related to understanding human language.

Parsing is a central problem for NLP. It consists in decomposing a sentence in its components (nouns, verbs, adverbs, etc.) and built the syntactic relation between them – the parsing tree. It is a very complex problem due the ambiguity in possible decompositions. For instance the sentence "Alice drove down the street in her car" has at least two possible dependency parses:

The first corresponds to the (correct) interpretation where Alice is driving in her car; the second corresponds to the (absurd, but possible) interpretation where the street is located in her car. The way humans disambiguate these options is through "common sense" – they know that streets cannot be located in cars. For machines incorporating this world information is very challenging.

Google launched a recent tool SyntaxNet (code based on Tensorflow available at Github SyntaxNet) to solve the hard parsing problem – a 20 to 30 words sentence can have thousands of syntactic structures (see the blog post on Google research Google parser). They used a globally normalized transition-based neural network model that achieves state-of-the-art part-of speech tagging, dependency parsing and sentence compression. The model is a simple feed-forward neural network that operates on a task-specific transition system, yet achieves comparable or better accuracies than recurrent models. The SyntaxNet English language parser, Parsey McParseface, is considered the best parser surpassing, in some cases, human level accuracy. Recently they expand the service to cover about 40 languages.

3.1. Distributed Representations

One of the core problems in NLP is related to the high-dimensionality of data, thus a huge search space) and inference of grammatical rules. Hinton (D. E. Rumelhart 1986) was one of the first to propose the idea that words could be represented via distributed (dense) representations. The advantage of distributed representations is that semantics can easily be accessible and knowledge be transferred from different domains - even on different languages.

Learning a distributed (vectorized) representation for each word is called word embedding. Word2vec is the most popular approach to create a distributed representation of words; It's a publicly available library providing an efficient implementation of skip-gram based vector representations for words. The model and implementation is based on the work of Mikolov et al. (Mikolov & Sutskever 2013). In addi-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

tion to the implementation, the authors also provide vector representations of words and phrases learnt by training this model on Google News Dataset (about 100 billion words). Vectors can be up to 1000-dimensional containing 3 million words and phrases. An interesting feature of these vector representations is that they capture linear regularities in the language. For example the result of the vectorized words equation: "Madrid" - "Spain" + "France" is "Paris".

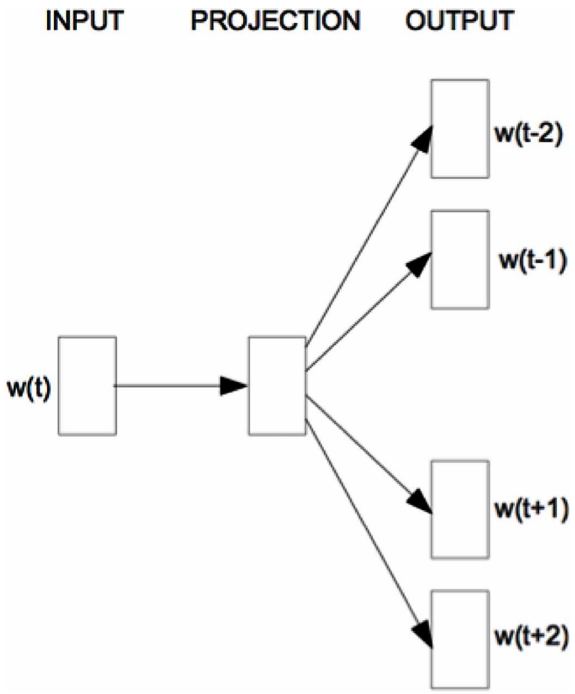
After the "bag of words" (BOW) - with the TFIDF trick - Word2vec is probably the most used method for NLP problems since it's relatively easy to implement and very useful in understanding (in an unsupervised way) hidden relations in the data. There is a very good and well-documented Python implementation of word2vec, Gensim. It can be used with pre-trained vectors or trained to learn the embeddings from scratch given a large training corpus - normally on the order of million of documents.

Le et al. also proposed a method to encode whole (variable size) paragraphs using a similar technique; they called it Paragraph Vector. The paragraph vector is shared across all contexts generated from the same paragraph but not across paragraphs. Each word vector is shared across paragraphs. Kiros et al. (R. Kiros 2015) introduced the idea of Skip-Though Vectors using unsupervised learning to encode sentences. The model used a RNN to reconstruct the surrounding sentences of an encoded passage.

3.2. Language Translation

Natural Language Translation is a very hard problem that has been defeating a satisfactory solution since the eve of AI in the 50's. Traditional DNNs have some limitations in dealing with this problem, like the requirement that the inputs and targets should be encoded with vectors of fixed dimensionality.

Figure 6. The schematic representation of word2vec algorithm—word W(t) is used to predict context words W(t-2)…W(t+2)—here a context window of K=5 is considered



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

For sequences of arbitrary length this is a serious limitation[2]. Unlike the traditional statistical machine translation, recent DNN models create a single neural network to jointly represent the distributions of both languages and maximize the translation score. Most models use a scheme of encoder–decoders and encode a source sentence into a fixed-length vector from which a decoder generates the respective translation. The most common model consists of RNN with LSTMs unities that process input sequences and compress it into a large, but fixed, dimensional vector. This vector is later used by another LSTM to extract the output sequence. The second LSTM is essentially a recurrent neural network language model except that it is conditioned on the input sequence.

Sutskever et al. (I. Sutskever 2014) used this approach to achieve state-of-the-art performance of the conventional phrase-based machine translation system on an English-to-French translation task - a BLEU score of 34.81, outperforming the best previous neural network NLP systems, and matching the best published results for non-neural network approaches. When their system is used to re-rank candidate translations from another system, it achieved a BLEU score of 36.5. The implementation involved eight GPUS and training took 10 days to complete.

3.3. Voice Recognition

Automatic Speech Recognition (ASR) is the problem of translating voice transcripts into pioneer the application of Deep Bidirectional LSTM to this problem achieving top results in the TIMIT database. Graves et al. applied applied Deep Bidirectional LSTM to this problem achieving top results in the TIMIT database. They used end-to-end designed for discriminative sequence transcription with recurrent neural networks, namely Connectionist Temporal Classification and Sequence Transduction. These methods do not require forced alignments to pre-segment the acoustic data, as they directly optimize the probability of the target sequence conditioned on the input sequence and are able to learn an implicit language model from the acoustic training data.

A Baidu team recently proposed an advanced Automatic Speech Recognition (ASR) system for translating voice into text (Amodei 2015). The improvement performance of the algorithm is due to deep learning replacing feature extraction modules with a single neural model. The Baidu system, called *Deep Speech 2*, consists of a pipeline that approaches the accuracy of humans in several languages. This system was built on end-to-end deep learning using bidirectional RNN trained in clean and noise environments. In English the speech system was trained on 11,940 hours of speech, and in Mandarin for 9,400 hours. Data synthesis was used to augment the data during training. Training a single model at these scales requires tens of exaFLOPs that would require 3-6 weeks to execute on a single GPU. A recent experiment showed that users, using this system on smartphones, input text faster by speaking then typing.

4. APPLICATIONS TO IMAGE AND VIDEO PROCESSING

Probably the area where most advances were made by DL was in image processing. The human visual perceptual system achieves remarkable object recognition performance even in noisy environments or under geometric transformations or background variation. Recent advances in DNN, particularly using Convolutional Neural Networks, lead to a revolution in image processing and analysis achieving, or even surpassing, human level performance. In this section we will show some applications of deep learning into image processing.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Performance of deep speech 2 compared with humans

Read Speech			
Test Set	**DS1**	**DS2**	**Human**
WSJ eval '92	4.94	3.60	5.03
WSJ eval '93	6.94	4.98	8.08
LibriSpeech test-clean	7.89	5.33	5.83
LibriSpeech test-other	21.74	13.25	12.69

4.1. From ImageNet to Caption Generation

CNNs were one of the first deep learning models - biologically inspired by the mammal visual cortex. In a CNN each unit takes its input from a local receptive field on the layer below forcing it to extract local features. Furthermore units within a plane or feature map are constrained to share a single set of weights, this makes the operations performed by a feature map to be shift invariant. The weight sharing technique also reduces the number of free parameters, thus reducing risk of overfiting. The features of CNNs can be summarized as:

- Convolutional neural network with a sequence of convolution and pooling layers.
- Use convolution to extract spatial invariant features.
- Subsample using spatial average of maps.
- Multilayer neural network (MLP) as a final classifier.
- Sparse connection matrix between layers (weigh sharing) to avoid large computational cost and reduce overfitting.

Complete CNNs are formed by stacking together multiple convolutional layers (each with feature-maps planes and local receptive fields). Sub-sampling layers are also added improving invariance to shift and distortions. During 1990's it become evident that deeper networks perform better. However, larger networks required lots of computational resources and overfit easily since not much data was available.

Krizhevsky et al. (A. Krizhevsky 2012, R. Kiros 2015, Amodei 2015) pioneered the use of a graphical processor units (GPUs) for a fast implementation of a CNN containing 650,000 neurons and 60 million parameters – by contrast LeNet-5 had 60,000 - winning with a top-5 error rate of 15.3% - a far better result than the state of the art technique techniques that reached only 26.2%. Besides using larger data set and bigger networks, these authors used aggressive regularizations techniques to avoid overfitting, like data augmentation (applying slight distortions in shapes, rotations and colours) and dropout to shrink co-adaptions of neurons. This last technique allows a single neuron to learn more robust features without relying on other neighbour neurons.

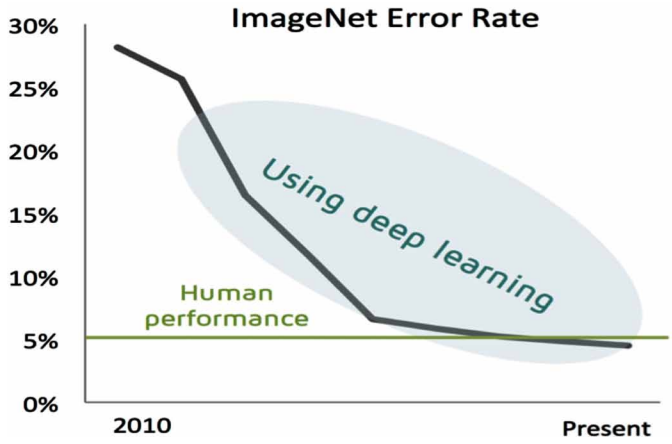
A trained CNN can be used for classification on the ImageNet data set replacing the final classification layer with a regression network. This regression network is simply an MLP with two hidden layers of 4,096 and 1,024 units, connected finally to the output layer of 4 units which predicts the coordinates of the bounding boxes. The final layer is class-specific having 1,000 versions (one for each class) while the rest of the regression network shares weights. During localization only the regression network's weights are updated, and the remaining larger network only acts as a feature extractor.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Image captioning with CNN



Figure 8. Performance evolution on ImagNet competition



In 2014, Google introduced Inception5 (Google LeNet), a very deep CNN (with 20 layers) that won the ImageNet contest with an error rate of only 6.7% - this work showed the importance of using very deep models to abstract higher level features from the images.

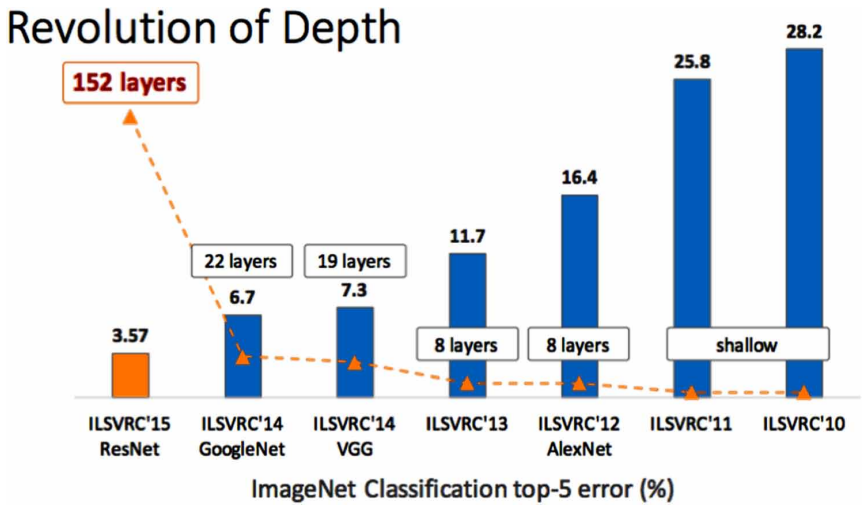
Recently, a Microsoft team achieved a super-human performance on ImageNet with an error rate of only 3.7% with a network named ResNet (residuals network) (K. He 2015) as well as MS COCO data set – the code is available on github. ResNet is based on a simple idea: feed the output of two successive convolutional layer bypassing the input to the next layers - bypassing a single layer did not give much improvements while two layers can be seen as a classifier itself. They were able to train networks of up to 1000 layers.

4.2. Video and Satellite Images

Donahue et al. investigated a CNN model with RNN and proposing a recurrent convolutional architecture for end-to-end large-scale visual challenging task, like activity recognition, image captioning, and video description. This model departs from fixed visual representations and was able to learn compositional

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Deep residual learning for image recognition (from He, Zhang, Ren, & Sun, 2016)



representations in space and time. The model is a fully differentiable RNN, capable to learn long-term dependencies, and very appealing since it can map variable-length videos to natural language text. The model is fully trained with backpropagation. The authors show that this model can achieve good results for discriminative or generative text generation tasks. They evaluate the model on TACoS multilevel dataset containing 44,762 video/sentence pairs obtaining a BLUE score of 28.8.

In Mamoshina (2016), the authors combined a CNN and a LSTM for jointly learn the embedding of video and text in order to generate automatic annotation of videos. Due to lack of datasets, the authors relied on photo annotation data and used knowledge transfer techniques. They achieved good accuracy in the subject, verb, and object (SVO) metrics but still far from human level – probably due to the lack of training data.

Satellite image classification is a complex problem involving remote sensing, computer vision, and machine learning. The problem is challenging due to the high variability of the data. Basu et al (S. Basu 2015) proposed a method based on Deep Belief Networks and careful preprocessing of satellite image achieving 97.95% accuracy in two public datasets. One dataset consists of 500,000 image patches covering four broad land cover classes: barren land, trees, grassland and other; 400,000 patches were chosen for training and the remaining 100,000 for testing.

Cargometrics is a startup that uses VHF radio tracking as well satellite images processing through deep learning algorithms to analyze maritime traffic data in order to help predict commodity prices – it tracks the movement of 120 000 ships across the world. Their work is being used by hedge funds to identify pricing and securities opportunities.

Terrapattern uses DL to perform similarity based search for unlabeled satellite photos. It provides an open-ended interface for visual query-by-example, the user clicks in a spot on Terrapattern's map and it will find other locations that look similar.

Vicarious developed deep learning algorithms for vision, language, and motor control. They are mainly focused on visual perception problems, like recognition, segmentation, and scene parsing. They claim that their system requires orders of magnitude less training data than traditional machine learning technique in deploying generative probabilistic models.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

5. OTHER APPLICATIONS

5.1. Self-Driven Cars

Deep Learning plays a consider role in Self-driving cars technology by analyzing a disparate set of signals, being video the most challenging. Following the recent success of Google driverless car, almost all carmakers are considering this option in future versions of they cars. Some models under test are Toyota Prius, Audi TT, and Lexus RX450h. Tesla S3 will probably be the first production car with self-driving capabilities included by default. All of these models rely on Deep Learning technology for object recognition, planning, routing and object avoidance.

Google has developed its own custom vehicle relying on a 64-beam laser detector that allows the vehicle to generate a detailed 3D map of its environment. The algorithm use these maps and combines them with high-resolution maps of the world, producing a sufficiently detailed model for self-navigation. Google tested their fleet of vehicles over 2 million km – Tesla tested their own vehicles on more than 130 million km.

Nvidia launched a GPU specific for these cars and Baidu plans to start self-driving cars tests on roads in the United States over 2016 - the core technology is based on DL algorithms. Baidu and Google are pushing towards governmental regulation of self-driven cars claiming it only requires minor changes in actual infrastructure. The goal of Baidu is to run a shuttle service in Chinese cities by 2018.

Autonomous driving requires intuitive psychology. The self-driving car needs to have some common sense understanding or inferring pedestrian behavior and beliefs (do they think it is safe to cross the street? Are they paying attention?) as well as desires (where do they want to go? Are they in a rush?). Similarly, other drivers on the road have similarly complex mental states underlying their behavior (do they want to change lanes or ass another car?). This type of psychological reasoning, along with other types of model-based causal and physical reasoning, are likely to be especially valuable in challenging and novel driving circumstances for which there is little relevant training data.

A recent study concluded that self-driving taxis could reduce the traffic in cities as much as 90%.

5.2. Chatbots

Chatbots, also called Conversational Agents or Dialog Systems, are algorithms designed to have human level conversation capabilities. The goal of bots is to achieve a level of natural conversations indistinguishable from humans. There are two types of bots, Retrieval-based, that use a repository of predefined responses and some heuristics to pick an appropriate response based on the input and context and Generative models that generate automatically responses from past experience and context. Most of the latter rely on Deep Learning technology.

Generative models can be quite sophisticated and learn from creating latent representation of data. They are very flexible but require a large corpus of text (training data). Sequence-to-Sequence learning methods like (I. Sutskever 2014) has a great potential, but most production systems are still based on hard rules for dialogue retrieval. Short text conversations are easier to optimize, especially in closed domain knowledge where the space of possible inputs and outputs is somewhat limited to a specific context.

Vinyals et al. used the sequence-to-sequence framework to create a conversational model able to predict the next sentence given previous sentences in a conversation (O. Vinyals 2015). It was designed as an end-to-end model requiring few crafted features. They trained it in a large conversational dataset and it

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

was competent enough to deliver good advice using an IT helpdesk dataset while showing common sense reasoning in a movie conversational dataset. However, the bot fails to be consistent in the conversations.

Generative models are very powerful, but grammatical mistakes can be costly, so companies still rely on old retrieval technologies. However, as companies get more data, generative models will become the norm - but probably with some human supervision to prevent them from "inappropriate behavior", like happened to the Microsoft Twitter chatbot, Tay

Most big companies are using, testing or considering the implementation of chatbots in their services and operations. Using its experience with its personal assistant Cortana, Microsoft has recently opened a development framework for Chatbots implementation and also released Luis.ai an API for language understanding. Facebook has launched M, a chatbot for its message service and acquired Wit.ai, a company that works on voice recognition technology. IBM has also available a very simple API to embed its powerful knowledge inference machine, Watson, into a conversational bot. Most of these services can be easily incorporated in conversational services like Twitter, Whatsapp, Skype, WeChat, Telegraf or Slack. For instance, Slack allow simple or complex conversation automating based on hard or soft rules.

Retail banks and FinTech start-ups are now exploring the use of Chatbots for digital experiences for checking bank account balances, finding nearby ATMs, making payments and even advising how to spend your money more wisely. Tractica predicts US$ 40.6bn will be spent on artificial intelligence systems for enterprises from 2015 to 2014. Zendesk launched a chatbot to automate answers to customer queries, after other competitors, like Slack, start automate some conversations. Deep learning for chatbots on Wildml, offfers an excellent tutorial on how to build bots from scratch using Ubuntu forum data.

In a recent post (Hybrid Intelligence) Clarie Corthell stated that: "humans will always be part of the system. So we shouldn't focus on how to replace humans in a system, rather how to optimally integrating the contributions of humans and computers." I partially agree: humans will definetly have to be in the loop for the next years, but as technology evolves we have to change.

5.3. Art

Perhaps the most extraordinary achievements of DL were obtained in an unsuspicious area: art. Gatys et al. applied Convolutional Neural Networks to obtain a representation of the style of an artist (input image) using the feature space originally designed to capture texture information. By including the feature correlations of multiple layers, they obtain a stationary, multi-scale representation of the input image. They proved that the representations of content and style in the Convolutional Neural Network are separable. Both representations can be independently manipulated to produce new, perceptually meaningful images. To demonstrate this finding, they generate images that mix the content and style representation from two different source images – see Figure 10 (see a demo online on Ostagram).

A Swedish company, Peltarian, released a neural network that can perform sophisticated computer generated choreography by extracting high-level features from raw sensor data. The system, called chor-rnn, uses Recurrent Neural Networks for generating novel choreographic material in the nuanced choreographic language and style of an individual choreographer. It also can create higher-level compositional cohesion, rather than just generating sequences of movement. The neural network is trained on raw motion capture data and that can generate new dance sequences for a solo dancer.

In literature, Hitoshi Matsubara, used a DL based algorithm to generate a short story that was shortlisted, among 10 others, from more than 1000 human written pieces.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. Artificial generated images using CNNs trained on two sets of images



In music, Google Brain's creative AI launched a project Magenta dedicated to the creation of music and art through machine learning. They released the first music track - which show the potential ANN to generate creative music.

5.4. Biomedical

In the era of Big Data, modern biology and medicine are confronted with a massive scale of data that requires novel computational tools for analysis and translation of the results. Success of the analysis depends on the ability to interpret such complex and biased data.

5.4.1. Biomarkers and Genomics

One of the main directions of biomedicine is the transition from biological data to valid biomarkers that reflect phenotype and physical state of the body. Biomarkers are critical in measuring clinical trials outcomes, detecting diseases, and monitoring disease states, particularly for diseases as heterogeneous as cancer (see Mamoshina, 2016, for a review).

Exponential growth in modern genomics technology, particularly with the introduction of Next Generation Sequencing (NGS), allows production of a massive amount of genomics data by reducing cost and time consumption of sequencing. Analysis of sequencing data implies analysis of the function and structure of genome. All of this could be performed *in silico* using modern computational approaches.

Analysis of transcriptomics includes analysis of various types of transcripts (mRNA, IncRNA, miRNA, etc.) that carry various types of functional information, from splicing code to biomarkers of various diseases. Transcriptomics data are often obtained from different types of platforms (various microarray platforms, sequencing platforms) that differ by the gene set measured and method of signal detection. Cross-platform analysis requires normalization techniques, which can be a major challenge[22]. DNNs are particularly well-suited for cross-platform analysis with their generalization ability and as such have great potential to benefit transcriptomics data analysis.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Some startups working on these areas are:

- Stratified medical.
- Atomwise.
- Quantitative Medicine.
- Insillico medicine.
- Deep Genomics.

5.4.2. Healthcare

Deep Learning in already having a strong impact in healthcare industry due to: i) increase capacity and accuracy of learning algorithms and ii) wide availability of large sets of healthcare data – made possible through the digitalization of (structured and unstructured) medical records, as well as personal genetic data and other personalized data originated from mobile devices.

Having achieved human level performance on object identification or facial recognition, DANN have a great potential in processing medical imaging, an area where subjectivity interpretation is common and contextualization is key to disambiguate several possible explanations.

DL is being applied to images—which remains its primary application today. Some companies—like Enlitic—are applying DL to recognize cancer in medical images, like X-rays and many automated image recognition tools are already used in hospitals. Diagnoses based on processing of medical images is, however, just a tiny fraction of the potential of DL in medical sector.

Some startups have emerged to apply deep learning to medical imaging such as BayLabs, Imagia, MD.ai, AvalonAI, Behold.ai or Kheiron Medica. iCarbonX has a near term goal of predicting the onset of diseases from genomic, medical and lifestyle data.

DL was also extensively used for model and understanding complex biomedical processes like protein interactions, RNA decoding.

6. CONCLUSION AND FUTURE DIRECTIONS

We are experiencing a revolution in Artificial Intelligence. This fast-moving wave of technology innovations powered by AI is impacting business and society in such a velocity that is hard to predict its implications. One thing is sure: cognitive computing powered by AI will be able to perform most tasks and empower humans even in creative tasks. AI will impact jobs that had seemed impossible to automate a few years ago, from doctors to legal clerks.

By 2020, the market for machine learning applications will reach $40 billion, IDC, a market research firm, estimates - 60 percent of those applications, the firm predicts, will run on the platform software of four companies (Amazon, Google, IBM and Microsoft) most of it relying on Artificial Intelligence.

Probably the most immediate impact of GAI supported by DNN will not be in self-driving cars or robotics but in customer service. Services like sending a specific email, a mobile push, or a customer pass for a specific shop or event; predictive analytics to help support decisions and call centers. Contact centers deal with very mundane interactions that soon will be serviced through automated messaging like chat bots and personal assistants. AI can help suggest how to deliver a conversation; user interests and product. It can even use the data for secondary proposes, like risk assessment based on previous interactions.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

6.1. Research

From the point of view of technology, I would like to highlight three key research areas I found the most promising:

- Deep Generative Adversarial Neural Networks (GANs).
- Recurrent Neural Networks.
- Multimodal learning models.

GANs are a powerful approach to explore unsupervised and semi-supervised learning where the goal is to discover the hidden structure within data without relying on external labels. One of the most interesting models is Generative Adversial Networks (GANs) where a generative networks is trying to confuse a discriminative one in distinguish real data from faked (generated) one (Alexey Kurakin 2016)14]. The Deep generative models have widespread applications including probability density estimation, image denoising and inpainting, data compression, scene understanding, representation learning, 3D scene construction, semi-supervised classification, and hierarchical control.

New methods to train Recurrent Neural Networks and Multimodal Learning (learning from different data types – video, image, voice, text) are presently very active areas of research and very promising in solving problems related with control and object manipulation, image and video caption and text processing.

6.2. Risks

DL is a powering this tremendous transformative AI technology. A study by Carl Benedikt Frey and Michael Osborne, published in 2013, found that 47% of jobs in America were at risk of being replaced in a near term. More recently Bank of America Merrill Lynch predicted that by 2025 the "annual creative disruption impact" from AI could amount to $14 trillion, including a $9 trillion reduction in employment costs thanks to AI-enabled automation of knowledge work; The McKinsey Global Institute says AI is contributing to a transformation of society "happening ten times faster and at 300 times the scale, or roughly 3,000 times the impact" of the Industrial Revolution.

6.3. Startups

In 2015 only 1% of software have AI capabilities. In 2020 that number will be 50%. DL thus represents a huge opportunity for Startups and investors. In a recent review by The Economist, Nathan Benaich stated that "In 2015 a record $8.5 billion was spent on AI companies, nearly four times as much as in 2010, according to Quid, a data-analysis company. The number of investment rounds in AI companies in 2015 was 16% up on the year before, when for the technology sector as a whole it declined by 3%."

Accel, a venture capital firm headquartered in Palo Alto, announced a $500 million fund for early stage companies and $1.5billion fund for late stage businesses. Of their nine areas of focus for the future, Artificial Intelligence topped the list of priorities. New Enterprise Associates, also a worldwide venture capital firm, is the second top investor in Artificial Intelligence. The firm has invested in WellTok's CafeWell.com website which combines social networking technology with knowledge of healthcare, making recommendations based on health goals allowing users to take control of their wellbeing.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Most AI based venture capital deals were in deep learning/machine learning. Naveen Rao, founder and Chief Executive of Nervana Systems explains, "the method allows computers to process tasks through 'neural networks', systems that mimic the structure of the human brain in order to learn, improve and, in the case of AlphaGo, become capable of decision-making that resembles human intuition more than it does brute computation."

REFERENCES

Amodei, D. (2015). Deep Speech 2: End-to-End Speech Recognition in English and Mandarin. *CoRR*.

Basu, Ganguly, Mukhopadhyay, DiBiano, Karki, & Nemani. (2015). DeepSat – A Learning framework for Satellite Imagery. *CoRR*.

Brenden, M. L., Salakhutdinov, R., & Tenenbaum, J. B. (2015). Human-level concept learning through probabilistic program induction. *Science*, 1332. PMID:26659050

Domingos, P. (2012). A few useful things to know about machine learning. *Communications of the ACM*, 55.

Goodfellow, I., Bengio, Y., & Courville, A. (2006). *Deep Learning*. Boston: MIT Press.

He, Zhang, Ren, & Sun. (2015). Deep Residual Learning for Image Recognition. *CoRR*.

Hinton & Salakhutdinov. (2006). Reducing the dimensionality of data with neural networks. *Science*, *313*(5786).

Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, *18*(7), 1527–1554. doi:10.1162/neco.2006.18.7.1527 PMID:16764513

Huang, Yu, Liu, Sedra, & Weinberger. (2016). Deep Networks with Stochastic Depth. *CoRR*.

Kiros, Zhu, Salakhutdinov, Zemel, Torralba, Urtasun, & Fidler. (2015). Skip-Thought Vectors. *CoRR*.

Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). Imagenet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems. Urran Associates, Inc.

Kurakin, Goodfellow, & Bengio. (2016). Adversarial Examples in the Physical World. *CoRR*.

Long, M., Cao, Y., Wang, J., & Jordan, M. I. (2015). *Learning Transferable Features with Deep Adaptation Networks*. ICML.

Mamoshina, P., Vieira, A., Putin, E., & Zhavoronkov, A. (2016). Applications of deep learning in biomedicine. *Molecular Pharmaceutics*, *13*(5), 1445–1454. doi:10.1021/acs.molpharmaceut.5b00982 PMID:27007977

Mikolov & Le. (2014). Distributed Representations of Sentences and Documents. *CoRR*.

Mikolov, T., Chen, K., Corrado, G. S., Dean, J., & Sutskever, I. (2013). Distributed representations of words and phrases and their compositionality. *Proceedings of NIPS*, 3111.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ribeiro, B., & Lopes, N. (2011). Deep belief networks for financial prediction. *International Conference on Neural Information Processing.* Springer. 10.1007/978-3-642-24965-5_86

Rumelhart, D. E., Smolensky, P., McClelland, J. L., & Hinton, G. (1986). Parallel distributed models of schemata and sequential thought processes. In *Psychological and Biological Models* (pp. 5–57). Cambridge, MA: MIT Press.

Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, *61*, 85–117. doi:10.1016/j.neunet.2014.09.003 PMID:25462637

Srivastava, R. K., Greff, K., & Schmidhuber, J. (2015). *Highway network.* arXiv preprint arXiv:1505.00387

Sutskever, I., Vinyals, O., & Le, Q. (2014). Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems*.

Venugopalan, Xu, Donahue, Rohrbach, Mooney, & Saenko. (2015). Translating Videos to Natural Language Using Deep Recurrent Neural Networks. *CoRR*.

Vieira, A., & Barradas, N. (2003). A training algorithm for classification of high-dimensional data. *Neurocomputing*, *50*, 461–472. doi:10.1016/S0925-2312(02)00635-5

Vinyals & Le. (2015). A Neural Conversational Model. *CoRR*.

ENDNOTES

1 Regarding these points, note that DL is a very active area of research and many of these difficulties are being addressed – some of them are partially solved while others probably never will be.

2 The quality of translation is measured by BLUE - it is the geometric mean of the n-gram precisions for all values of n between 1 and some upper limit – typically 4.

This research was previously published in Ubiquitous Machine Learning and Its Applications edited by Pradeep Kumar and Arvind Tiwari; pages 39-67; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Section 3
Tools and Technologies

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 24
Agreement Technologies for Conflict Resolution

Vicente Julian
Universidad Politécnica de Valencia, Spain

Stella Heras
Universidad Politécnica de Valencia, Spain

Victor Sanchez-Anguix
Coventry University, UK

Carlos Carrascosa
Universidad Politécnica de Valencia, Spain

ABSTRACT

Recently, artificial intelligence, has emerged as a new source of scientific works in conflict resolution. The interest in conflict resolution lies in diverse reasons. One of the main reasons is that computational systems have gradually shifted towards a distributed paradigm where heterogeneous entities should include computational conflict resolution mechanisms, such as proposed by agreement technologies. This chapter gives an overview of these technologies, which are needed in order to ensure the accomplishment of the global system goal and to solve conflicts. Among agreement technologies, automated negotiation is proposed as one key mechanism in conflict resolution due to its analogous use in human conflict resolution. Automated negotiation consists of an automated exchange of proposals carried out by software agents on behalf of their users. Another key technology is argumentation, which provides a fruitful means of dealing with conflicts and knowledge inconsistencies. Agents can reach agreements by engaging in argumentation dialogues with their opponents in a discussion.

INTRODUCTION

Agreement Technologies (AT) is a new discipline, which covers a range of specific techniques for dealing with interactions in dynamic, open environments (Ossowsky, 2013; Sierra, Botti & Ossowski, 2011). They address issues such as finding ways to negotiate, agree and cooperate with other agents, and developing appropriate means of forming and managing coalitions. These aspects are extremely related with the concept of conflict resolution. Conflict is an omnipresent phenomenon in human society, which is related with the necessity to achieve an agreement between entities. Conflict spans from individual

DOI: 10.4018/978-1-7998-0951-7.ch024

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

decision making trade-offs such as deciding what to do next (sleep, eat, work, play), to complex scenarios including politics and business. Over the last decades, computer science has tried to study conflict from a computational perspective. Conflict is a ubiquitous phenomenon and arises in many areas of our lives. It arises and needs to be dealt with in social settings, such as a group of friends deciding on a vacation or contract negotiation in business and politics (Browder, 2000), as well as in individual settings related to action selection (e.g., how to weigh one's preferences and decide what product to buy). Even if we are not aware of it, we are continuously facing conflict and attempting to solve it.

Different proposals based on agreement technologies have been developed with conflict resolution as their main research goal. For instance, automated negotiation approaches have been proposed as set of algorithms and protocols whose mission is providing effective deals in electronic marketplaces (Jennings et al, 2001; Parsons, Rodriguez-Aguilar & Klein, 2011; Sanchez-Anguix, Julian, Botti, & García-Fornes, 2013). Researchers in argumentation aim to solve conflicts by means of dialogue games, speech particles and information rebuttal (Rahwan et al, 2003). Works in cooperation and coordination try to give a joint solution to a problem, usually with no global information or control (Rebollo, Carrascosa, & Palomares, 2014). Reputation is another way to solve conflicts when an entity must to decide about other entities (Pinyol, Sabater-Mir, 2013). Finally, agent organizations, inspired by the metaphor of human organizations, allow achieving global goals (and solve conflicts) by using norms, plans and structures formally specified (Argente et al 2011; Hübner, Sichman & Boissier).

This chapter explores in detail several works done in the area of agreement technologies, which can be used as a way to engage new conflict resolution approach along with new challenges to face by these technologies. During the chapter we discuss different types of agreement and agreement processes related to conflict resolution, describing and relating the different technologies and application areas involved. First, the different related technologies are briefly introduced. After this, a detailed analysis of negotiation and argumentation technologies is introduced in the following sections.

AGREEMENT TECHNOLOGIES

Open multi-agent systems are distributed systems where heterogeneous agents, with their own goals, can enter and leave the system during the life of the system (Hewitt, 1991). For instance, we can think of an electronic commerce platform as an open system where users, human or even automated software, act according to their own interests: in the case of sellers to maximize their own profits, and in the case of buyers to acquire some goods at relatively good price. Since agents have different goals, act based on their goals, and they can be heterogeneous (i.e., humans and software agents may show different behaviors), it is feasible to find situations where an agent's goals conflict with other agents' goals. If we refer ourselves to the example of electronic commerce, the buyer may want to buy the product at a low price, while the seller may want to maximize its revenue. In these situations, mechanisms that allow agents to coordinate, regulate their behavior, and solve conflicts are needed.

Electronic commerce is not the only application where conflict may make act of presence. For instance, in the last few years, grid computing (Foster, Kesselman & Tuecke, 2001) has emerged as a new paradigm of computation where different entities collaborate to accomplish several tasks. In grid computing, entities share several resources: from hardware resources (e.g., computing nodes) to software resources (e.g., services). How should those resources be distributed among the different tasks or users of the system? Presumably, users want the best response time for their tasks, and resource owners want

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

to take the highest profit of their resources. Resource allocation is a delicate matter, especially when collaboration requires cross-boundary relationships. Software mechanisms that solve conflict in these scenarios are needed.

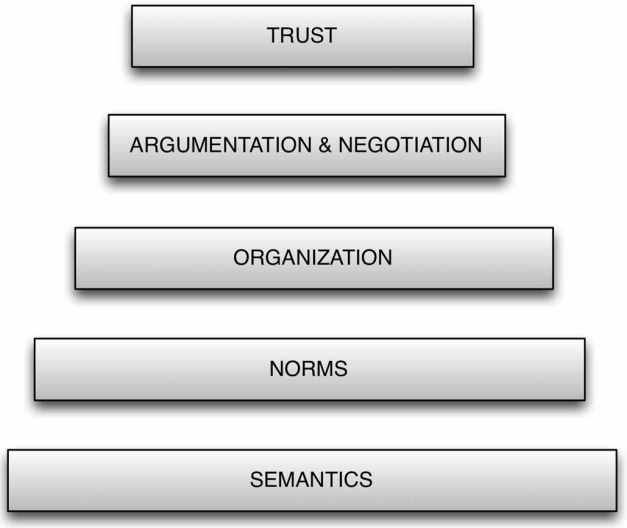
Even purely cooperative applications like rescue applications (Kitano & Tadokoro, 2001) are not alien to conflict. In multi-robot systems for rescue applications, information is usually distributed among the different robotic agents. Coordination among these entities is a problem itself, which becomes more acute when agents' opinions and information conflict. How should these entities solve conflict and rescue as many persons as possible while making an efficient use of the computational resources? Again, software tools and mechanisms are necessary to tackle with conflicts.

The concept of Agreement Technologies (Luck & McBurney, 2008; Ossowsky, 2013; Sierra et al 2011) has been coined in the last few years as an umbrella term for addressing all of those technologies that are envisioned to collaborate, directly or indirectly, to the resolution of conflicts in software systems, and specifically in MAS systems. Even though which works can be considered agreement technologies are arguable (since the contribution to the resolution of a conflict may be indirect), some authors distinguish between several challenges that need to be solved in the so-called agreement technologies.

In this chapter, we position ourselves with the taxonomy/challenges introduced by Sierra et al (2011). Despite the fact that negotiation and argumentation will be the focus of this chapter, we think that it is important to briefly describe every technology involved in agreement technologies. Next, we briefly describe each of the challenges mentioned by Sierra et al. (see Figure 1): (i) Semantics: The current trend of service-oriented computing (Papazoglou & Georgakopoulos, 2003) has changed the way in which complex systems are built. Nowadays, software is built by using diverse services offered by very different providers. Given the heterogeneity of service providers, it is logic to think that service information is provided in different communication languages, and even using different terms to address the same concept (i.e., different ontologies). Whenever a software system needs to cooperate or solve a conflictive situation with other systems, it requires of mechanisms that allow understanding other software systems by matching and aligning ontologies and semantic concepts (Noy, 2004) (Choi, Song & Han, 2006); (ii) Norms: Most distributed applications are no longer static but open, and agents can exhibit a varied spectrum of behaviors. One possible way of "solving" conflict is avoiding conflict, establishing mechanisms that preclude agents of reaching a conflictive situation. Normative systems (Boella, Van Der Torre & Verhagen, 2006; Criado, Argente, & Botti, 2013) are envisioned with such purpose (among others). The society of agents is regulated by norms, which define which actions/states are to be punished in the system (e.g., to avoid conflict) and which actions/states are to be rewarded (e.g., promote actions that avoid conflict); (iii) Organizations: Agents usually have limited computational capabilities. Therefore, if a complex problem needs to be solved, agents need to join together as a group and coordinate to reach such complex goal. Agent organizations (Argente et al, 2011; Esparcia & Argente, 2011; Horling & Lesser, 2008) may be seen as large and implicit coordination mechanisms that establish the roles to be played by agents and the interaction protocols to be carried out among organizational members. In this sense, agent organizations may conflict by strictly defining the structure and interactions of the group; (iv) Trust: Trust mechanisms (Sabater & Sierra, 2005; Such, Espinosa, Garcia-Fornes & Botti, 2011), usually used in concordance with reputation models, are devised to help agents to select whom they should interact with. Trust is formed from one's own past experiences with other agents. To put it simply, positive experiences should bias one agent to collaborate and interact with the other party, whereas negative experiences should bias one agent to avoid interactions with the other party. Reputation is built according to the opinion that agent societies have on individuals. Trust and reputation mechanisms may

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Agreement technologies architecture



help to reduce conflict by interacting with good partners; and, finally, (v) Negotiation & Argumentation, which can be considered the main components, from the point of view of conflict resolution. For this reason, these two topics will be explained in detail in the following sections.

NEGOTIATION

Negotiation is the main technology of agreement technologies that makes possible for agents to solve conflicts. In this case, imitating how humans solve conflicts, agents negotiate looking for an agreement that is acceptable for all of the involved entities. It would not be possible for agents to solve conflict without this technology, at least not in an efficient way. Despite negotiation being crucial for solving conflicts, it should be highlighted that it requires of other technologies as semantics, norms, organizations, and trust to help in the resolution of conflicts. Semantics may help heterogeneous agents to form a negotiation domain (e.g., negotiation problem) that is understandable by all of the parties involved in the conflict situation. Then, society's norms may be used to formally force agents to respect established agreements. Otherwise, agents would violate agreements whenever it suits them. Organizations establish a framework where roles and possible interaction protocols are formalized, giving room negotiations with clear rules of interaction (e.g., negotiation protocols), and helping agents to identify and search conflicting agents based on the information provided by roles (e.g., sellers and buyers are classical roles in conflict). Trust and reputation may guide agents to select negotiation opponents that are more likely to guarantee a good service. Hence, every technology in agreement technologies collaborates along negotiation in leading conflict situations towards good terms. Despite being part of a new topic like agreement technologies, automated negotiation has been studied by scholars for a few years.

Negotiation can be defined as a process in which two or more parties make a joint decision. The parties first verbalize contradictory demands and then move towards agreement by a process of concession-making or search for new alternatives (Pruitt, 1981). Analogously, automated negotiation consists of an automated

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

search process for an agreement between two or more software parties. Two different research lines can be distinguished in automated negotiation models: game-theoretic models and heuristic models. Since the decade of the 50's, automated negotiation has been studied in game theory. Game theory researchers focus on reaching optimal solutions under assumptions of unbounded computational resources, complete/partial information regarding the strategies and preferences of other parties. Some of the most important theoretical results come from game theory, like the work of Nash (1950), Rubinstein (1982), Binmore et al (1986), and more recent studies like Fatima et al. (2004) and In and Serrano (2003). Although game theory studies are interesting from a theoretical point of view, most of them make strong assumptions that may not hold in real applications. For instance, computational resources are of extreme importance for agents since they may be scarce and shared among different tasks. Thus, negotiation should not always assume un-bounded computational resources. Additionally, since agents are heterogeneous, not all of the agents know the same strategies. Identifying which strategies each agent knows may be a hard task that can only be successful after several negotiations. The same goes for the knowledge regarding the opponents' preferences, reservations values, and so forth. Hence, models that tackle uncertainty and limit the use of computational resources are mandatory for some situations.

Heuristic models tackle the problem mentioned above. They do not calculate the optimum agreement, but they obtain results that aim to be as close as possible to the optimum. Heuristic models assume imperfect knowledge regarding the opponent and the environment, and aim to be computationally tractable while obtaining good results. The reader is assumed to have some working knowledge on heuristic models for automated negotiation. In other case, the reading of several introductory texts and reviews like (Guttman, Moukas & Maes, 1998; Beam, Segev, 1997; Kraus, 1997; Jennings et al, 2001; Lomuscio, Wooldridge & Jennings, 2003; Lopes, Wooldridge & Novais, 2008) is recommended. The amount of literature in automated negotiation is vast and immense, ranging from bilateral negotiations, to multi-party negotiations. An extensive review of all of the problems in automated negotiation would be an almost non-feasible task. Following, we discuss some of the most important works in the area of automated negotiation and bilateral negotiation.

Concession Strategies

The classic view of artificial intelligence with respect to negotiation in incomplete information settings is that agents need to eventually concede in order to reach an agreement (Faratin, Sierra, & Jennings, 1998; Bui, Venkatesh, & Kieronska, 1999; Hindriks & Tykhonov, 2008). However, agents can concede in very different magnitudes and in different rounds of the negotiation. Concession strategies determine how the agents concede and when these concessions are carried out.

The most influential work regarding concession strategies is, perhaps, the work of Faratin et al. (1998). The authors proposed concession strategies that are a mix of different families of concession tactics. The authors divide concession tactics into three different families:

- **Time-Dependent Tactics:** These tactics take into account the remaining time in the negotiation to carry out concessions. In this family, we can distinguish between linear tactics, boulware tactics, and conceder tactics. On the one hand, linear tactics carry out the same amount concession at each negotiation round until the reservation value is reached. On the other hand, conceder tactics con-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ceder very rapidly towards the reservation value in the first interactions, whereas boulware tactics concede very slowly during the first negotiation rounds, but it concedes faster as the negotiation process approaches the deadline.

- **Behavior-Dependent Tactics:** In the case of behavior-dependent tactics, the concession carried out by the agent depends on the negotiation movements performed by the opponent in the previous rounds. The classic tactic in this family is tit-for-tat, which mimics the concession carried out by the opponent in the previous round. Other variants of tit-for-tat include random absolute tit-for-tat, which performs the absolute concession carried out by the opponent in the last offer plus/minus a small deviation, and averaged tit-for-tat which takes the window of past opponent offers and carries out the average concession carried out by the opponent.
- **Resource-Dependent Tactics:** This family of tactics computes concession based on the scarceness of a resource in the environment and resource consumption (i.e., time, product quantity, messages, etc.). In general, the scarcer the resource, the more eager should be the agent to maintain/obtain such resource.

Another classic concession based model for bilateral multi-issue negotiations is the Agent Based Market Place (ABMP) framework proposed in (Jonker, Robu & Treur, 2007; Jonker & Treur, 2001). ABMP is a negotiation framework, based on additive utility functions, where proposed bids are concessions to previous bids. The amount of concession is regulated by the concession factor (i.e., reservation utility), the negotiation speed, the acceptable utility gap (the maximal difference between the target utility and the utility of an offer that is acceptable), and the impatience factor, which governs the probability of the agent leaving the negotiation process. Additionally, the framework includes other remarkable characteristics such as the possibility of sharing preference information with the other party, and guessing heuristics that allows agents to determine the ranking of issues and issue values based on the bid history.

Similarity Mechanisms in Negotiation

One of the traditional mechanisms proposed in the literature for solving conflicts is the use of similarity mechanisms. They can be used to solve a current conflict based on solutions given to previous conflicts or as mechanisms that implicitly approximate offers to opponents' preferences. Basically, the two similarity mechanisms more widely used are Case Based Reasoning (Leake, 1994;(Lopez de Mantaras, 2001) and similarity heuristics (Faratin, Sierra & Jennings, 2002; Lai, Sycara & Li, 2008).

Sycara proposed a mediator that uses case based reasoning for solving conflicts in the labor domain (i.e., PERSUADER) (Sycara, 1989, 1990, 1991). PERSUADER takes as input a set of conflicting goals and outputs an agreed plan of actions. The system keeps track of the agreements found in past negotiations and, once a new conflict situation is present, it looks for the most similar past situation. The retrieved agreement is adapted to the present conflict situation, since the rationale behind this heuristic is that similar conflict situations should yield similar solutions.

Another popular use of similarity mechanisms is implicitly approximating one's own proposals to the preferences of the opponent. This is usually carried out by means of similarity heuristics that look for trade-offs. A trade-off consists in decrementing the benefit obtained from some negotiations issues that are not important for us but are important of the other agent, in order to get the decremented benefit as an equivalent increase in the benefit obtained by other issues that are important for us but are not important for the other agent. Faratin et al. (2002) introduced the use of similarity heuristics in

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

bilateral multi-issue negotiations to compute similarity between pairs of others. Given a certain utility u demanded by one of the agents, this agent proposes the offer with utility u that is the most similar to the previous offer proposed by the opponent. The idea behind this heuristic is that the more similar the offer is to the previous opponent offer, the more acceptable it is for the opponent. For computing the similarity between two offers, a fuzzy similarity criterion between issue values. The main drawback of fuzzy similarity heuristics is that they require domain knowledge regarding the similarity between issue values for the opponent.

The use of similarity heuristics was reintroduced again by Lai et al (2008). In this work, a bilateral negotiation protocol for multi-issue negotiations, where agents are capable of sending up to k different offers per round, is presented. The k offers sent by agents are selected from the iso-utility curve, which contains all of the offers with a certain utility. The offer that is selected is the one that is the most similar to the previous opponent offer that reported the most utility. The other k-1 offers are selected randomly from the iso-utility curve. In this case, the similarity heuristic employed is the Euclidean distance. As a similarity measure, Euclidean distance may be less powerful than fuzzy similarity, but it has the advantage of being more general and not requiring domain knowledge.

Bayesian Learning in Negotiation

When reviewing the use of Bayesian learning in negotiation, we cannot forget about the seminal work of Zeng and Sycara (1998). In this article, authors argue about the benefits of using Bayesian models in negotiation. They study a bilateral negotiation case where the buyer attempts to learn the reservation price of a seller by updating its beliefs with Bayesian learning. Despite the fact that it introduces the use of Bayesian learning in negotiation, the applicability of the article is limited since it only focuses on single issue models.

Bayesian classifiers have been used to model the preferences of negotiating agents. In Bui et al. (1999), the authors propose a multi-party cooperative negotiation mechanism for the distributed meeting scheduling domain. Agents start from an initial set of possible agreements and jointly look for good collective agreements by partitioning the set of possible agreements in a tree until a set with only one agreement (leaf node) that is acceptable by all of the agents is found. From the joint set of possible agreements, each agent proposes a partition of such set where the final agreement will be looked for. Agents decide on which set should be explored from all of the partitions that have been proposed. If all of the agents agree on the partition to be selected, the partition becomes the new joint set of possible agreements and the refinement process continues. Otherwise, agents exchange preferences on the proposed partitions and the partition that maximizes the preferences of the group is selected as the next joint set of agreements. In order to save messages exchanged, the agents employ Bayesian classifiers to learn the preferences of other agents according to the information gathered from the cur- rent and past negotiations. Intervals of utility are used as classes and partitions represent attributes of the Bayesian model.

Genetic Algorithms in Negotiation

Genetic Algorithms (GAs) (Goldberg, 1989) have also contributed to the state-of-art in automated negotiation. They are general optimization and learning algorithms based on the evolutionary processes found in the nature. Candidate solutions for a problem form the genetic population of the algorithm, which gradually converges towards high quality solutions by applying genetic operators like mutation

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

and crossover. GAs are general, which means that they do not rely on a specific problem structure. Additionally, they can be used as an implicit learning and adaptation mechanism in environments where dynamics and structure is also uncertain. This is perhaps what makes GA an adequate approach to negotiation problems, since they can be used to learn and adapt both to the opponent and the environment.

The seminal work of GA's in Automated Negotiation is Oliver et al. (1996). They focused on evolving negotiation strategies for bilateral multi-issue negotiations where agents' preferences are represented by means of additive utility functions. In the proposed negotiation framework, a negotiation strategy is a set sequential rules (i.e., rules that are applied in sequential order according to the round), where a rule is a utility threshold that determines if an offer from the opponent is acceptable and a counter-offer to be made to the other party in case that the opponent's offer is not acceptable. Faratin et al. (1998) introduced a negotiation framework for bilateral negotiations where agents' concession strategies can be classified into time-dependent strategies, behavior-dependent strategies and resource-dependent strategies. Matos et al. (1998) proposed a framework where the concession to be carried out in each negotiation issue is a linear combination of the concession of the families of concessions strategies proposed by Faratin. The main research goal of Matos et al. (1998) is determining which the optimal negotiation strategies in different negotiation environments are. For this purpose, an evolutionary process is proposed where the weights given to the concession strategies for each negotiation issue represent a candidate solution in a genetic algorithm. Populations of sellers and buyers with different negotiation strategies negotiate in a round robin way. After each round robin round, negotiation strategies are assigned a fitness value which takes into account the utility obtained in the negotiations and the numbers of messages exchanged. The highest fitness negotiation strategies for sellers and buyers become the parents of the next population of negotiation strategies, which is obtained by the application of genetic operators like mutation and crossover. Eventually, the population of negotiation strategies for sellers and buyers converges towards an optimal set of strategies for the environment under study. The advantage of this proposal with respect to Oliver (1996) is that the evolutionary process does not depend on the number of negotiation rounds but on the number of negotiation issues, which results in a more tractable search space.

Offline Learning in Negotiation

By offline learning we refer to a learning process that is carried out after or before the negotiation process starts. Hence, the model is not updated during the negotiation and it requires of several iterations of the negotiation game to learn an educated model. From the works that we have already reviewed, we can highlight some works like Buffett and Spencer (2007) where the learning of the Bayesian classifiers is carried out before the negotiation starts. However, there are also other approaches that have advocated for the use of learning before or after the negotiation process.

For instance, Coehoorn and Jennings (2004) propose the use of kernel density (Sheather & Jones, 1991) for the estimation of the weights of the opponent's additive utility function. The negotiation model revolves around the idea that a rational agent gradually con- cedes towards its reservation utility, and a rational agent should concede less on the most important issues at the start of the negotiation and concede more on the least important issues. Given this assumption, the agent calculates for each pair of consecutive offers the concession carried out in each issue, and an educated guess of the weight based on such concession. Each tuple, composed of the difference between pairs of consecutive offers, the estimated weight, and the probability density of the weight, forms a three dimensional kernel that is used along the other kernels to calculate an estimation of the real issue weight.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Another set of approaches that heavily rely on offline learning are those approaches based on Artificial Neural Networks (Carbonneau, Kersten & Vahidov, 2008; Carbonneau, Kersten & Vahidov, 2011). In Carbonneau et al. (2008), authors propose a neural network that takes as input the negotiation history of a bilateral negotiation with continuous issues and an offer to make an estimation of the opponent's counter-offer. The major drawback of this approach is that it requires that an artificial neural network is trained per negotiation case. Similarly, the same authors propose an improvement over their previous work in (Carbonneau, Kersten & Vahidov, 2011). It aims to make a predictive model that does not depend on the negotiation case. The model takes pairs of negotiation issues as inputs of the neural network, where one of the issues is considered the primary issue (i.e., independent variable) and the other issue is considered the secondary issue (i.e., dependent variable). The neural network may also take historical information from each issue like the minimum value, maximum value, average value, etc.

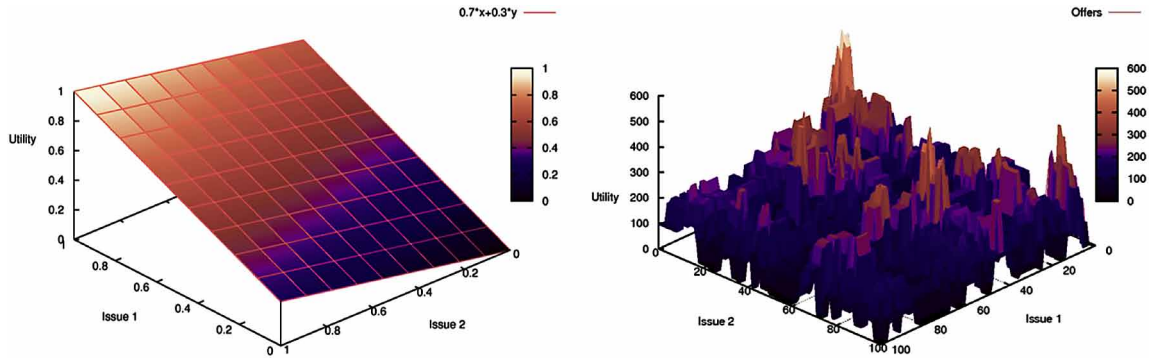
Complex Interdependent Utility Functions

Negotiation processes normally consist in the exchange of proposal between the involved parties. One of the key issues in negotiation strategies is the way in which the agents' preferences are represented. This issue strongly affects how proposals are evaluated and how offers should be generated. In processes where just a single issue is involved, it is quite clear how to evaluate and generate proposals: the value of the issue. However, it is not easy to give a valuation when multiple issues are involved. The multi-attribute utility theory (Wallenius, 2008; Keeney & Raiffa, 1993) comes into play in this case. This theory provides mechanisms for the evaluation of proposals composed of multiple issues. Classic multi-attribute theory has considered that issues are independent. Issue independence means that the value of negotiation issues does not affect the valuation of other issues. Hence, a classic way of representing such preferences is by means of linear additive utility functions. Despite the fact that linear additive functions perform well in some simple domains, there are scenarios where they become poorly suited (Klein, Faratin, Sayama & Bar-Yam, 2003). Just as an example, we could think of a water market domain where two parties negotiate over the exploitation of several water resources. One of the parties' desires to satisfy its water needs whereas the other party has rights over several water exploitations. In this negotiation, the different issues are the water exploitations to be included in the deal. Even though the provider offers a proposal whose amount of water may satisfy the buyer, the value of the proposal may turn into a low utility for the buyer if the water sources are too distant. Thus, some issues have a negative effect over the value of others, and preferences can no longer be represented as classic linear additive utility functions. There is a need to provide complex utility functions that are capable of representing interdependences between negotiation issues.

Negotiation strategies that perform well in domains with linear utility functions may not perform equally in the case of complex interdependent utility functions. In fact, the search space for each agent is much more complicated, needing new negotiation strategies adapted to these complex functions. Figure 2 shows an illustrative example of the search space in the case of linear utility functions and complex interdependent utility functions. The left figure shows the search space of a two issues linear utility function, whereas the right figure shows the search space of a two issues complex interdependent utility function using the model introduced by Ito et al. (2008), which will be reviewed later. As it can be observed, the optimization problem faced by each agent is more complex in the case of interdependent utility functions, leading to the need of new mechanisms that tackle these domains.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. An example of linear (left) and complex (right) utility functions



In this section we have review some important works related with automated negotiation, and more specifically in bilateral negotiation. We have analyzed the adequateness of the different models proposed in the literature for conflict resolution, and we identify those mechanisms that may prove more interesting for solving conflicts. Next section introduces argumentation as a complementary technology to achieve success in negotiation processes among intelligent entities.

ARGUMENTATION

The high dynamism of MAS requires agents to have a way of harmonizing the conflicts that come out when they have to collaborate or coordinate their activities. In those situations, agents need a mechanism to argue (persuade other agents to accept their points of view, negotiating the terms of a contract, etc.) and reach agreements. Argumentation provides a fruitful means of dealing with conflicts and knowledge inconsistencies in MAS. Within the history of research done in argumentation in Artificial Intelligence (AI), there have been a wide variety of mutual contributions. The argumentation theory has produced important benefits on many AI research areas, from its first uses as an alternative to formal logic for reasoning with incomplete and uncertain information to its more recent applications in Multi-Agent Systems (MAS) (Bench-Capon & Dunne, 2007; Rahwan & Simari, 2009). Currently, the study of argumentation in this area has gained a growing interest. The reason behind is that having argumentation skills increases the agents' autonomy and provides them with a mechanism to engage in dialogues and to solve conflicts with other agents demonstrating an intelligent behavior.

An autonomous agent should be able to act and reason as an individual entity on the basis of its mental state (beliefs, desires, intentions, goals, etc.). As a member of a MAS, an agent interacts with other agents whose goals could come into conflict with those of the agent. Moreover, if dynamic and open MAS are considered, the knowledge that an agent has about the environment, its neighbors and its mental state can change in the course of time. In addition, agents can have a social context that imposes dependency relations between them and preference orders among a set of potential values to promote/demote. Therefore, agents must have the ability of reaching agreements that harmonize their mental states and that solve their conflicts with other agents by taking into account their social context and

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

values. Argumentation is a natural way of reaching agreements between several parties with opposing positions about a particular issue. The argumentation techniques, hence, can be used to facilitate the agents' autonomous reasoning and to specify interaction protocols between them (Rahwan, 2006).

Argumentation and Artificial Intelligence

From the beginning of AI research, the development of knowledge-bases, reasoning methods and cognitive architectures able to reply human reasoning has been a core area of interest. The work done in such area is typically known as common sense reasoning research. A reasoning method of this kind must include the following features: (i) the ability to reason with knowledge that is assumed to be true or false in the absence of any evidence that shows the opposite, which is called default reasoning; (ii) the ability to manage uncertain knowledge; (iii) the ability to reason quickly over a wide range of domains; and (iv) the ability to reason and take decisions in presence of incomplete knowledge and subsequently, to revise the beliefs and decisions that were taken when concrete knowledge is acquired, which is called non-monotonic reasoning.

Initially, argumentation theory was adopted in AI due to the inability of the classic propositional logic to reason and give explanations in presence of uncertain or imprecise information (Reiter, 1980). The main problem with classical logic is its monotonic condition, which implies that the acquisition of new information cannot modify the conclusions that were inferred to that moment and thus, it is not applicable as common sense reasoning method. This problem already appeared in rule-based expert systems, where several rules could conflict or even be invalidated by the acquisition of new information. The process of drawing conclusions by using rules that can be defeated by new information is called defeasible reasoning. When defeasible rules are linked up to reach to a conclusion, the proofs that support such rules turn into arguments. The arguments can defeat each other, given rise to an argumentation process. To determine the winning arguments, they must be compared by establishing which beliefs are justified. Therefore, argumentation theory has been studied in AI to deal with the process of argument searching and more concretely (Bench-Capon & Dunne, 2007): (i) to distinguish between valid and invalid arguments; (ii) to identify the protocols and rules that manage the argumentation processes; (iii) to define the argument components and their interaction; and (iv) to determine the conditions under which further discussion becomes redundant.

There are several works that provide an extensive review of the argumentation research that has been done in AI throughout history. According to (Bench-Capon & Dunne, 2007) and (Rahwan & Simari, 2009), the foundations of argumentation in AI lie in the studies done to extend non-classical logic to manage argumentation, the argumentation models that are based on dialogue processes and the diagrammatic treatments of the argument structure. We refer the reader to this important work for further details. Nowadays, the argumentation research in AI is experiencing a new reactivation, mainly motivated by recent and interesting contributions developed in MAS. On the one hand, the argumentation theory has been studied in MAS to manage the agent's practical reasoning. Practical reasoning is a well-known area in philosophy, but which historically has received less attention in AI than the theoretical reasoning. This type of reasoning analyses which specific action should be performed in a particular situation, instead of the theoretical reasoning objective of deciding the truthfulness of beliefs. Moreover, practical reasoning does not presuppose, as theoretical reasoning does, that the fact of reaching an objective is always adequate or profitable, but it must select the best objectives to perform and decide afterwards whether their realization is worthwhile. This fits the reality of a MAS, where each individual agent has its own

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

point of view and its particular objectives and interests. However, the theoretical reasoning about the state of the world and the effects of the potential actions to perform is also essential. Therefore, both types of reasoning must be considered in MAS. In (Rahwan & Amgoud, 2006), an argumentation-based approach for practical reasoning has been proposed. In this work, Dung's abstract argumentation framework (Dung, 1995) is instantiated to generate consistent desires and plans to achieve them. Dung's framework is defined as a pair <A, R> where A is a set of arguments and $R \subseteq A \times A$ is a binary attack relation on A. The works developed by Atkinson in her thesis and her subsequent research are also other important contributions to the modelling of argumentation processes that allow the agents to reason about what is the best action to execute (Atkinson, 2005).

Argumentation in MAS

Moreover, the argumentation techniques have been applied to manage the agents' autonomous reasoning and the interaction between them (Rahwan, 2006). In open MAS, the introduction of new information may give rise to new arguments that reinforce or weaken certain beliefs. Therefore, the argumentation techniques can be used as a way of revising the agents' beliefs in presence of incomplete or uncertain information. The work proposed in (Capobianco, Chesñevar & Simari, 2005) applies argumentation to keep the consistency of the agents' mental state in changing environments by using an appropriate representation of the environment and a mechanism that integrates the new information in the beliefs update process. Argumentation has also been applied in MAS as a selection means between conflicting desires (Amgoud, 2003) and objectives (Amgoud & Kaci, 2004), as a qualitative means of reasoning about the expected value of the realization of certain actions (Fox & Parsons, 1998) and as generator of plans (Hulstijn & van der Torre, 2004) (Simari, Garcia & Capobianco, 2004).

Argumentation provides MAS with a framework that assures a rational communication. The dialogue typology of Walton and Krabbe (Walton & Krabbe, 1995) has been adopted in MAS to classify the different types of dialogues between the agents depending on the objective of the interaction. Other concepts of the argumentation theory (i.e. dialogue games (Hamblin, 1970) and argumentation schemes (Walton, 1996) have also been applied to structure the dialogue between agents with different points of view according to the interaction rules that have been previously agreed. A wide range of approaches that formalize interaction protocols by using different dialogue games have been published (McBurney & Parsons, 2002).

Some examples of dialogue game protocols about specific types of dialogues are: inquiry, persuasion, negotiation and deliberation. As pointed out before, argumentation schemes have several characteristics that make them very useful in defining the communication between agents. In the case of Walton's argumentation schemes, the critical questions are arguments that can be presented by an opponent to criticize the claim that the scheme poses, thus providing the argument with a clear structure that reduces the computational cost of generating and evaluating arguments. Reed & Walton (2005) propose a formal framework to specify argumentation schemes for agent's communication by using the markup language AML, based on XML. Also, the work presented in Atkinson and Bench-Capon (2007) provides firm foundations for an approach to practical reasoning based on presumptive argument.

Among current research on argumentation in MAS, to study the effect of argumentation strategies in the interaction between agents is also a recent trend. Here, there are different approaches to the study of strategies in argumentation frameworks. On the one hand, preliminary works studied the concept of strategy as developing heuristics for move selection in argumentation dialogues. A first contribution was

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

provided in Bench-Capon (1998). In this work the author defines a Toulmin dialogue game machine and proposes some heuristics for move selection. The acceptability of the arguments is computed by using some Toulmin-like rules. A similar work is the one presented in Amgoud and Maudet (2002), which proposes heuristics for move selection on the context of persuasion and negotiation dialogues. This research defends a three-level approach of strategy, inspired on naturally occurring dialogues between humans. The levels identified are: maintaining the focus of the dispute, building one's point of view or destroying the opponent's one, and selecting the method to fulfil the objective set at levels 1 and 2.

While levels 1 and 2 refer to strategy (planning the line of argumentation), level 3 refers to tactic (the mean to reach the aims fixed at the strategically level). Then, the account for strategy proposed follows three steps to develop strategies: (i) define some agent profile: agreeable (accept whenever possible), disagreeable (only accept when there is no reason not to), open-minded (only challenge when necessary), argumentative (challenge whenever possible), or elephant child (question whenever possible); (ii) choose to build or destroy; and (iii) choose some appropriate argumentative content.

This work computes argument acceptability by using a more general Dung-like argumentation framework. In a subsequent work, the author studied the notion of strategy for selecting offers during a negotiation dialogue (Amgoud & Kaci, 2005), proposing different agent's profiles and different criteria for the notions of acceptability and satisfiability of offers. Also, in (Amgoud & Hameurlain, 2006) it is argued that there is no consensus on the definition of a strategy and on the parameters necessary for its definition. Consequently, there are no methodology and no formal models for strategies. This work defends that a strategy is a two steps decision process: (i) to select the type of act to utter at a given step of a dialogue, and (ii) to select the content, which will accompany the act. Thus, an agent tries to choose among different alternatives the best option, which according to its beliefs will satisfy at least its most important goals. There are two types of goals: strategic goals, which help an agent, on the basis of the strategic beliefs, to select the type of act to utter; and functional goals, which help an agent to select, on the basis of the basic beliefs, the content of a move. Then, the work proposes a formal model for defining strategies. The model takes as input the strategic and the functional goals together with the strategic and basic beliefs and returns the next move (act plus its content) to play. Then, the model assesses each alternative by constructing the set of supporting arguments for each one and evaluating their quality.

The agent profiles of Amgoud and Maudet (2002) were also considered in Kakas, Maudet & Moraitis (2005) to develop different types of strategies. This work proposes an argument-based framework for representing communication theories of agents that can take into account the conformance to society protocols, private tactics of individual agents, strategies that reflect different types of personal attitudes (agents' profiles) and adaptability to the particular external circumstances at the time when the communication takes place. Although the authors do not provide a clear structure and definition for their notion of agent society, social relations between agents are captured in the form of preference rules that affect the tactic component of an agent and help it to decide the next move in a dialogue.

On the other hand, a different approach follows a game-theoretic view to the study of argumentation strategies. This is the case of the work proposed in Roth & Rotolo (2007), where the probability of a conclusion is calculated using a standard variant of defeasible logic, in combination with standard probability calculus. In this approach the exchange of arguments is analyzed with game-theoretic tools, yielding a prescriptive account of the actual course of play. Other game-theoretic approach for the study or argumentation strategies in negotiation dialogues was presented in Rahwan & Larson (2009). This approach uses the paradigm of Argumentation Mechanism Design (ArgMD) for designing and analyzing

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

argument evaluation criteria among self-interested agents using game-theoretic techniques. Mechanism design (MD) is a sub-field of game theory concerned with determining the game rules that guarantee a desirable social outcome when each self-interested agent selects the best strategy for itself.

The approach analyses strategy-proofness under grounded semantics for a specific type of arguments, the so-called focal arguments (the arguments that agents are especially interested in being accepted). In a preliminary work (Rahwan & Larson, 2008), the authors restricted the analysis to the case where agents use a specific type of preference criteria, the individual acceptability maximizing preference criteria. Following these criteria, every agent attempts to maximize the number of its arguments that are accepted. In further research the ArgMD approach has been applied to more realistic situations in which each agent has a single focal argument it wishes to have accepted. Authors demonstrate for both preference criteria that if each agent's type (characterized as the set of argument that an agent can bring up) corresponds to a conflict-free set of arguments which does not include (in)direct defeats, the grounded direct argumentation mechanism for this argumentation framework is strategy-proof. Opposite to the heuristic-based approaches, the goal of this game-theoretic approach is to design rules that ensure, under precise conditions, that agents have no incentive to manipulate the outcome of the game by hiding arguments or lying (how to ensure the truth in an argumentation framework).

Within this section it has been shown how argumentation techniques have been successfully used to reach agreements and solve conflicts that assure the coherence of the agents' mental state and to structure their interaction in disagreement situations. Parsons et al. (1998) proposed a seminal theoretical framework that unifies argumentation-based reasoning and communication for negotiation in MAS. After that, Rahwan et al. (2003) analyzed this and other argumentation-based negotiation frameworks. A wide review of the situation of the argumentation research in AI was also published in the special issue on argumentation of the journal Artificial Intelligence (Bench-Capon & Dunne, 2007) and in the book Rahwan & Simari (2009). Moreover, an effort to consolidate the work done in argumentation languages and protocols, argument visualization and editing tools and, generally, in argumentation frameworks for MAS, was performed by the ASPIC project. As a result, a new standard for argument interchange in MAS, the Argument Interchange Format (AIF), was proposed to serve as a convergence point for theoretical and practical work in this area (Willmott, 2006). All these advances show how the study of argumentation in AI, and more concretely in MAS, is currently a research area that has a high activity and a growing interest.

CONCLUSION

In this chapter, we have reviewed the state-of-the-art in automated negotiation from the point of view of Agreement Technologies. A special emphasis was put to analyze which current negotiation models are more convenient for conflict resolution. The analysis allowed us to identify that from the point of view of negotiation, the use of time-based concession tactics, the use of complex and interdependent utility functions for representing agents' preferences, and the use of similarity heuristics and genetic algorithms as learning mechanisms may be of interest in order to achieve more efficient automated conflict resolution tools.

Moreover, this chapter has also introduced several concepts and works of argumentation technology that have been applied to model the reasoning and behavior of agents in MAS. Argumentation technology has been proved as a powerful tool for solving conflicts in the context of a society in a MAS. Neverthe-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

less, the application of argumentation to agent societies is a new area or research with few contributions to date. Commonly, the term agent society is used in the argumentation and AI literature as a synonym for an agent organization or a group of agents that play specific roles, follow some interaction patterns and collaborate to reach global objectives. Many works in argumentation in MAS that refer to the term 'agent societies' follow this approach, which is not targeted to the study of the structure of agent societies and the underlying social dependencies between agents. The influence of the agent group and the social dependencies between agents in the way agents can argue must be further investigated.

REFERENCES

Amgoud, L. (2003). A formal framework for handling conflicting desires. In *7th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU-03*, (LNAI) (vol. 2711, pp. 552–563). Springer. 10.1007/978-3-540-45062-7_45

Amgoud, L., & Hameurlain, N. (2006). A formal model for designing dialogue strategies. In *5th International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS-06*. ACM Press. 10.1145/1160633.1160706

Amgoud, L., & Kaci, S. (2004). On the generation of bipolar goals in argumentation-based negotiation. In *Argumentation in Multi-Agent Systems: State of the art survey* (LNAI), (Vol. 3366, pp. 192–207). Springer.

Amgoud, L., & Kaci, S. (2005). On the study of negotiation strategies. In *AAMAS 2005 Workshop on Agent Communication*. ACM Press.

Amgoud, L., & Maudet, N. (2002). Strategical considerations for argumentative agents (preliminary report). In *9th International Workshop on Non-Monotonic Reasoning* (LNAI), (pp. 399–407). Springer.

Argente, E., Botti, V., Carrascosa, C., Giret, A., Julian, V., & Rebollo, M. (2011). An abstract architecture for virtual organizations: The THOMAS approach. *Knowledge and Information Systems*, *29*(2), 379–403. doi:10.100710115-010-0349-1

Argente, E., Botti, V., Carrascosa, C., Giret, A., Julian, V., & Rebollo, M. (2011). An Abstract Architecture for Virtual Organizations: The THOMAS approach. *Knowledge and Information Systems*, *29*(2), 379–403. doi:10.100710115-010-0349-1

Atkinson, K. (2005). *What Should We Do? Computational Representation of Persuasive Argument in Practical Reasoning*. (PhD thesis). Liverpool University.

Atkinson, K., & Bench-Capon, T. (2007). Practical reasoning as presumptive argumentation using action based alternating transition systems. *Artificial Intelligence*, *171*(10-15), 855–874. doi:10.1016/j.artint.2007.04.009

Beam, C., & Segev, A. (1997). Automated negotiations: A survey of the state of the art. *Wirtschaftsinformatik*, *39*(3), 263–268.

Bench-Capon, T., & Dunne, P. (2007). Argumentation in artificial intelligence. *Artificial Intelligence*, *171*(10-15), 619–938. doi:10.1016/j.artint.2007.05.001

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bench-Capon, T. J. (1998). Specification and implementation of Toulmin dialogue game. In *International Conferences on Legal Knowledge and Information Systems*. IOS Press.

Binmore, K., Rubinstein, A., & Wolinsky, A. (1986). The Nash bargaining solution in economic modelling. *The Rand Journal of Economics*, *17*(2), 176–188. doi:10.2307/2555382

Boella, G., Van Der Torre, L., & Verhagen, H. (2006). Introduction to normative multiagent systems. *Computational & Mathematical Organization Theory*, *12*(2), 71–79. doi:10.100710588-006-9537-7

Browder, G. (2000). An analysis of the negotiations for the 1995 mekong agreement. *International Negotiation*, *5*(2), 237–261. doi:10.1163/15718060020848758

Buffett, S., & Spencer, B. (2007). A bayesian classifier for learning opponents' preferences in multi-object automated negotiation. *Electronic Commerce Research and Applications*, *6*(3), 274–284. doi:10.1016/j.elerap.2006.06.008

Bui, H., Venkatesh, S., & Kieronska, D. (1999). Learning Other Agents' Preferences in Multi-Agent Negotiation Using the Bayesian Classifier. *International Journal of Cooperative Information Systems*, *8*(4), 275–293. doi:10.1142/S0218843099000149

Capobianco, M., Chesñevar, C. I., & Simari, G. R. (2005). Argumentation and the dynamics of warranted beliefs in changing environments. *Autonomous Agents and Multi-Agent Systems*, *11*(2), 127–151. doi:10.100710458-005-1354-8

Carbonneau, R., Kersten, G., & Vahidov, R. (2011). Pairwise issue modeling for negotiation counteroffer prediction using neural networks. *Decision Support Systems*, *50*(2), 449–459. doi:10.1016/j.dss.2010.11.002

Carbonneau, R., Kersten, G. E., & Vahidov, R. (2008). Predicting opponent's moves in electronic negotiations using neural networks. *Expert Systems with Applications*, *34*(2), 1266–1273. doi:10.1016/j.eswa.2006.12.027

Choi, N., Song, I. Y., & Han, H. (2006). A survey on ontology mapping. *SIGMOD Record*, *35*(3), 34–41. doi:10.1145/1168092.1168097

Coehoorn, R., & Jennings, N. (2004). Learning on opponent's preferences to make effective multi-issue negotiation trade-offs. In *The 6th International Conference on Electronic Commerce (ICEC'04)*, (pp. 59–68).

Criado, N., Argente, E., & Botti, V. (2013). THOMAS: An Agent Platform For Supporting Normative Multi-Agent Systems. *Journal of Logic and Computation.*, *23*(2), 309–333. doi:10.1093/logcom/exr025

Dung, P. M. (1995). On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming, and n-person games. *Artificial Intelligence*, *77*(2), 321–357. doi:10.1016/0004-3702(94)00041-X

Esparcia, S., & Argente, E. (2011). *Defining Virtual Organizations Following a Formal Approach. In Agents and Artificial Intelligence 2011*. Revised Selected Papers.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Faratin, P., Sierra, C., & Jennings, N. (1998). Negotiation Decision Functions for Autonomous Agents. *International Journal of Robotics and Autonomous Systems*, *24*(3-4), 159–182. doi:10.1016/S0921-8890(98)00029-3

Faratin, P., Sierra, C., & Jennings, N. (2002). Using Similarity Criteria to Make Issue Trade-Offs in Automated Negotiations. *Artificial Intelligence*, *142*(2), 205–237. doi:10.1016/S0004-3702(02)00290-4

Fatima, S. S., Wooldridge, M. J., & Jennings, N. R. (2004). An agenda-based framework for multi-issue negotiation. *Artificial Intelligence*, *152*(1), 1–45. doi:10.1016/S0004-3702(03)00115-2

Foster, I., Kesselman, C., & Tuecke, S. (2001). The anatomy of the grid: Enabling scalable virtual organizations. *International Journal of High Performance Computing Applications*, *15*(3), 200–222. doi:10.1177/109434200101500302

Fox, J., & Parsons, S. (1998). Arguing about beliefs and actions. In *Applications of Uncertainty Formalisms* (Vol. 1455, pp. 266–302). Springer. doi:10.1007/3-540-49426-X_13

Goldberg, D. E. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Boston, MA: Addison-Wesley Longman Publishing Co., Inc.

Guttman, R. H., Moukas, A. G., & Maes, P. (1998). Agent-mediated electronic commerce: A survey. *The Knowledge Engineering Review*, *13*(2), 147–159. doi:10.1017/S0269888998002082

Hamblin, C. L. (1970). *Fallacies*. Methuen Co. Ltd.

Hewitt, C. (1991). Open information systems semantics for distributed artificial intelligence. *Artificial Intelligence*, *47*(1-3), 79–106. doi:10.1016/0004-3702(91)90051-K

Hindriks, K. V., & Tykhonov, D. (2008). Opponent modelling in automated multi-issue negotiation using Bayesian learning. In *The 7th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'08)*.

Horling, B., & Lesser, V. (2008). Using quantitative models to search for appropriate organizational designs. *Autonomous Agents and Multi-Agent Systems*, *16*(2), 95–149. doi:10.100710458-007-9020-y

Hübner, J. F., Sichman, J. S., & Boissier, O. (2002). A model for the structural, functional, and deontic specification of organizations in multiagent systems. *Advances in Artificial Intelligence*.

Hulstijn, J., & van der Torre, L. (2004). Combining goal generation and planning in an argumentation framework. In *10th International Workshop on Non-Monotonic Reasoning* (LNAI). Springer-Verlag.

In, Y., & Serrano, R. (2003). Agenda restrictions in multi-issue bargaining (II): Unrestricted agendas. *Economics Letters*, *79*(3), 325–331. doi:10.1016/S0165-1765(02)00321-X

Ito, T., Klein, M., & Hattori, H. (2008). A multi-issue negotiation protocol among agents with nonlinear utility functions. *Multiagent and Grid Systems*, *4*(1), 67–83.

Jennings, N. R., Faratin, P., Lomuscio, A. R., Parsons, S., Wooldridge, M. J., & Sierra, C. (2001). Automated negotiation: Prospects, methods and challenges. *Group Decision and Negotiation*, *10*(2), 199–215. doi:10.1023/A:1008746126376

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Jennings, N. R., Faratin, P., Lomuscio, A. R., Parsons, S., Wooldridge, M. J., & Sierra, C. (2001). Automated Negotiation: Prospects, Methods and Challenges. *Group Decision and Negotiation*, *10*(2), 199–215. doi:10.1023/A:1008746126376

Jonker, C. M., Robu, V., & Treur, J. (2007). An agent architecture for multi-attribute negotiation using incomplete preference information. *Autonomous Agents and Multi-Agent Systems*, *15*(2), 221–252. doi:10.100710458-006-9009-y

Jonker, C. M., & Treur, J. (2001). An Agent Architecture for Multi-Attribute Negotiation. In *The 17th International Joint Conference on Artificial Intelligence (IJCAI'01)*.

Kakas, A., Maudet, N., & Moraitis, P. (2005). Modular Representation of Agent Interaction Rules through Argumentation. *Autonomous Agents and Multi-Agent Systems*, *11*(2), 189–206. doi:10.100710458-005-2176-4

Keeney, R. L., & Raiffa, H. (1993). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. Cambridge University Press. doi:10.1017/CBO9781139174084

Kitano, H., & Tadokoro, S. (2001). RoboCup Rescue: A Grand Challenge for Multiagent and Intelligent Systems. *AI Magazine*, *22*(1), 39–52.

Klein, M., Faratin, P., Sayama, H., & Bar-Yam, Y. (2003). Negotiating Complex Contracts. *Group Decision and Negotiation*, *12*(2), 111–125. doi:10.1023/A:1023068821218

Kraus, S. (1997). Negotiation and cooperation in multi-agent environments. *Artificial Intelligence*, *94*(1-2), 79–97. doi:10.1016/S0004-3702(97)00025-8

Lai, G., Sycara, K., & Li, C. (2008). A decentralized model for automated multi-attribute negotiations with incomplete information and general utility functions. *Multiagent and Grid Systems*, *4*(1), 45–65.

Leake, D. B. (1994). Case-based reasoning. *The Knowledge Engineering Review*, *9*(01), 61–64. doi:10.1017/S0269888900006585

Lomuscio, A., Wooldridge, M., & Jennings, N. (2003). A Classification Scheme for Negotiation in Electronic Commerce. *Group Decision and Negotiation*, *12*(1), 31–56. doi:10.1023/A:1022232410606

Lopes, F., Wooldridge, M., & Novais, A. Q. (2008). Negotiation among autonomous computational agents: Principles, analysis and challenges. *Artificial Intelligence Review*, *29*(1), 1–44. doi:10.100710462-009-9107-8

Lopez de Mantaras, R. (2001). Case-Based Reasoning. In Machine Learning and Its Applications (LNCS), (Vol. 2049, pp. 127-145). Springer. doi:10.1007/3-540-44673-7_6

Luck, M., & McBurney, P. (2008). Computing as Interaction: Agent and Agreement Technologies. In *Proc. IEEE Conference on Distributed Human-Machine Systems*.

Matos, N., Sierra, C., & Jennings, N. (1998). Determining Successful Negotiation Strategies: An Evolutionary Approach. In *The 3rd International Conference on Multi Agent Systems (ICMAS '98)*. Washington, DC: IEEE Computer Society. 10.1109/ICMAS.1998.699048

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

McBurney, P., & Parsons, S. (2002). Dialogue games in multi-agent systems. Informal Logic. *Special Issue on Applications of Argumentation in Computer Science*, *22*(3), 257–274.

Nash, J. F. (1950). The bargaining problem. *Econometrica*, *18*(2), 155–162. doi:10.2307/1907266

Noy, N. F. (2004). Semantic integration: A survey of ontology-based approaches. *SIGMOD Record*, *33*(4), 65–70. doi:10.1145/1041410.1041421

Oliver, J. R. (1996). On Artificial Agents for Negotiation in Electronic Commerce. In *The 29 Hawaii International Conference on System Sciences (HICSS'96)*. Washington, DC: IEEE Computer Society. 10.1109/HICSS.1996.495355

Osborne, M. J., & Rubinstein, A. (1999). *A Course in Game Theory*. MIT Press.

Ossowski, S. (2013). *Agreement Technologies. In Law, Governance and Technology Series* (Vol. 8). Springer.

Papazoglou, M. P., & Georgakopoulos, D. (2003). Service-oriented computing. *Communications of the ACM*, *46*(10), 25–28.

Parsons, S., Rodriguez-Aguilar, J. A., & Klein, M. (2011). Auctions and bidding: A guide for computer scientists. *ACM Computing Surveys*, *43*(2), 10. doi:10.1145/1883612.1883617

Parsons, S., Sierra, C., & Jennings, N. R. (1998). Agents that reason and negotiate by arguing. *Journal of Logic and Computation*, *8*(3), 261–292. doi:10.1093/logcom/8.3.261

Pinyol, I., & Sabater-Mir, J. (2013). Computational trust and reputation models for open multi-agent systems: A review. *Artificial Intelligence Review, Springer Netherlands.*, *40*(1), 1–25. doi:10.100710462-011-9277-z

Pruitt, D. G. (1981). *Negotiation Behavior*. Academic Press.

Rahwan, I. (2006). *Argumentation in multi-agent systems*. Autonomous Agents and Multiagent Systems. *Guest Editorial*, *11*(2), 115–125.

Rahwan, I., & Amgoud, L. (2006). An argumentation-based approach for practical reasoning. In *5th International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS-06*. ACM Press. 10.1145/1160633.1160696

Rahwan, I., & Larson, K. (2008). Mechanism design for abstract argumentation. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems* (vol. 2, pp. 1031–1038). ACM Press.

Rahwan, I., & Larson, K. (2009). *Argumentation and Game Theory*. Argumentation in Artificial Intelligence.

Rahwan, I., Ramchurn, S. D., Jennings, N. R., Mcburney, P., Parsons, S., & Sonenberg, L. (2003). Argumentation-based negotiation. *The Knowledge Engineering Review*, *18*(04), 343–375. doi:10.1017/S0269888904000098

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Rahwan, I., Ramchurn, S. D., Jennings, N. R., McBurney, P., Parsons, S., & Sonenberg, L. (2003). Argumentation-based negotiation. *The Knowledge Engineering Review*, *18*(4), 343–375. doi:10.1017/S0269888904000098

Rahwan, I., & Simari, G. (Eds.). (2009). *Argumentation in Artificial Intelligence*. Springer.

Rebollo, M., Carrascosa, C., & Palomares, A. (2014). Follow the leader in a consensus network as a solution to manage a smart grid: the Balearic Islands case. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems* (AAMAS '14). International Foundation for Autonomous Agents and Multiagent Systems.

Reed, C., & Walton, D. (2005). Towards a formal and implemented model of argumentation schemes in agent communication. *Autonomous Agents and Multi-Agent Systems*, *11*(2), 173–188. doi:10.100710458-005-1729-x

Reiter, R. (1980). A logic for default reasoning. *Artificial Intelligence*, *13*(1-2), 81–132. doi:10.1016/0004-3702(80)90014-4

Roth, B., & Rotolo, A. (2007). Strategic argumentation: a game theoretical investigation. In *Proceedings of the Eleventh International Conference on Artificial Intelligence and Law*. ACM Press.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, *50*(1), 155–162. doi:10.2307/1912531

Sabater, J., & Sierra, C. (2005). Review on computational trust and reputation models. *Artificial Intelligence Review*, *24*(1), 33–60. doi:10.100710462-004-0041-5

Sanchez-Anguix, V., Julian, V., Botti, V., & García-Fornes, A. (2013). Tasks for agent-based negotiation teams: Analysis, review, and challenges. *Engineering Applications of Artificial Intelligence*, *26*(10), 2480–2494. doi:10.1016/j.engappai.2013.07.006

Sheather, S. J., & Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society. Series B. Methodological*, 683–690.

Sierra, C., Botti, V., & Ossowski, S. (2011). Agreement computing. *KI-Künstliche Intelligenz*, *25*(1), 57–61. doi:10.100713218-010-0070-y

Sierra, C., Botti, V., & Ossowski, S. (2011). *Agreement Computing*. KI-Kunstliche Intelligenz.

Simari, G. R., Garcia, A. J., & Capobianco, M. (2004). Actions, planning and defeasible reasoning. In *10th International Workshop on Non-Monotonic Reasoning* (LNAI), (pp. 377–384). Springer.

Such, J. M., Espinosa, A., Garcia-Fornes, A., & Botti, V. (2011). Partial identities as a foundation for Trust and Reputation. *Engineering Applications of Artificial Intelligence*, *24*(7), 1128–1136. doi:10.1016/j.engappai.2011.06.008

Sycara, K. P. (1989). *Multiagent compromise via negotiation*. In Distributed. *Artificial Intelligence*, *2*, 119–137.

Sycara, K. P. (1990). Negotiation Planning. *An AI Approach*, *46*(1), 216–234.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sycara, K. P. (1991). Problem restructuring in negotiation. *Management Science*, *37*(10), 1248–1268. doi:10.1287/mnsc.37.10.1248

Wallenius, J., Dyer, J. S., Fishburn, P., Steuer, R., Zionts, S., & Deb, K. (2008). Multiple Criteria Decision Making, Multiattribute Utility Theory: Recent Accomplishments and What Lies Ahead. *Management Science*, *54*(7), 1336–1349. doi:10.1287/mnsc.1070.0838

Walton, D. (1996). *Argumentation Schemes for Presumptive Reasoning*. Routledge Publishers.

Walton, D., & Krabbe, E. C. W. (1995). *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. State University of New York Press.

Willmott, S., Vreeswijk, G., Chesñevar, C., South, M., McGinnis, J., Modgil, S., ... Simari, G. (2006). Towards an argument interchange format for Multi-Agent Systems. In *3rd International Workshop on Argumentation in Multi-Agent Systems, ArgMAS-06*. ACM Press.

Zeng, D., & Sycara, K. (1998). Bayesian learning in negotiation. *International Journal of Human-Computer Studies*, *48*(1), 125–141. doi:10.1006/ijhc.1997.0164

KEY TERMS AND DEFINITIONS

Agreement Technologies: All of those technologies that are envisioned to collaborate, directly or indirectly, to the resolution of conflicts in software systems, and specifically in MAS systems.

Argumentation: A verbal, social, and rational activity aimed at convincing a reasonable critic of the acceptability of a standpoint by putting forward a constellation of propositions justifying or refuting the proposition expressed in the standpoint.

Negotiation: Process in which a joint decision is made by two or more parties.

Open Multi-Agent System: A distributed system where heterogeneous agents, with their own goals, can enter and leave the system during the life of the system.

Reputation: What a society says about a target (probably an entity) regarding his behavior.

Trust: An expectation about an uncertain behavior.

Virtual Organization: A set of individuals and institutions that need to coordinate resources and services across institutional boundaries.

This research was previously published in Interdisciplinary Perspectives on Contemporary Conflict Resolution edited by Davide Carneiro and Paulo Novais; pages 147-167; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 25
Virtual Soar-Agent Implementations:
Examples, Issues, and Speculations

Jeremy Owen Turner
Simon Fraser University, Canada

ABSTRACT

This chapter provides a brief overview of those virtual agent implementations directly inspired by the cognitive architecture: Soar. This chapter will take a qualitative approach to discussing examples of virtual Soar-agents. Finally, this chapter will speculate on the future of Soar virtual characters. The goals of this chapter are sixfold. The first goal is to explain why cognitive architectures are becoming increasingly important to virtual agent design(s). The second goal is to convey why this chapter focuses exclusively on virtual agents that utilize the Soar architecture. The third goal is to explore some of Soar's technical details. The fourth goal is to showcase a few diverse examples where Soar is beginning to have a design impact on virtual agents. The fifth goal addresses Soar's limitations – when applied to agent design in virtual environments. The final goal speculates on ways Soar can be expanded for virtual agent design(s) in the future.

INTRODUCTION

Cognitive architectures have been around for decades.[1] Many of the more canonical architectures such as CLARION (CLARION webpage, n.d.) and Soar (Soar webpage, n.d.) have been continually updated over the years for agent implementation by external developers, academics and hobbyists. However, most virtual agents (e.g. bots, Non-player-characters [NPCs]) created up to the present day, still do not directly utilize cognitive architectures for interaction with the agent's embodied environment. Instead, such agents often depend on cognition-less systems that may include the use of: finite-state machines, subsumption-level (i.e. reactive) algorithms (Brooks, 1991), hierarchical search interaction trees, heuristic search, and brute-force techniques, or some combination thereof. In recent years, computing power has exponentially increased and this has allowed for a serious reconsideration of cognitive architecture

DOI: 10.4018/978-1-7998-0951-7.ch025

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

implementations. This new technological climate is beginning to encourage tractable innovations in artificial general intelligence (AGI) generally, and experimentation with cognitive architectures in particular. Academics and developers are increasingly interfacing cognitive architectures with virtual agents for more deliberation and cognitive potential in their interaction design.

This chapter provides an opportunity to qualitatively discuss and tally some of the state-of-the-art of those virtual agents that employ a particular canonical cognitive architecture, Soar (State Operator And Result)[2] (invented by Laird, Newell & Rosenbloom, ca. 1983). There are numerous cognitive architectures existing today (e.g. ACT-R, CLARION, RASCALS etc.) but due to size and other limitations explained below, Soar will be the focus of this chapter.

The thematic goals of this chapter are sixfold. The first goal is to explain what a cognitive architecture is and why they are becoming increasingly important to virtual agent design(s). The second goal is to convey why this particular chapter focuses exclusively on virtual agents that utilize the Soar architecture. The third goal is to explore some technical details that are uniquely characteristic of Soar's handling of cognition. The fourth goal is to showcase - via a few diverse examples - the continued legacy and design impact Soar is beginning to have on virtual agents. The fifth goal is to address Soar's limitations – when applied to NPC/bot design in virtual environments. The final goal is to speculate on additional ways Soar can be expanded for virtual agent design in the near future.

CAVEAT: VIRTUAL AGENTS MIGHT NOT EVEN REQUIRE COGNITIVE ARCHITECTURES

Design motivations behind whether or not to deploy a cognitive architecture often varies according to whether the virtual agent is meant to academically evaluate embodied and/or extended cognition or instead, to implement a virtual character (NPC) for social and/or ludic entertainment purposes. With these contextual contingencies in mind, most contemporary virtual agents are not required to be cognizant of the visual inputs beyond the raw numerical data underlying each object and path. Basically, a virtual agent can successfully interact and navigate its environment without being directly aware of it in any introspective way. An agent can react to numbers alone. For example, a bit-mapped image or any other virtual object can be represented simply as a stream of numeric representations (e.g. binary digits). These numbers provide the agent with coordinates, proximity-relations, and other manually coded identification data that allow the agent to respond in a completely reactive way. In many video-games, for example, an NPC can appear to be intelligent simply by mindlessly reacting to the changes in numerical inputs afforded by the dynamic changes occurring via game-play interaction. However, cognitive architectural enhancement might be useful for any agent-based cognitive task involving reasoning and mnemonic assistance. Otherwise, a cognitive architecture may only be useful in this numerically-reduced case if we wish to imbue the agent with the ability to contemplate its environment as an "[...] idealized version of the underlying reality" (Best et al., 2006, p. 186). This underlying numerical reality might be modeled to enable embodied, extended, situated, and distributed (EESD) cognition so that virtual agents can actively "[...] structure or manipulate their environments" to "simplify, sequence, or otherwise, support cognitive processes" (Smart & Sycara, 2015, p. 3837).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The Need for Integrating Cognitive Architectures with Virtual Agents

Most virtual agents to date – whether embodied in video-games, interaction simulators or social virtual worlds - are not enabled with any cognitive architectures or algorithms modeling deeper cognitive processes. Cognitive representation of some kind is crucial if designers and developers wish to create artificially intelligent agents capable of some manner of introspective thought and deliberation – before executing an action. Due to the simplicity of construction and the current cost of available computational resources, behavioral regulation comes instead in the form of finite-state machines, and/or purely reactive systems where a particular input immediately corresponds to an (observable) output. Some designers see the behavioral affordances of finite-state machine-driven characters as "brittle" (Best et al., 2006, p. 190). Further, as the number of states (and state-transitions) increase in more behaviorally realistic games, the more states the programmers have to manually account for. Further, each of these states would require the programmer to account for an increasing number of pre-programmed responses while entering, transitioning to, or experiencing these additional states. At worst, "[...] human players often learn to game the finite state machine and take advantage of the idiosyncrasies of the opponents" (Best et al., 2006, p. 190). If virtual agents are one day intended to be perceived by gamers and other interacting avatars as "believable characters", then it seems that one should expect an increasing integration in the near-term of cognitive architectures as well as cognition-inspired algorithms and models in a virtual agent's deliberative and behavioral design.

COGNITIVE ARCHITECTURES

A cognitive architecture is semantically distinct from any other system or algorithmic structure that parses and models representations of cognition. To qualify this distinction, one must first clarify those categories that are not considered to be cognitive architectures. Firstly, an agent imbued with cognitive potential might utilize a cognitive architecture as one of many possible modular components for deliberation and (inter)action. However, the agent itself, is not architectural. Cognitive architectures are ontologically distinct from cognitive agents in that they resemble computer architectures. Like computer architectures, cognitive architectures specialize in parsing: "memories, processing and control components, representations of data, and input/output devices" (Laird, 2012, p. 5-6). A cognitive agent using any cognitive architecture ideally involves Artificial General Intelligence (AGI) not for computation's sake, but only for "[symbolic] representation, acquisition, and use of knowledge to pursue goals." (Laird et al., 2012, p. 6). Secondly, a cognitive model resembles a cognitive architecture however, "[...] a number of task-specific cognitive models can be developed that all draw on the same underlying (cognitive) mechanisms [which would comprise an architecture]" (Smart et al., 2014a, p. 3). A cognitive (infra)architecture can summarize a "common set of underlying mechanisms" to unify and explain "[...] a number of otherwise disparate behaviors" (Smart et al., 2014a, p.3).

A cognitive architecture is fundamentally, more general-purpose in that it can be applied towards a host of different agents and multi-agent systems. An agent uses a portion of an architectures' "fixed and task independent" (Smart et al., 2014a, p. 6) attributes. These attributes are customarily identified as: memories, natural-language patterns, data-structures and algorithms. These attributes learn and encode specific knowledge configurations, problem-specifications, rational deliberations, and actions for goal-directed intelligent performance in a stochastic environment (e.g. virtual world) (Smart et al., 2014a,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

p. 8). Cognitive architects usually understand that their cognitive architecture cannot be so general purpose as to handle every possible contingency and action. Due to the stochastic nature of an external environment,cogntive agents are assumed to "exist in an environment and [...] pursue tasks similar to those we find in the world we inhabit". Therefore, cognitive agents – regardless of architectural type - must be custom tailored to handling situations unique to their task domain (Smart et al., 2014a, p. 28). Some cognitive architectures assert their cognitive influence on a single agent by focusing on introspection and metareasoning capabilities. To clarify, metareasoning is the "process of [recursively] reasoning about [its own] reasoning cycle" (Cox & Raja, 2011, p. 4).The strong AI community traditionally employs a robust cognitive architecture (e.g. Soar) to model an agent's reasoning or metareasoning capabilities in the form of a "decision cycle within an action-perception loop […] (Cox & Raja, 2011, p. 4).

Why Discuss Soar and Not Other Cognitive Architectures?

Other than Soar, there are a few canonical cognitive architectures that can be (and are being used) for tractable virtual agent implementations. The three cognitive architectures - briefly discussed below in chronological order - are by no means representative of all cognitive architectures but are sufficient for encapsulating three basic architectural types as belonging to one canonical corpus.

The most established cognitive architecture for virtual agents is known as ACT-R [Adaptive Control of Thought-Rational] (Anderson ca. 1990)]. ACT-R was an early model for representing brain-like cognition and was re-iterated from previous models known as "HAM" (1973) and "ACT-*" (1983).[3] ACT-R is one of the first cognitive architectures to explore and implement the relationships between deliberation, conceptualization memorization, and action. Further, ACT-R is an early instantiation of a hybrid architecture that dynamically draws from: symbolic, rule-based, and statistically tunable sub-symbolic representations of memory and cognition (Best et al., 2006, p. 188).

ACT-R has been implemented primarily for use by NPCs (non-player characters)[4] in first-person shooter video games for pathfinding, maze navigation, and strategic team communication (Best et al., 2006). Additionally, ACT-R has been occasionally deployed for more cerebral social simulations modeling human cognition and/or human-surrogate attributes.[5] For example, some virtual agents use a centralized intelligence network to send and receive news items in order to semantically interpret and predict terrorist attacks (Smart et al., 2014). The bulk of these particular ACT-R projects are military affiliated and/or directly funded. As mentioned previously, ACT-R is perhaps the most widely used cognitive architecture for virtual agent implementations. Therefore, it was decided that this chapter's contribution would be a focus on a cognitive architecture that is just as canonical as ACT-R but was not quite as popular and whose virtual agent implementations had not yet been summarized into one text.

Another canonical cognitive architecture is CLARION (Connectionist Learning with Adaptive Rule Induction ON-Line) (Sun et al., 2006). CLARION, like ACT-R, is a hybrid architecture that is focused on distinguishing between implicit and explicit cognition by using means of distributed representation.[6] For implicit cognition, CLARION can also parse subconscious processes, sub-symbolic knowledge (Sun, 2006, p. 82-83) and employ tacit knowledge. This implicit knowledge is effectively a priori knowledge which, like ACT-R, is represented and managed using bottom-level functional approximations of neural-nets. CLARION's neural nets are quite elaborate as they can support: back-propagation, feed-forward behaviors (Sun, 2006 p. 82), and Q-learning techniques (Watkins, 1989 in Sun, 2006., p. 83, 85). For explicit cognition, CLARION focuses on levels of: conscious awareness, active introspection, and intentional deliberation. CLARION frequently generates elements of focused top-down knowledge from

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

bottom-up intuition in the form of: deductions, inductions, abductions, hypothesis testing, and explicit symbolic representation (Sun, 2006, p. 83-84). CLARION is also one of the few cognitive architectures capable of representing meta-cognition (Sun, 2006, p. 92-94). Of all the canonical cognitive architectures, CLARION theoretically has the most potential for implementing virtual agents with cognitive faculties for self-awareness, innate motivations and drives, complex strategic deliberation, and even learning from reactive behavior. However, CLARION's issue is the reverse of ACT-R in that there are very few known virtual agent implementations of CLARION. Because of this under-representation, the main active developer of CLARION has even decided to translate its entire code-base from C# into Python in order to entice character designers to use CLARION for their virtual agents.[7] Therefore, a discussion of CLARION will be reserved for a future paper on virtual agent implementations. It is hoped that tractable CLARION examples will become more plentiful over time.

There are many other cognitive architectures. Some of these architectures have emerged from fringe beginnings. However, due to interesting implementations with virtual agents, the occasional architecture had the potential to become canonical in its own right. Take the example of RASCALS (Rensselaer Advanced Synthetic Character Architecture for "Living" Systems) (Bringsjord et al., 2005). The RASCALS cognitive architecture was developed to explore the mathematical modeling of morality in AGI-enabled agents. RASCALS "[...] uses simple logical systems (first-order ones) for low-level (perception & action) and mid-level cognition, and advanced logical systems (e.g., epistemic and deontic logics) for more abstract cognition" (Bringsjord et al., 2005, p. 31). RASCALS' primary virtual agent was a personified embodiment of moral evil existing in the social virtual world Second Life ("Second Life", n.d.)[8] - aptly named, "E". RASCALS has been shown to be technically distinguished from CLARION due to the additional aforementioned emphasis of natural language processing capabilities (Bringsjord et al., 2005, p. 31). RASCALS can also be distinguished from earlier versions of other canonical cognitive architectures based on production rules such as Soar. Once again, RASCALS, like CLARION, has not included many virtual agent examples from which to examine. In fact, the RASCALS project is no longer formally maintained by Bringjord's research team. Therefore, the rest of this discussion will focus primarily on Soar. Soar is an architecture with enough examples to examine and also deserves a higher-profile for its under-represented virtual agent implementations.

Soar

Soar ("State Operator And Result") is a production rule-based AGI cognitive architecture with military and entertainment clientele. However, Soar's creators claim that Soar-agent developers do not need previous programming nor computer-science experience to understand its architectural syntax. Soar is also appealing as a cognitive architecture because it is generalizable enough to handle different: problem-space, tasks, states, and operators. Ultimately, Soar is known for handling multi-faceted problem-domain approaches to a particular task (Laird, 2012, p. 4).

Canonical Features of Soar

Soar is highly scalable to support a large number of rules. In recent versions, Soar has also provided reinforcement learning capabilities. The most recent versions of Soar even have found a way to represent affective states and mental imagery processing (Lathrop & Laird, 2009). Unlike other Rule Based Systems (RBS), Soar-agents have the ability to dynamically update their beliefs about the external

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

environment via "computationally inexpensive truth maintenance algorithms"("SoarTech Corporate Brochure", n.d.). Having a flexible belief system allows a Soar-agent to adapt to multiple ontological commitment opportunities within its particular task-domain. Soar's growing number of memory banks are completely associative, and currently store procedural, semantic, and episodic memories. There is even active speculation that Soar might include a predictive prospective memory bank in the near future (Li & Laird, 2013).

Soar's chunking mechanism is also rule-based and is a form of "explanation based learning" (Laird, 2012, p. 8) that is consolidated in episodic memory. Rule-based chunking allows a Soar-agent to compress a batch of pre-conditions from a specific episodic experience into a single generalized pre-condition for later possible activation of an operator that would be governed by a production rule. Soar's chunking procedure governs all learning from input through deliberation on the object-level and back to output via proposing, then selecting an operator for future action(s) in the external environment. Any novel or unfamiliar real-time episodic event perceived by the agent "returns a result to [the agent's] supergoal" (Laird, 2012, p. 8). The agent's actions and interpretations from this event are immediately "chunked" into a production rule. This chunked rule contains new conditional statements that relate directly to the episodic experience just witnessed and interpreted by the agent. Just like Soar's operators and states, this rule-chunk is stored associatively between the agent's short-term episodic memory and its long-term procedural memory banks for cross-retrieval. These episodic "interpretations" are based on a priori or previously acquired knowledge relating to apparently similar situations. For future interaction scenarios where a similar episodic event matches some of the conditions in that rule-chunk, the agent can retrieve suggested preferences for the most appropriate operators or actions to use in that new situation. Over time, Soar's reinforcement learning feature "[…] tunes the actions of rules that create preferences for operator selection" by rewarding the agent for acting out "good" preferred operators and actions (Laird, 2012, p. 18).

Soar also employs "automatic subgoaling" (an extension of the chunking mechanism). After recognizing a conflict (impasse) during a novel decision-making scenario, the Soar-agent generates a new state known as a "subgoal". As this new state is based on rule-based knowledge, Soar's freshly generated subgoal can oscillate between superstate and substate designations. Soar's corporate literature claims that other Rule Based Systems (RBS) can only deal with one fixed state at a time. The Soar agent automatically generates monitoring substates that "decompose complex operators into similar operators" (Laird, 2012, p. 66). Therefore, Soar's assignment of a monitoring role to these substates implies that Soar also offers rudimentary on-demand metareasoning or self-reflection capabilities (Laird, 2012, p. 66).

Soar's characteristic feature is its decision cycle. Soar's decision cycle is a "fixed processing mechanism that works in five phases: input, elaboration, decision, application, and output" (Lehman, Laird, & Rosenbloom, 2006, p.17). The Soar-agent must be able to parse continuous percepts (e.g. observations, perceptions) from a stochastic environment in order to make informed real-time decisions on which operators to select next. The Soar-agent handles this perpetually updating percept stream of new symbols by temporarily storing them first in a local "perceptual short-term memory" before adding them to a more global and action-driven working memory (Laird, 2012, p. 18). Next, the Soar-agent reflects on recently occurred episodes stored in its short-term working memory by elaborating on the elements of these events as conditions and suggested actions.

The Soar-agent elaborates on these learned episodes by comparing the history of its operator-preferences (thoughts) and operations (actions) it made at the time with conditional statements it had acquired from the rules it had chunked and stored from previous experience into its long-term memory. The Soar-agent

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

then has the opportunity to decide on a course of action. If any rules generated (learned) from a recent episode match with any other remembered rules, those matching rules will fire in parallel. Once these matching combination of rules fire, the Soar-agent will have updated to its next state which contains new "suggestions, or preferences, for selecting the current operator" (Lehman, Laird, Rosenbloom, 2006, p.17). From a list of highly rated and preferred operators, the Soar-agent can then propose, select, and apply the operator to be used for the next action. The programmer also has the luxury of manually preferring operators by adding a "+" symbol to a favored operator. With the reinforcement learning procedure, the Soar-agent updates an operator with a "+" symbol and a corresponding numerical rating value automatically when rewarded for an action that was produced by that operator.

Soar's architecture is known mostly for its ability to simultaneously fire an extremely large number of hard-coded production rules without having these rules initially undergo a matching procedure. Soar's operators and states are also rule-driven. In fact, Soar differentiates itself from other Rule Based Systems (RBS) in that it is most concerned with selecting the best operator that will impact the agent's next state. In other RBS, the focus is on hierarchically rating each rule and sequentially determining which single rule should fire next (Laird, 2012, p. 4). Soar instead determines the best combination of rules to be fired in parallel, given the "knowledge and structure'" of a particular task-context (Laird, 2012, p. 4-5). These operators, states, and rules are completely modular atomic elements and can be recycled from one memorized context to another for action-selection. This means that Soar saves on rule-firing speed and resolves any conflicting logical contradiction or decision impasse by being highly associative with different elements of each of its parallel long and short term memory banks. However, this convenient association and modularity among memorized rules comes at a structural cost. For example, Soar can retrieve any other chunk (concept with data and a form) from one memory bank and automatically turn it into a sub-goal for use in episodic context. This cross-retrieval process is executed without requiring any knowledge of the previous associative context. Consequently, Soar is task-agnostic as it only focuses on how the parallel optimization of a large number of hard-coded rules result in operational changes "[…] to the features and values" in the agent's state (Lehman, Laird, & Rosenbloom, 2006, p. 17).

Soar's Technical Integration with Virtual Agents

This chapter is focused on historical virtual agent implementation examples rather than on their technical interfacing protocols between agent, cognitive architecture and virtual environment. However, it can be stated in passing that generally speaking, integration between Soar and character-driven virtual environments are contingent on the technical affordances of a particular software application. For example, NPCs residing in gaming engines such as Unreal Tournament can be interfaced with Soar using a control methodology and data specifications that are packaged by the company SoarTech as a "Human Behavior Model Interface Standard" (HBMIS, van Lent et al., 2004). It is entirely feasible that this interface standard can be migrated across similar gaming engines such as Unity. The most established gaming engine implementation treats each human behavioral model as an asynchronous external software module that (for now) communicates with a proprietary "UnRealScript" function interface (van Lent et al., 2004, p. 3). Each model exclusively processes Unreal's simulation data through the Soar General Input/Output mechanism (SGIO, Laird et al., 2002) before passing this data through its own decision cycle (van Lent et al., 2004, p. 3). Fortunately, these models are designed so that their processing cycles do not interfere with the game engine's own processing cycles (van Lent et al., 2004, p. 4).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

SoarTech has made further high-level modifications to this Unreal/SGIO interface to account for state initialization and representation, (multiple) independent procedural threading, and group action sequences (van Lent et al., 2004, p. 3-4). For data control, SoarTech has also created their own low-latency DLL file that is executed once the Unreal engine initializes (van Lent et al., 2004, p. 4). In order to account for Unreal's real-time rendering capabilities, the SGIO mechanism must be able to execute and return game loops at a rapid speed. In fact, deliberative latency might even compromise the virtual environment's frame-refresh rate (i.e. below 30 frames per second, van Lent et al., 2004, p. 4). Therefore, Soar's once impressive ability to refresh a new decision cycle every 50 milliseconds might still prove to be too temporally cumbersome for cognition to believably function at the speed of contemporary game-play (van Lent et al., 2004, p. 4). One of the agent designers mentioned in this chapter has developed a plausible workaround to this interface issue. In this case, the high-level interface virtually encapsulates three selectable lower-level interfaces that operate at run-time instead of real-time (Magero et al., 2004, p. 879). Magerko's research team has optimised the performance capabilities of SGIO by "embedding [...] multiple copies [of] Soar within the UT process using a C API" (Magerko et al., 2004, p. 879).

Gaming engines in general (esp. VBS2, Unity3D and Unreal Tournament) can now be interfaced with Soar using SoarTech's "PoL" ("Patterns of Life") behavior modeling infrastructure (SoarTech, "Entity-Level Patterns of Life for Virtual Environments, promotional brochure, 2013). This PoL infrastructure comes bundled with sample virtual agents for evaluation. These agents feature parameterized and schedulable patterns that correspond to self-monitoring behaviors, beliefs, desires and intentions (SoarTech, "Entity-Level Patterns of Life for Virtual Environments, promotional brochure, 2013).

Outside of the gaming context, other Soar-agent designers have created novel integration solutions such as implementing an elaborate messaging interface that links various technical components and software application components. These disparate software components can even run on different machines, if required. For example, the proprietary components that created the cognition for the first virtual agent example (STEVE, Rickel & Johnson, 2000, p. 11) passed messages between the VIVIDS simulation engine (USC Behavioral Technologies Laboratory, 1997), Vista Viewer (for visual interfacing, Lockheed Martin, 1995), and Entropic's Text-to-Speech software. These messages were dispatched using another piece of proprietary software known as, ToolTalk (SunSoft, 1994).

Soar Examples

These four Soar implementation examples were chosen to showcase the various ways a rule-based cognitive architecture can enhance the interactive and deliberative potential of a virtual agent. These examples are presented chronologically in historical order from 2000-2013. These Soar-agent examples address military (e.g. industrial and cultural sensitivity training), entertainment (e.g. video-games), and fine-arts (e.g. dance) contexts.

The Soar Training Expert for Virtual Environments (STEVE)

The first example involves a US Navy-funded "Soar Training (Virtual) Expert for Virtual Environments" known colloquially by "his" personified acronym, "STEVE" (Rickel & Johnson, 2000). STEVE is a virtual humanoid who interacts with naval officers wearing a head-mounted display (HMD) in a virtual reality (VR) environment (Rickel & Johnson, 2000, p. 101-105). STEVE can respond to text-to-speech queries as well as reason about the motivations and goals for enacting the correct operational procedures

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

for maintaining has turbine engines on naval ships[9] (Rickel & Johnson, 2000, p. 96). In order to carry out these training procedures in a believable manner, STEVE can orient his avatar body, walk around, and point towards salient virtual objects within the turbine-engine simulation while identifying each object by name (Rickel & Johnson, 2000, p. 97). Further, STEVE can shift his gaze from the inquiring officer to the queried virtual objects that he is currently identifying for pedagogical purposes.

STEVE uses a perception module to monitor as well as semantically interpret all of the appropriate messages and events that result in updating the (virtual) world-state (Rickel & Johnson, 2000, p. 106). STEVE constantly relays the current world-state in a way that is continuously interpretable by the STEVE's cognition module (ie. where Soar resides) (Rickel & Johnson, 2000, p. 106). This cognition module is then responsible for deciding what STEVE's next action-sequence should be. Once one decision cycle has transpired, STEVE transmits motor-commands to the motor control module which allows STEVE's voice to speak and avatar body to move (Rickel & Johnson, 2000, p. 107). STEVE's cognition module carries out serial procedures independently from the other modules while the perception and motor control modules work in parallel (Rickel & Johnson, 2000, p. 107). STEVE is unique from many other virtual agent implementations because Soar in this case, does not supply STEVE's virtual mind with higher-level cognitive processes (e.g. cerebral deliberation and meta-cognition). By contrast, Soar represents STEVE's lowest level of cognition (Rickel & Johnson, 2000, p. 106-107). Soar was only modeled for lower-level cognition because the original Soar architecture was never optimized to assist with those demonstration and conversational activities more familiar to the perceptual and actuational (motor-control) modes of operation (Rickel & Johnson, 2000, p. 106-107)

Figure 1. STEVE giving a demonstration of the turbine engine simulator
Permission granted by the author (Rickel & Johnson, 2000)



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

STEVE repeatedly deployed his cognition module to choose the next appropriate action (Rickel & Johnson, 2000, p. 114). STEVE engaged the cognitive module by perpetually looping Soar's trademark decision cycle. Based on the hardware specifications circa 2000, STEVE used to complete up to ten decision cycles per second (Rickel & Johnson, 2000, p. 106-107). Even back then, STEVE appeared to be very responsive to his environment (Rickel & Johnson, 2000, p. 115). From each percept received by the perception module, STEVE responds to a trainee's question or demonstrates a new procedure by executing a behavioral operator. These operators are represented by production rules that effect a change in the world-state and might even influence the next sequence of proposed operators themselves (Rickel & Johnson, 2000, p. 116). Once an operator has been applied, STEVE then determines which preferred operator to select and apply for the next time-step. These operators conform to action-categories related to: object manipulation, object-movement, self-proximity, and condition checking while the actions themselves correspond with STEVE's communicative faculty (Rickel & Johnson, 2000, p. 116). STEVE hierarchically compares, prefers, selects and then applies operators based on their contextual priority. STEVE's first priority is usually to answer the trainee's queries. However, if there are no queries, STEVE might instead attempt to steer the conversation towards focusing on more of his own verbal explanations and demonstrations of the training course (Rickel & Johnson, 2000, p. 114-115). In other cases, production rules might prompt STEVE to move to the next appropriate demonstration location or compel him to point at a particular object of immediate importance (Rickel & Johnson, 2000, p. 114-115). STEVE's communicative acts are in the form of Soar production-rules but ultimately have the restricted interactive characteristics of finite-state machines (Rickel & Johnson, 2000, p. 115).

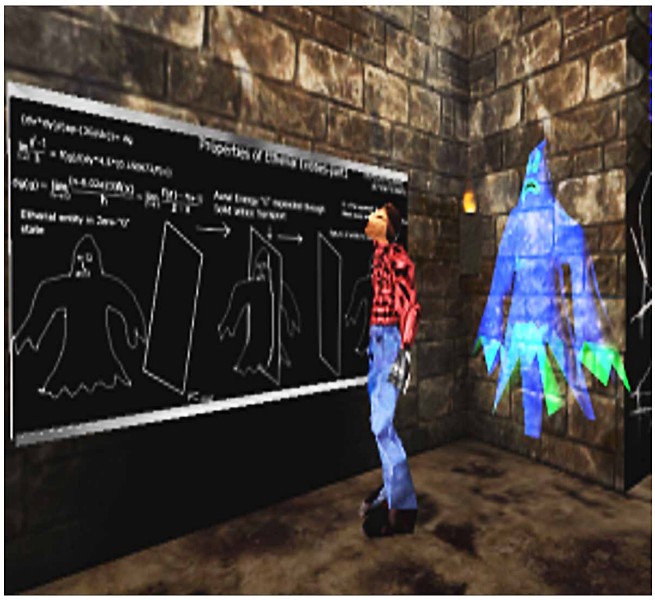
Soar-NPCs in Unreal Tournament

These virtual Soar-agents are NPC (Non-Player Character) humans that use the Soar General Input/Output (SGIO, Laird et al., 2012) interface in order to communicate with Soar's architecture and the Unreal custom video-game environment (Magerko et al., 2004, p. 879). In this video game known as "Haunt 2", the player is the ghost and the player's goal is to uncover the mystery behind his/her death by subtly haunting the rule-driven NPCs using a limited ethereal mode of non-verbal communication. The player also cannot directly manipulate objects in the game and so, can only psychologically influence the haunted human NPCs to manipulate objects on behalf of the ghost's otherworldly will (Magerko et al., 2004, p. 878). The player-as-ghost can either scare the NPC into acting on the ghost's behalf or temporarily possess a more psychologically receptive NPC. When an NPC is in a possessed state, the ghost-player can actually observe and directly manipulate the NPC's Soar-mediated cognitive processes as they occur (Magerko et al., 2004, p. 878). However, each NPC has its own idiosyncratic psychological tolerance, bias, goals, motivations, needs (e.g. comfort), drives (e.g. hunger, thirst etc.), and a unique knowledge-representation of the background context (Magerko et al., 2004, p. 879).

Plot points in Haunt 2 are divided into preconditions and postconditions that can be optimally mediated by Soar operators to address both general and specific narrative outcomes. Preconditions describe "[…] what has to be in true in the world in order for this plot point to be relevant". In contrast, postconditions describe "[…] what actions should take place once these conditions are met" (Magerko et al., 2004, p. 880-881). In Haunt 2, there is no explicit mention of in-conditions which would describe what should be continually happening in the world to retain the truth of that particular condition. The closest thing to in-conditions in Haunt 2 is a "timing constraint" that provides situations when certain plot points lose their relevance over time (Magerko et al., 2004, p. 881). In addition to the player-character (PC)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. A screenshot of the player haunting Soar-NPCs as a ghost in the video-game HAUNT 2 Permission for this image granted by the author. Accessed online from http://ai.eecs.umich.edu/~Soar/sitemaker/projects/haunt/images/blackboard.jpg on March 16, 2015
(Magerko et al,2004).



and NPC, Haunt 2 has a "Director" that runs a simulation modeling the player's actions in order to hypothesize over "[…] what declarative knowledge [the ghost] is acquiring through his [event triggered or time-lapsed] experiences [by keeping] an up-to-date internal structure describing the state of the world" (Magerko et al., 2004, p. 882). For example, the Director can alter the NPC's communication style (e.g. to be more outwardly "social"), and/or control the NPC's motor-functions and available text-strings for dialogue with other NPCs (Magerko et al., 2004, p. 880). The Director determines the success or failure of a world-state hypothesis depending on whether the player advances the story's plot points towards the expected narrative conclusion or instead, thwarts the activation of a plot-point by ignoring some particular pre-condition (Magerko et al., 2004, p. 880).

Virtual Soar-Agents using the Cultural Cognitive Architecture (CCA)

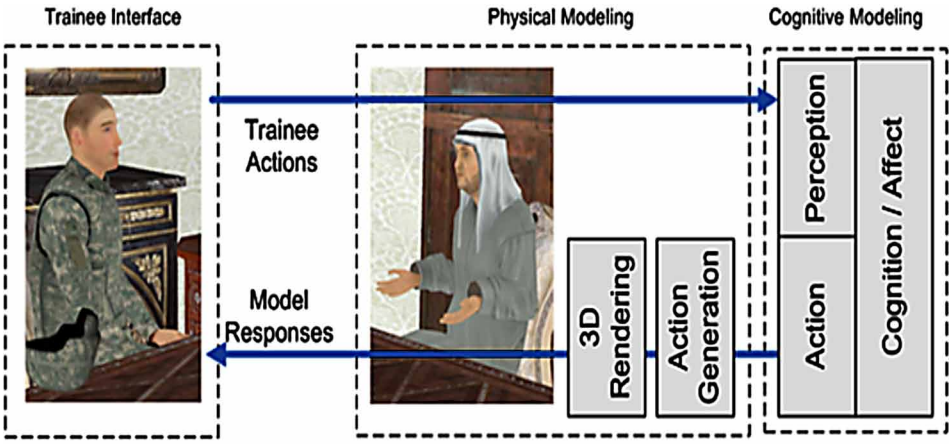
Some Virtual Soar-agents have been designed to interpret and communicate the complex ethnographic nuances of behavioral protocols, values, and social norms from other cultures. The military contractors known as "SoarTech" ("SoarTech", n.d.) have designed a Cultural Cognitive Architecture (CCA) within SOAR[10] in order to cognitively enhance conversational virtual agents for cross-cultural training purposes (Taylor & Sims, 2009, p.1). Soar in particular was chosen for its: "scalable long-term memory, […] associative memory activation [functions], mixed automatic and deliberative reasoning [capabilities], goal-directed, problem-solving behavior, ability to incorporate multiple reasoning methods, […], and graph-based working memory, which maps well to Scripts/Frames approaches to schema representation" (Taylor et al., 2007, p. 2; Minsky, 1975; Schank & Abelson, 1975).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Built upon the Cultural Schema Theory (CST), the CCA draws from an ontology that contains a growing corpus of embodied ethnographic knowledge (Taylor et al., 2007, p. 4). Outside of virtual environments, this ethnographic knowledge is usually acquired when human agents are embedded within a particular culture for a sufficient length of time (Taylor & Sims, 2009, p. 1). The cultural schema itself encodes the social hierarchies, known protocols, predicted social outcomes, culturally salient features of the environment as well as the particular contextually-sensitive goals that are likely to drive characters from particular ethnographic backgrounds (Taylor & Sims, 2009, p. 5). A culturally instantiated schema is subject to "[…] a process of activation, decay, and deactivation based on relevance to events perceived in the environment" (Taylor et al., 2007, p. 4). SoarTech claims that the CCA and CST are generalizable enough to be applied to understanding various cultures[11]. If any cultural differences cannot be captured with generality, the CCA generates leaf nodes to represent any noticeable cultural idiosyncrasies (Taylor & Sims, 2009, p. 5).

The Soar-CCA implements its virtual agents by balancing an emotional system dealing with an event-dependent appraisal mechanism with a cognitive system that establishes higher-level goals and generated behaviors (Taylor & Sims, 2009, p. 5-6). These conversational virtual agents exist in the 3D animation engine and behavioral simulator, "Vcommunicator". Vcommunicator has access to "[…] over 60 culturally diverse virtual human models as well as 40 facial expressions and 500 gestures, and can automatically lip-sync to over 22 mouth shapes that map to over 100 speech sounds of the International Phonetic Alphabet." (Taylor & Sims, 2009, p. 4). A culturally successful agent will be able to ascertain the relevance of goals, and the benchmark set for societal norms while assessing the relative novelty of a particular ethnographic ritual (Taylor & Sims, 2009, p. 4). SoarTech claims that a virtual agent drawing from the Soar-CCA can dynamically assess the most typical emotional state (e.g. relative standards of "rudeness" and "politeness") based on the episodic sum of the outputs from the emotional system and the appraisal value collected from a specific cultural encounter (Taylor & Sims, 2009, p. 4). Overall, each virtual character receives both the recommended action-sequence and the expected emotional state while having parameterized control over situations where the relative emphasis of certain culturally appropriate gestures is required (Taylor & Sims, p. 2009, p. 7).

Figure 3. This image depicts a virtual soldier attempting cross-cultural negotiations with a Middle-Eastern Soar-CCA Virtual Agent. Permission to use this image granted by the author (Taylor & Sims, 2009, SoarTech).



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The SoarTech team has conceded to some challenges regarding virtual character implementations of the CCA. For one, it was difficult for the team to determine the ideal granularity for cognitive and physical actions. Dealing with the granularity at the atomic level might be well-tailored to very specific cultural instances but would likely not be generalizable enough to be interoperable with multiple cultural contexts. More generalized models are easier to engineer. However, if the models are too generalized, then this would create a culture-agnostic character which might not be seen as believable or realistic (Taylor & Sims, 2009, p. 9).

In theory, but not entirely realized in practice, the CCA-agents are potentially scalable to handle rich data models and a fluent understanding of various natural languages. However, with increasing complexity and size comes an equal increase in available computational resources. Within the virtual world context, this poses no real obstacle because the "[…] fidelity of the data would be only as good as the fidelity of the interactions available in the virtual world" and most virtual world data is limited to culturally unrealistic gestures and inauthentic chat-responses anyway (Taylor & Sims, 2009, p. 9). To date, however, there does not appear to be any firm evidence of Soar-agents (including CCA-agents and any other AI-enhanced agents) that have sufficiently transcended the limitations of finite-state organization to handle natural language understanding and generation at an adult human level. At the very least, some noteworthy attempts (see Lindes et al., 2015) have been made to use Soar for grounded language acquisition and linguistic mapping. In one particular case, a Soar-agent can "[…] associate the sentence 'put a large red block in the table' with an instantiated operator which will achieve the intended goal" (Mohan et al., 2012., p. 124).

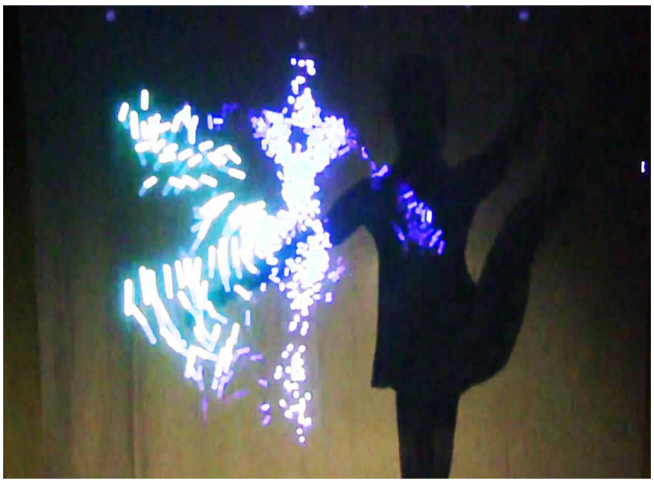
CAA-agents are similar to the Director in Haunt 2 as it can select goals that will encourage following the currently activated narrative script towards its ritualistic conclusion (Taylor et al., 2007, p. 4). Soar's impasse threshold-breaking mechanism is mirrored in CAA's ability to prefer, apply, and fire those current script-relevant operators and rules from working memory that already have the highest activation potential (Taylor et al., 2007, p. 3-4). The issue here is that the CAA-agent will always act based on what is determined to be the most culturally appropriate action rather than in its own self-interest (Taylor et al., 2007, p. 3-4).

A Virtual Soar-Agent using Augmented Reality for Improvised Choreography

VAI (Viewpoints Artificial Intelligence, Jacob et al., 2013) is an augmented reality virtual Soar-agent designed for interactive improvised choreography in real-time. VAI is projected onto a 2D surface in order to interact with human-participants in a turn-taking manner (Jacob et al.,, 2013, p. 4). VAI is influenced by the collision physics of the human participant's shadow and can learn movement sequences that are scanned from a human's skeletal vertex points acquired via a Microsoft Kinect system ("Kinect for Windows", n.d.). Cognitively, VAI accesses "perception, reasoning and action modules" in order to calculate relative vertex values of: joint position, tempo (pacing), energy[12] levels, and velocity. These values are conveniently represented for Soar as "predicates" (Jacob et al., 2013). These Soar-ready predicates convert VAI's ambient movements into color values when gestures are being made at low-acceleration rates (Jacob et al., 2013, p. 4). VAI stores all input and output gestures into an "internal interaction history" (Jacob et al., 2013, p. 4) which likely encompasses various stages of Soar's working memory. VAI either randomly accesses these working memory elements – composed of various dance-sequence durations - in order to determine the best movement operator to propose, select, and apply, or ignores the interaction history entirely (Jacob et al., 2013, p. 4). In response, VAI can choose between refusing

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. A snapshot of VAI taking turns improvising a dance routine with a human participant. Permission for this image granted by the author
(Jacob et al., 2013)



to apply an operator, repeat the input gesture (i.e. mimicking the dancing human participant), or perform some novel choreographic variation on said gesture. VAI's creators have asserted their desire to enable the possibility for VAI to propose and apply operators based on instances of similarity based reasoning principles (i.e. the similarity between "source and target […] predicates", Jacob et al., 2013, p. 5). VAI also has a heuristic for showing an awareness of the human participant's presence and dance-movement choices while also showing the capacity for improvisation and creative autonomy. This heuristic allows VAI to follow a human participant's repeated gesture twice (consecutively) before performing an improvised gestural variation of its own devising (Jacob et al., 2013, p. 5).

VAI's limitations were more the result of its perceived levels of interaction and not so much due to any cognition parsing issues produced by Soar. For example, some people were too impatient with the turn-taking approach to VAI's choreographed interactions (Jacob et al., 2013. p. 4-5). Otherwise, some human participants either found VAI's interactions to be arbitrary (random), or merely copying the human movements. In fact, some human participants were deluded into thinking the VAI agent possessed more deliberative capabilities than were actually present under-the-hood (Jacob et al., 2013, p. 4-5). The latter can be seen as beneficial for those designers wishing to elicit conceivable rather than accessible interactive states in the mind of the interactor.[13]

Criticism of Soar-Implementations

Current versions of Soar – although centered on rules – work best with a robust knowledge base containing a large number of hand-coded productions. Ultimately, Soar works best for purely deterministic simulated environments but has been criticized for not being able to solve even rudimentary logic problems based on hidden, partially observable or incomplete empirical knowledge (Bringsjord et al., 2005, p. 35). For many virtual agent implementations, cognitive architectures such as Soar do not usually require full scale logic operations. This is because they are usually "[…] too slow for real-time perception-and-action-heavy computation" (Bringsjord et al., 2005, p. 35) anyway. Soar might be scalable while keep-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ing the computational cost relatively low but eventually, Soar's cognition lags when a large number of production-rules need to successfully interpret a dynamically changing virtual environment. Basically, Soar's chunking mechanism can optimize limited computational resources but a stochastic simulation modeled after the "real world" will cause the automatic-subgoaling mechanism to gradually tax the Soar's working and long-term memory capabilities.

The general issue of abstracting out semantically-appropriate information from a stochastic virtual environment and "specifying the [symbolically represented] conditions under which productions can apply" (Best et al., 2006, p. 212) provides ongoing engineering challenges to those particular cognitive architectures that are production rule-driven (esp. Soar). At best, a production-oriented cognitive architecture can cope with generalized symbolic approximations of environmental conditions through some rudimentary form of case-based reasoning process. In this instance, a cognitive agent can focus on a single "canonical case" that can be applied to most particular interactive scenarios where conditional decision making applies (Best et al., 2005, p. 213). Culturally speaking, cognitive elements representing a "canonical case" – and its associated natural language structures - might not even be sufficiently decomposable into production rules and rule-chunks.

In virtual environments optimized for multi-agent system interactions, cognitive architectures are usually under-represented as they are criticized for neglecting multi-agent implementations. This is because cognitive architectures were originally to focus on the egocentric cognitive concerns of a single agent instead of allocentrically mapping the overall cognitive climate of interacting agents. According to Smart et al., developers of these multi-agent swarms, usually "[...] assume that individual agents have very rudimentary cognitive processing capabilities, and in some cases, aspects of agent cognition (e.g., beliefs or attitudes) are represented by simple numerical values that do not respect the constraints imposed by the human cognitive system" (Smart et al. 2014a., p. 1).

Speculations for Future Virtual Soar-Agent Implementations

Overall, the above examples are diverse enough from each other to address the range of Soar's potential and limitations when cognitively interfacing with a virtual agent. It is up to each of the above researchers whether or not they wish to re-iterate their virtual agents with newer Soar-related features. Regardless, there are some plausible research directions these and other researchers could take to develop the next-generation virtual Soar-agents. For example, STEVE could utilize Soar's reinforcement learning capability to emergently discover how to train naval officers with equipment operating in unknown or hostile virtual simulations (e.g. a war-zone). In addition, the NPCs in Haunt 2 could be modified to explicitly model separate "in-conditions" from pre and post conditions. Further, CCA-agents could do with a connectionist representation in future versions of Soar – similar to CLARION - where cultural sensitivities can be gradually learned in a supervised manner using tunable weights and/or with reinforcement signals. Perhaps, a virtual-agent can be incrementally rewarded and/or have neural connections solidified (i.e. Hebbian learning activity increased) for producing behaviors (as well as internal thoughts) that seem to match the protocols of other cultures (e.g. scripts for culturally "correct" postures, semantic relationships, and routines). This would be a compelling alternative to classically tuning Soar's rules and operators themselves. Ideally, the issue of a virtual agent's "self-interest" situated within an uncertain cultural context would be more elaborately addressed if Soar was modified to draw from a software ontology (Noy et al., 2001) that contained recursive predicates and slots that dealt with "self-awareness". Monitoring substates is sufficient for finding common ground between complex operators (Laird, 2012.,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

p. 66). However, operators themselves are not necessarily a tractable representation of what one culturally would consider to be a form of self-awareness. As for VAI, it may prove difficult to have Soar's decision cycle operate in a non-serial manner. Turn-taking, however, was an issue for some impatient human participants so VAI's creators might consider either advocating more parallelism in Soar and/or implementing VAI with a different cognitive architecture more suited to parallel decision making.[14]

Generally speaking Soar's current endeavor to develop additional affective-state models and mental imagery (i.e. imagination) representations is encouraging. Soar will also likely be adopted by a newer generation of virtual agent designers once a distinct prospective memory store is finally implemented in future iterations. Imagination resides in episodic memory[15] as a list of proposed (but not preferred) operators, rules, and actions. These proposed rules, operators and, actions are the atoms of imagination in Soar. These atoms can be further elaborated on via fictional metadata stored in semantic memory, or as plausible (but untested) interactivity algorithms, and/or heuristics stored in procedural memory. Alongside speculations of a prospective memory Soar implementation, a parallel sub-component of episodic memory known as "elaborative memory" or "fantasy memory"[16] would generate imagined novel episodes as myths/fictions. These episodes would propose equally fictional operators, rules, and actions.[17] The elaborative memory would be driven by sub-goals and sub-states. Elaborative memory would determine which sub-goals/sub-states should be rapidly retrieved from other memory banks. Chunking would involve elaboration by fabricating memories, operators, rules etc. The rule firings themselves could even resemble mirror neurons as they could be simulated internally without actually firing to express cognitive deliberation (Maffre et al., 2001).

Prospective memory on its own would definitely benefit those aforementioned virtual Soar agents that navigate semi-deterministic virtual spaces. The possibility of Soar including a prospective memory as one of its features would be even more impressive if the virtual Soar-agent interface this type of memory with a robust logic component – autonomous from the authority of production rules. Further, a future version of Soar could have its own equivalent representation of CLARION's *rule-extraction refinement algorithm"* (aka. RER, Sun 2006., p. 86). Soar could approximate how CLARION actually connects declarative and procedural knowledge. Like CLARION, Soar's own version of the RER algorithm could be encoded as an explicit knowledge element for use by CLARION's deductive and inductive logic components in the service of actionable inference. In other words, Soar can understand in its own way how CLARION's showcased RER algorithm has enabled a CLARION-agent to learn from sub-symbolic knowledge and find ways to convert this knowledge into logic and/or ontology symbols. If the implicit knowledge component is sufficiently salient and capable of semantic refinement, base empirical experience can even be converted into a Soar production rule. Prospective memory combined with logical inference and a new representation of the RER algorithm may even be able to predict future world-states in a stochastic environment with some actionable degree of confidence.

CONCLUSION

In conclusion, this chapter has showcased the various ways in which Soar can be integrated with virtual agents representing multi-faceted roles such as: military trainers, video-game NPCs, culturally sensitive interpreters, and improvised choreographers. This chapter has discussed the relevance for using a cognitive architecture in the first place and has contextualized the strengths and weaknesses of implementing virtual agents with Soar.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Soar's legacy remains as an established scalable and stable production-rule architecture still maintained by an active development team seeking further iteration and improvement to its fundamental design. The speculations mentioned in this chapter merely serve to inspire cognitive architects to push the boundaries of interactions driven by mental-processes even farther than anticipated.

REFERENCES

Anderson, J. R. (1996). ACT: A simple theory of complex cognition. *The American Psychologist*, *51*(4), 355–365. doi:10.1037/0003-066X.51.4.355

Anonymous. (n.d.). *CLARION Cognitive Architecture.* Retrieved February 27, 2015, from https://sites.google.com/site/clarioncognitivearchitecture

Anonymous. (n.d.). *Kinect for Windows.* Retrieved March 15, 2015, from http://www.microsoft.com/en-us/kinectforwindows/

Anonymous. (n.d.). *LSL access portal.* Retrieved March 16, 2015, from http://wiki.secondlife.com/wiki/LSL_Portal

Anonymous. (n.d.). *Second Life. LSL access portal.* Retrieved March 16, 2015, from http://secondlife.com

Anonymous. (n.d.). *Soar Home Page.* Retrieved February 27, 2015, from http://Soar.eecs.umich.edu/

Anonymous. (n.d.). *SoarTech Corporate Brochure*. Retrieved February 28, 2015, from http://ai.eecs.umich.edu/Soar/sitemaker/docs/misc/SoarRBSComparison.pdf

Anonymous. (n.d.). *SoarTech* webpage. Retrieved March 15, 2015, from http://www.Soartech.com/about/

Best, B. J., & Lebiere, C. (2006). Cognitive agents interacting in real and virtual worlds. In R. Sun (Ed.), Cognition and Multi-Agent Interaction: From Cognitive Modeling to Social Interaction. New York: Cambridge University Press, 2006.

Bringsjord, S., Khemlani, S., Arkoudas, K., McEvoy, C., Destefano, M., & Daigle, M. (2005). Advanced synthetic characters, evil, and E. *Game-On*, *6*, 31–39.

Brooks, R. (1991). Intelligence Without Representation. *Artificial Intelligence*, *47*(1-3), 139–159. doi:10.1016/0004-3702(91)90053-M

Cox, M. T., & Raja, A. (2011). Metareasoning: An Introduction. In M. T. Cox & A. Rajam (Eds.), *Metareasoning: Thinking About Thinking* (pp. 3–14). Cambridge, MA: MIT Press. doi:10.7551/mitpress/9780262014809.003.0001

Crawford, C. (2003). *The Art of Interactive Design: A euphonius and illuminating guide to building successful software* (pp. 77–90). San Francisco: No Starch Press.

Isbister, K. (2006). *Better Game Characters By Design: A Psychological Approach.* San Francisco: Elsevier.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Jacob, M., Coisne, G., Gupta, A., Sysoev, I., Verma, G., & Magerko, B. (2013). Viewpoints AI. *Proceedings of the Ninth Annual AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*.

Laird, J. (2012). *The Soar Cognitive Architecture*. Cambridge, MA: MIT Press.

Laird, J., Hucka, M., Huffman, S., & Rosenbloom, P. (1991). An analysis of Soar as an integrated architecture. *ACM SIGART Bulletin*, *2*(4), 98–103. doi:10.1145/122344.122364

Laird, J. E., Assanie, M., Bachelor, B., Benninghoff, N., Enam, S., Jones, B., . . . Wallace, S. (2002). A Test Bed for Developing Intelligent Synthetic Characters. *AAAI 2002 Spring Symposium Series: Artificial Intelligence and Interactive Entertainment*.

Laird, J. E., Kinkade, K. R., Mohan, S., & Xu, J. Z. (2012). Cognitive robotics using the Soar cognitive architecture. *Cognitive Robotics AAAI Technical Report*, *WS-12*, 6.

Laird, J. E., & Rosenbloom, P. S. (1993). Integrating execution, planning, and learning in Soar for external environments. Academic Press.

Laird, R., Newell, A., & Rosenbloom, P. S. (1987). Soar: An Architecture for General Intelligence. *Artificial Intelligence*, *33*(1), 1–64. doi:10.1016/0004-3702(87)90050-6

Lathrop, S. D., & Laird, J. E. (2009, March). Extending cognitive architectures with mental imagery. *Proceedings of the second conference on artificial general intelligence*.

Lehman, J. F., Laird, J. E., & Rosenbloom, P. S. (1996). A gentle introduction to Soar, an architecture for human cognition. *Invitation to Cognitive Science, 4,* 212-249.

Leonard, A. (1997). *Bots: The Origin of New Species*. San Francisco: HardWired.

Li, J., & Laird, J. E. (2013). The Computational Problem of Prospective Memory Retrieval. *Proceedings of the 17th International Conference on Cognitive Modeling*.

Lim M.Y., Dias, J., Aylett, R., & Paiva, A. (2010). *Creating Adaptive Affective Autonomous NPCs. Autonomous Agent Multi-Agent Systems.* Doi:10.100710458-010-9161-2

Maffre, E., Tisseau, J., & Parenthoen, M. (2001). Virtual Agent's Self-Perception in Story Telling. In *Virtual Storytelling: Using Virtual Reality Technologies for Storytelling.* Berlin: Springer-Verlag.

Magerko, B., Laird, J., Assanie, M., Kerfoot, A., & Stokes, D. (2004). AI characters and directors for interactive computer games. In *IAAI'04 Proceedings of the 16th conference on Innovative applications of artifical intelligence*.

Mateas, M. (1997). *Computational Subjectivity in Virtual World Avatars.* AAAI Technical Report FS-97-02, 1-6.

Minsky, M. (1975). A Framework for Representing Knowledge. In The Psychology of Computer Vision (pp. 211-277). New York: McGraw Hill.

Moon, J., & Anderson, J. R. (2012). Modeling Millisecond Time Interval Estimation in Space Fortress Game. *34th Annual Conference of the Cognitive Science Society*, Sapporo, Japan.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Noy, N. F., & McGuinness, D. L. (2001). *Ontology Development 101: A Guide to Creating Your First Ontology.* Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001.

Reidl, M. O., & Stern, A. (2006). Failing Believably: Toward Drama Management with Autonomous Actors in Interactive Narratives. In S. Göbel, R. Malkewitz, & I. Iurgel (Eds.), *TIDSE 2006, LNCS 4326* (pp. 195–206). Berlin: Springer-Verlag. doi:10.1007/11944577_21

Rickel, J., & Johnson, L. W. (2000). Task-Oriented Collaboration with Embodied Agents in Virtual Worlds. In J. Cassell, J. Sullivan, & S. Prevost (Eds.), *Embodied Conversational Agents* (pp. 95–122). Cambridge, MA: MIT Press.

Ritter, F., Kim, J. W., & Sanford, J. P. (2012). *Soar Frequently Asked Questions List*. Retrieved January 30, 2016, from http://acs.ist.psu.edu/projects/soar-faq/soar-faq.html#G3

Schank, R. C., & Abelson, R. P. (1975). *Scripts, plans, and knowledge*. Yale University.

Smart, P. R., Richardson, D. P., Sycara, K., & Tang, Y. (2014a). Towards a cognitively realistic computational model of team problem solving using ACT-R agents and the ELICIT experimentation framework. 19th ICCRTS C2 Agility: Lessons Learned from Research and Operations. Southhampton University School of Electronics and Computer Science.

Smart, P. R., & Sycara, K. (2015, July) Situating Cognition within the Virtual World. *6th International Conference on Applied Human Factors and Ergonomics*, Las Vegas, NV.

Smart, P. R., Sycara, K., & Lebiere, C. (2014b, September). Cognitive Architectures and Virtual Worlds: Integrating ACT-R with the XNA Framework. *Annual Fall Meeting of the International Technology Alliance*, Cardiff, UK.

Sun, R. (2006). The CLARION cognitive architecture: Extending cognitive modeling to social simulation. In Cognition and multi-agent interaction, (pp. 79-99). New York, NY: Cambridge University Press.

Taylor, G., Quist, M., Furtwangler, S., & Knudsen, K. (2007, May). Toward a hybrid cultural cognitive architecture. *CogSci Workshop on Culture and Cognition.* SOARTech.

Taylor, G., & Sims, E. (2009). Developing Believable Interactive Cultural Characters for Cross-Cultural Training. In Online Communities and Social Computing (pp. 282-291). Springer Berlin Heidelberg.

van Lent, M., McAlinden, R., Probst, P., Silverman, B. G., O'Brien, K., & Cornwell, J. (2004). Enhancing the behaviorial fidelity of synthetic entities with human behavior models. Departmental Papers (ESE), 300.

ADDITIONAL READING

Anstey, J., & Pape, D. (2002, October). Scripting the Interactor: An Approach to VR Drama. *IC&C '02, October 16-26, 2002. ACM 1-58113-465-7/02/0010.* Pp. 150-156.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Anstey, J., Pape, D., & Sandin, D. (2000). *The Thing Growing: Autonomous Characters in Virtual Reality Interactive Fiction. Virtual Reality Proceedings* (pp. 71–78). New Brunswick, New Jersey: IEEE; doi:10.1109/VR.2000.840366

Ayiter, E., Glasauer, S., & Moswitzer, M. (2013). LPDT2: La plissure du texte 2. In D. Harrison (Ed.), *Digital Media and Technologies for Virtual Artistic Spaces* (pp. 75–90). Hershey, PA: IGI Global. doi:10.4018/978-1-4666-2961-5.ch006

Bates, J. (1994). The Role of Emotion in Believable Agents. *Communications of the ACM*, *37*(7), 122–125. doi:10.1145/176789.176803

Blumberg, B., & Galyean, T. (1995, August). Multi-Level Direction of Autonomous Creatures for Real-Time Virtual Environments. *Proceedings of SIGGRAPH 95, ACM SIGGRAPH, Los Angeles, CA, August 1995.* pp. 47-54. 10.1145/218380.218405

Brahnam, S. (2001). Creating Physical Personalities for Agents with Faces: Modeling Trait Impressions of the Face. *Proceedings of the Workshop on "Attitudes, Personality and Emotions in User-Adapted Interaction", held in conjunction with the Conference on User Modeling 2001.* Pp. 1-7.

Bringsjord, S. (1992). *What Robots can and can't be*. Norwell, MA: Kluwer Academic Publishers. doi:10.1007/978-94-011-2464-5

Bringsjord, S. (2008). Declarative/Logic-Based Computational Cognitive Modeling. In R. Sun (Ed.), *The Cambridge Handbook of Computational Psychology* (pp. 127–169). Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511816772.008

Bringsjord, S., Shilliday, A., Taylor, J., Werner, D., Clark, M., Charpentier, E., & Bringsjord, A. (2008). Toward logic-based cognitively robust synthetic characters in digital environments. *Frontiers in Artitifical Intelligence and Applications*, *171*, 87–98.

Burden, D. J. H. (2009, October). Deploying Embodied AI into Virtual Worlds. *28th SGAI International Conference on Artificial Intelligence. Knowledge-Based Systems*, *22*(7), 540–544. doi:10.1016/j.knosys.2008.10.001

Chong, H.-Q., Tan, A.-H., & Ng, G.-W. (2007). Integrated cognitive architectures: A survey. *Artificial Intelligence Review*, *28*(2), 103–130. doi:10.100710462-009-9094-9

Colton, S. (2008, Spring). Creativity versus the Perception of Creativity in Computational Systems. *Proceedings of the AAAI Spring Symposium on Creative Systems*.

El-Nasr, M., Yen, J., & Ioerger, T. R. (2000). FLAME - Fuzzy Logic Adaptive Model of Emotions. [Netherlands.]. *Autonomous Agents and Multi-Agent Systems*, *3*(3), 219–257. doi:10.1023/A:1010030809960

Friedenberg, J., & Silverman, G. (2014). *Cognitive Science: An Introduction to the Study of Mind* (2nd ed.). New York: SAGE publications.

Funge, J., Tu, X., & Terzopoulos, D. (1999, August). Cognitive Modeling: Knowledge, Reasoning and Planning for Intelligent Characters. *Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '99)*, pp. 29-38, Los Angeles, CA, August 1999. 10.1145/311535.311538

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Goertzel, B. (2009, June). OpenCogPrime: A cognitive synergy based architecture for artificial general intelligence. In *Cognitive Informatics, 2009. ICCI'09. 8th IEEE International Conference* (pp. 60-68). IEEE. 10.1109/COGINF.2009.5250807

Goertzel, B. (2014). GOLEM: Towards an AGI meta-architecture enabling both goal preservation and radical self-improvement. *Journal of Experimental & Theoretical Artificial Intelligence*, 1–13.

Goertzel, B., & Pennachin, C. (2007). *Artificial General Intelligence*. New York: Springer. doi:10.1007/978-3-540-68677-4

Gordon, A. S., Hobbs, J. R., & Cox, M. T. (2011). Anthropomorphic Self-Models for Metareasoning Agents. In Metareasoning: Thinking About Thinking (pp. 295-305). MIT Press: Cambridge. 335 pages. doi:10.7551/mitpress/9780262014809.003.0019

Hart, D., & Goertzel, B. (2008). OpenCog: A Software Framework for Integrative Artificial General Intelligence. *Artificial General Intelligence, 2008: Proceedings of the First AGI Conference* (Vol. 171, p. 468-472). IOS Press.

Hebb, D. O. (1949). *The Organization of Behavior*. New York: Wiley & Sons.

Hibbard, B. (2012). Model-based utility functions. *Journal of Artificial General Intelligence*, *3*(1), 1–24. doi:10.2478/v10229-011-0013-5

Holland, O. (2004). The Future of Embodied Artificial Intelligence: Machine Consciousness? Lecture Notes in Computer Science, 2004, Vol. 3139/2004, 629, DOI: . Pp. 74-85. doi:10.1007/978-3-540-27833-7_3

Kluwer, T., Adolphs, P., Xu, F., Uszkoreit, H., & Cheng, X. (2010). Talking NPCs in a Virtual Game World. *ACLDemos '10 Proceedings of the ACL 2010 System Demonstrations. Association for Computational Linguistics (Stroudsburg).*

Legg, S., & Hutter, M. (2005, July). A universal measure of intelligence for artificial agents. *International Joint Conference on Artificial Intelligence (*Vol. 19*)* p. 1509. Marwah, NJ: Lawrence Erlbaum Associates Ltd.

Lekavy, M., & Navrat, P. (2007, May/June). Expressivity of STRIPS-like and HTN-like Planning. Lecture Notes in Artificial Intelligence, Vol. 4496 Agent and multi-agent Systems. *Technologies and applications. 1st KES International Symposium, KES-AMSTA 2007, Wroclaw, Poland, May/June 2007*. - Germany, Springer-Verlag Berlin Heidelberg, 2007. pp. 121-130

Lim, M. Y., Dias, J., Aylett, R., & Paiva, A. (2010). *Creating Adaptive Affective Autonomous NPCs. Autonomous Agent Multi-Agent Systems* (pp. 287–311). Berlin: Springer.

Lindes, P., Lonsdale, D. W., & Embley, D. W. (2015, October). Ontology-Based Information Extraction with a Cognitive Agent. *Twenty-Ninth AAAI Conference on Artificial Intelligence*. Pp. 558-564.

Loyall, A. B., Reilly, W. S. N., Bates, J., & Weyhrauch, P. (2004). System for Authoring Highly Interactive Personality-rich Interactive Characters*". Eurographics/ACM SIGGRAPH Symposium on Computer Animation*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Marques, H. G., & Holland, O. (2009). Architectures for Functional Imagination. *Neurocomputing*, *72*(4-6), 743–759. doi:10.1016/j.neucom.2008.06.016

Marr, D. (1982). *Vision: A Computational Approach*. San Francisco: Freeman & Co.

McCormack, J., Romero, J., & Machado, P. (2007). Facing the Future: Evolutionary Possibilities for Human-Machine Creativity. In J. McCormack, J. Romero, & P. Machado (Eds.), *The Art of Artificial Evolution: A Handbook on Evolutionary Art and Music* (pp. 417–451). Berlin: Springer.

Minsky, M. (2006). *The Emotion Machine: Commonsense Thinking, Artificial Intelligence and the Future of the Human Mind*. New York: Simon & Schuster.

Mueller, E. T., & Dyer, M. G. (1985). Towards a Computational Theory of Daydreaming. *Proceedings of the Seventh Annual Conference of the Cognitive Science Society*, Irvine, CA, pp. 120-135.

Nareyek, A. (2004, February). AI in Computer Games. *Queue*, *1*(10), 58–65. doi:10.1145/971564.971593

Nau, D. S., Au, T. C., Ilghami, O., Kuter, U., Muñoz-Avila, H., Murdock, J. W., ... Yaman, F. (2005, March-April). Applications of SHOP and SHOP2. *IEEE Intelligent Systems*, *20*(2), 34–41. doi:10.1109/MIS.2005.20

Nivel, E., Thórisson, K. R., Steunebrink, B. R., Dindo, H., Pezzulo, G., Rodriguez, M., . . . Jonsson, G. K. (2013). *Bounded Recursive Self-Improvement.* arXiv preprint arXiv:1312.6764.

Nowak, K.L. (2004, January). The Influence of Anthropomorphism and Agency on Social Judgment in Virtual Environments. *Journal of Computer-Mediated Communication. Vol. 19, Issue No.2* Pp. 2-30.

Orkin, J. (2006, March). Three states and a plan: the AI of FEAR. *Game Developers Conference (*Vol. 2006, *p.* 4*).* 18 pages.

Perlis, D. (2011). There's No "Me" in "Meta" – Or Is There? In Cox, M.T.m Raja, A. (Eds.), Metareasoning: Thinking About Thinking (pp. 15-26). Cambridge: MIT Press: Cambridge. 335 pages.

Rabin, S. (2009). [Rabin, S., Ed.] Introduction to Game Development, Second Edition. Chapter 5.3: Artificial Intelligence: Agents, Architecture and Techniques. Newton Center, MA: Charles River Media. Pp. 538-574.

Ranathunga, S., Cranefield, S., & Purvis, M. (2011, May). Interfacing a Cognitive Agent Platform with a Virtual World: A Case Study using Second Life (Extended Abstract). *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2011).* Turner, Yolum, Sonenberg and Stone (eds.). May 2-6, 2011, Taipei, Taiwan, pp. 1181-1182.

Robinett, W. (1998). Interactivity and Individual Viewpoint in Shared Virtual Worlds. In Dodsworth Jr., C. (Ed.), Digital Illusion: Entertaining the Future with High Technology, Vol. 28, Issue 2. New York, NY: ACM Press, Ch. 21, pp. 331-342.

Roman, M., Sandu, I., & Buraga, S. C. (2011). Owl-Based Modeling of RPG Games. *Informatica*, *LVI*(3), 83–90.

Russell, S., & Norvig, P. (2010). Artificial Intelligence: A Modern Approach. Third Edition. Chps: Intro, 3, 12, 7-17. Upper Saddle River, NJ: Prentice Hall.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

SoarTech. (2013). *Entity-Level Patterns of Life for Virtual Environments*. Promotional Brochure.

Strannegård, C. (2005). *Anthropomorphic Artificial Intelligence. Philosophical Communications #33* (pp. 169–181). Department of Philosophy, Gothenburg University.

Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction: Vol. 1. *No. 1*. Cambridge, MA: MIT press.

Thórisson, K., & Helgasson, H. P. (2012). Cognitive architectures and autonomy: A comparative review. *Journal of Artificial General Intelligence*, *3*(2), 1–30. doi:10.2478/v10229-011-0015-3

Ventrella, J. (2010). *Virtual Body Language: Avatars, Smileys and Digital Puppets: Evolving the Future of Nonverbal Expression on the Internet. Beta Edition*. Petaluma, CA: Eyebrain Books.

Waterworth, E.A. & J.A. (2003). The Illusion of Being Creative. In Riva. G. Et al (Eds.), *Being There: Concepts, Effects and Measurements of User Presence in Synthetic Environments* (pp. 223-236), Chapter 15, pp. 223-236. Amsterdam, NL: IOS Press.

Whitelaw, M. (2004). *Metacreation: Art and Artificial Life*. Cambridge, MA: MIT Press.

Zhao, R., & Szafron, D. (2012). *Generating Believable Virtual Characters Using Behavior Capture and Hidden Markov Models. Advances in Computer Games* (pp. 342–353). Berlin: Springer.

Zilberstein, S. (2011). Metareasoning and Bounded Rationality. In Cox., M.T., Raja, A. 2011 [Eds.], Metareasoning: Thinking About Thinking (pp. 27-40). Cambridge, MA: MIT Press 335 pages. doi:10.7551/mitpress/9780262014809.003.0003

KEY TERMS AND DEFINITIONS

Bot: Social virtual world slang for a Virtual Agent.
Chunk: Episodic associations compressed into a singular generalized concept.
Cognitive Architecture: A unified collection of cognitive models used to regulate mental processes.
Declarative Knowledge: Memorized Information that pertains to factual information.
NPC: Video-game slang for a "Non-Player Character" (aka. a Virtual Agent).
Production: A type of rule that is dynamically governed by alterable conditional statements.
Virtual Agent: The academic name for an artificially intelligent entity in a virtual environment.

ENDNOTES

[1] Since at least 1971. For one of the earliest cognitive-architecture flow-charts, see page 3b & 5b of Atkinson, R. C., & Shiffrin, R. M. (1971). The control processes of short-term memory. Technical report no. 173. Institute for Mathematical Studies in the Social Sciences, Stanford University. 23 pages.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

[2] Soar was originally treated as an acronym. Therefore, Soar would normally be capitalized as SOAR. However, there is now evidence to suggest that the Soar community no longer uses the acronym convention to describe this cognitive architecture (Ritter, Kim, and Sanford, 2012).

[3] For more historical contextualization, read Anderson 1996 & Best et al. 2006 (in Sun, 2006, p. 187-189).

[4] Of note, the creator of ACT-R has interfaced his architecture a 2D virtual space-ship that is the main player-avatar of the videogame "Space Fortress" (Bothell, 2010). However, to limit the scope of this paper, a space-ship will, in in this case, not be considered as a "character". See Moon & Anderson, 2012.

[5] Smart et al. (2014) make the distinction that their abELICIT agents designed by Smart et al. (2014) for ACT-R are not really "[…] based on a cognitive architecture that is modeled after the human cognitive system" (Smart et al. 2014a, p. 5).

[6] For more information, see van Gelder, T. (2001). "Distributed vs. Local Representation" (pp. 236-*238).* In R.A. Wilson., F.C. Keil., [Eds.]. The MIT Encyclopedia of Cognitive Science (MITECS). Cambridge, MA: MIT Press. 964 pages.

[7] Based on email correspondence with the lead CLARION developer, Michael Lynch on March 04, 2015.

[8] Second Life is the most popular social virtual world with its own proprietary "Linden Scripting Language" (LSL) for virtual agent AI implementations.

[9] Apparently, STEVE can draw from different declarations related to other domains that support virtual worlds and associated objects, if needed (Rickel & Johnson, 2000, p. 96). Also, STEVE can re-tailor his training demonstration by to adjusting his plan to world-state changes and "failure modes" induced by human-interactors in the VR world (Rickel & Johnson, 2000., p. 97).

[10] To clarify, SoarTech developed a "[…] hybrid symbolic-numeric model" that is part of more comprehensive sub-architecture built entirely within Soar. (Taylor et al., 2007, p. 8).

[11] This research was intended to analyze and interpret Middle-Eastern (esp. Arabic) cultural protocols.

[12] Energy in this context, represents the "amount of movement" (Jacob et al., 2013, p. 4).

[13] An inverse of the "common-sense" idea of privileging accessible over conceivable states (advocated by Crawford, 2003).

[14] With Soar, only the decision cycle runs serially. The rules themselves can fire in parallel.

[15] In future Soar versions, the imagination will likely also reside in prospective memory.

[16] These are my terms to describe what should be future memory components in Soar.

[17] For example, Soar's syntax might include "elaborate-operator" as well as "propose-operator".

This research was previously published in Integrating Cognitive Architectures into Virtual Character Design edited by Steve DiPaola, Michael Nixon, Jeremy Owen Turner, and Ulysses Bernardet; pages 181-212; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 26
Applications of PNC in Artificial Intelligence

Dariusz Jacek Jakóbczak
Technical University of Koszalin, Poland

ABSTRACT

Interpolation methods and curve fitting represent so huge problem that each individual interpolation is exceptional and requires specific solutions. Presented method is such a new possibility for curve fitting and interpolation when specific data (for example handwritten symbol or character) starts up with no rules for polynomial interpolation. The method of Probabilistic Nodes Combination (PNC) enables interpolation and modeling of two-dimensional curves using nodes combinations and different coefficients γ. This probabilistic view is novel approach a problem of modeling and interpolation. Computer vision and pattern recognition are interested in appropriate methods of shape representation and curve modeling. PNC method represents the possibilities of shape reconstruction and curve interpolation via the choice of nodes combination and probability distribution function for interpolated points. It seems to be quite new look at the problem of contour representation and curve modeling in artificial intelligence and computer vision.

INTRODUCTION

Probabilistic modeling is still developing branch of the computer science: operational research (for example probabilistic model-based prognosis) (Lorton, Fouladirad & Grall, 2013), decision making techniques and probabilistic modeling (Pergler & Freeman, 2008), artificial intelligence and machine learning. There are used different aspects of probabilistic methods: stochastic processes and stochastic model-based techniques, Markov processes (Cocozza-Thivent, Eymard, Mercier & Roussignol, 2006), Poisson processes, Gamma processes, a Monte Carlo method, Bayes rule, conditional probability and many probability distributions. In this chapter the goal of probability distribution function is to describe the position of unknown points between given interpolation nodes. Two-dimensional curve (opened or closed) is used to represent the data points. So problem statement of this chapter is: how to reconstruct

DOI: 10.4018/978-1-7998-0951-7.ch026

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

(interpolate) missing points of 2D curve having the set of interpolation nodes (key points) and using the information about probabilistic distribution of unknown points. For example the simplest basic distribution leads to the easiest interpolation – linear interpolation. Apart from probability distribution, additionally there is the second factor of proposed interpolation method: nodes combination. The simplest nodes combination is zero. Thus proposed curve modeling is based on two agents: probability distribution and nodes combination. Curve interpolation (Collins, 2003) represents one of the most important problems in mathematics and computer science: how to model the curve (Chapra, 2012) via discrete set of two-dimensional points (Ralston & Rabinowitz, 2001)? Also the matter of shape representation (as closed curve-contour) and curve parameterization is still opened (Zhang & Lu, 2004). For example pattern recognition, signature verification or handwriting identification problems are based on curve modeling via the choice of key points. So interpolation is not only a pure mathematical problem but important task in computer vision and artificial intelligence. The chapter wants to approach a problem of curve modeling by characteristic points. Proposed method relies on nodes combination and functional modeling of curve points situated between the basic set of key points. The functions that are used in calculations represent whole family of elementary functions with inverse functions: polynomials, trigonometric, cyclometric, logarithmic, exponential and power function. These functions are treated as probability distribution functions in the range [0;1]. Probabilistic modeling represents a subject arising in many branches of mathematics, economics and computer science. Such modeling connects pure mathematics with applied sciences. Operations research similarly is situated on the border between pure mathematics and applied sciences. So when probabilistic modeling meets operations research, it is very interesting occasion. Our life and work are impossible without planning, time-tabling, scheduling, decision making, optimization, simulation, data analysis, risk analysis and process modeling. Thus, it is a part of management science or decision science. This comprehensive and timely publication aims to be an essential reference source, building on the available literature in the field of probabilistic modeling, operational research, planning and scheduling, data extrapolation in decision making, probabilistic interpolation and extrapolation in simulation, stochastic processes, and decision analysis. It is hoped that this text will provide the resources necessary for economics and management sciences, also for mathematics and computer sciences. Decision makers, academicians, researchers, advanced-level students, technology developers, and government officials will find this text useful in furthering their research exposure to pertinent topics in operations research and assisting in furthering their own research efforts in this field. Proposed method, called Probabilistic Features Combination (PFC), is the method of 2D curve interpolation and extrapolation using the set of key points (knots or nodes). Nodes can be treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination.

PFC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. Two-dimensional curve is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

BACKGROUND

An important problem in machine vision and computer vision (Ballard, 1982) is that of appropriate shape representation and reconstruction. Classical discussion about shape representation is based on the problem: contour versus skeleton. This chapter is voting for contour which forms boundary of the object. Contour of the object, represented by contour points, consists of information which allows us to describe many important features of the object as shape coefficients (Tadeusiewicz & Flasiński, 1991). In the chapter contour is dealing with a set of curves. Curve modeling and generation is a basic subject in many branches of industry and computer science, for example in the CAD/CAM software. The representation of shape has a great impact on the accuracy and effectiveness of object recognition (Saber, Yaowu & Murat Tekalp, 2005). In the literature, shape has been represented by many options including curves (Sebastian & Klein, 2003), graph-based algorithms and medial axis (Liu & Geiger, 1999) to enable shape-based object recognition. Digital curve (open or closed) can be represented by chain code (Freeman's code). Chain code depends on selection of the started point and transformations of the object. So Freeman's code is one of the method how to describe and to find contour of the object. An analog (continuous) version of Freeman's code is the curve α-*s*. Another contour representation and reconstruction is based on Fourier coefficients calculated in Discrete Fourier Transformation (DFT). These coefficients are used to fix similarity of the contours with different sizes or directions. If we assume that contour is built from segments of a line and fragments of circles or ellipses, Hough transformation is applied to detect contour lines. Also geometrical moments of the object are used during the process of object shape representation (Choraś, 2005).

Nowadays methods apply mainly polynomial functions, for example Bernstein polynomials in Bezier curves, splines (Schumaker, 2007) and NURBS (Rogers, 2001). But Bezier curves don't represent the interpolation method and cannot be used for example in handwriting modeling with key points (interpolation nodes). In comparison PNC method with Bézier curves, Hermite curves and B-curves (*B-splines*) or NURBS one unpleasant feature of these curves has to be mentioned: small change of one characteristic point can result in unwanted change of whole reconstructed curve. Such a feature does not appear in proposed PNC method. Numerical methods for data interpolation are based on polynomial or trigonometric functions, for example Lagrange, Newton, Aitken and Hermite methods. These methods have many weak sides (Dahlquist & Bjoerck, 1974) and are not sufficient for curve interpolation in the situations when the curve cannot be build by polynomials or trigonometric functions. Also there exists several well established methods of curve modeling, for example shape-preserving techniques (Dejdumrong, 2007), subdivision algorithms (Dyn, Levin & Gregory, 1987) and others (Kozera, 2004) to overcome difficulties of polynomial interpolation, but probabilistic interpolation with nodes combination seems to be quite novel in the area of shape modeling. Proposed 2D curve interpolation is the functional modeling via any elementary functions and it helps us to fit the curve during the computations. The method of Probabilistic Nodes Combination (PNC) enables interpolation and modeling of two-dimensional curves using nodes combinations and different coefficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function, also inverse functions. This probabilistic view is novel approach a problem of modeling and interpolation. Computer vision and pattern recognition are interested in appropriate methods of shape representation and curve modeling. PNC method represents the possibilities of shape reconstruction and curve interpolation via the choice of nodes combination and probability distribution function for interpolated points. It seems to be quite new look at the problem of contour representation and curve modeling in artificial intelligence

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

and computer vision. Function for γ calculations is chosen individually at each curve modeling and it is treated as probability distribution function: γ depends on initial requirements and curve specifications. PNC method leads to curve interpolation as handwriting modeling via discrete set of fixed knots. So PNC makes possible the combination of two important problems: interpolation and modeling. The method of Probabilistic Features Combination (PFC) enables interpolation and modeling of high-dimensional *N* data using features' combinations and different coefficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function. Functions for γ calculations are chosen individually at each data modeling and it is treated as *N*-dimensional probability distribution function: γ depends on initial requirements and features' specifications. PFC method leads to data interpolation as handwriting or signature identification and image retrieval via discrete set of feature vectors in N-dimensional feature space. So PFC method makes possible the combination of two important problems: interpolation and modeling in a matter of image retrieval or writer identification. Main features of PFC method are: PFC interpolation develops a linear interpolation in multidimensional feature spaces into other functions as *N*-dimensional probability distribution functions; PFC is a generalization of MHR method and PNC method via different nodes combinations; interpolation of *L* points is connected with the computational cost of rank $O(L)$ as in MHR and PNC method; nodes combination and coefficients γ are crucial in the process of data probabilistic parameterization and interpolation: they are computed individually for a single feature. Proposed method, called Probabilistic Features Combination (PFC), is the method of *N*-dimensional data interpolation and extrapolation using the set of key points (knots or nodes). Nodes can be treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PFC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. *N*-dimensional object is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function.

PNC IN OBJECT RECOGNITION

This chapter consists of some examples with PNC applications in Artificial Intelligence (Zhang & Lu, 2004). First example is dealing with PNC method in MHR version.

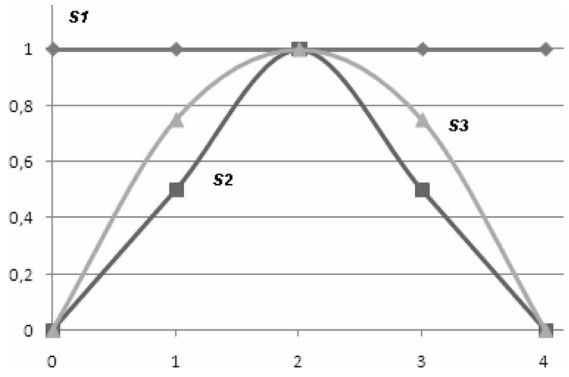
Object Recognition via Contour Reconstruction

The Example: Models

Assume that one part of *OBJECT1* contour is described by the set of points $\boldsymbol{S}_1 = \{(x_0, y_0^{(1)}), (x_1, y_1^{(1)}), \ldots, (x_M, y_M^{(1)})\}$ for $x_j < x_{j+1}$. Also one part of *OBJECT2* contour is described by the set of points $\boldsymbol{S}_2 = \{(x_0, y_0^{(2)}), (x_1, y_1^{(2)}), \ldots, (x_M, y_M^{(2)})\}$. Additionally exists $k = 0,1,..., M$ that $y_k^{(1)} = y_k^{(2)}$ (at least one common point of models). These sets of nodes $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ are models of contour fragment for *OBJECT1* and *OBJECT2*. For example $M = 4$: $\boldsymbol{S}_1 = \{(0;1), (1;1), (2;1), (3;1), (4;1)\}$ and $\boldsymbol{S}_2 = \{(0;0), (1;0.5), (2;1), (3;0.5), (4;0)\}$. Of course it is possible to deal with more models $\boldsymbol{S}_i$: let there is given the set $\boldsymbol{S}_3 = \{(0;0), (1;0.75), (2;1), (3;0.75), (4;0)\}$ for *OBJECT3*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. The sets of nodes S_1, S_2, S_3 and contour fragments of three objects



The nodes from Figure 1 are monotonic in coordinates x_i. MHR method is also working with nodes monotonic in coordinates y_i. For example here are three other models of three other objects: $\boldsymbol{S_1}'$={(1;0), (1;1), (1;2), (1;3), (1;4)}, $\boldsymbol{S_2}'$ = {(3;0), (2;1), (1;2), (2;3), (3;4)}, $\boldsymbol{S_3}'$ = {(2;0), (1.1;1), (1;2), (1.1;3), (2;4)}.

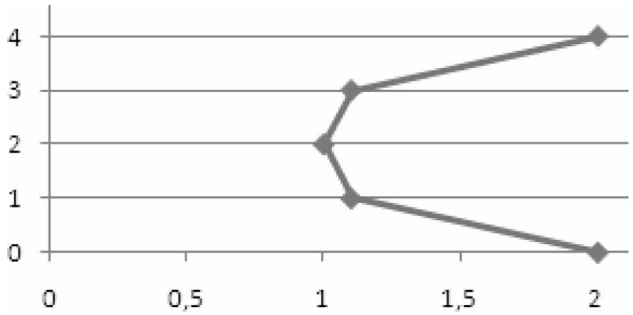
Five nodes is the minimum number for MHR method to find points between the nodes. Curves of the contour are smooth (Figure 1) or not smooth (Figure 2). The number of nodes influences precise reconstruction of the curve.

Detection

The unknown object (one part of its contour) is described by detected interpolation nodes according to MHR method: $\boldsymbol{S} = \{(x_0',y_0'), (x_1',y_1'),\ldots, (x_m',y_m')\}$ for $x_j' < x'_{j+1}$. Also assume: exists $i = 0, 1, ..., m$ that $(x_i',y_i')\in \boldsymbol{S_1}$, $(x_i',y_i')\in \boldsymbol{S_2}$ (common point of models and recognized object) and $x_0' \le x_0$, $x_m' \ge x_M$ because of using all nodes of $\boldsymbol{S_1}$ and $\boldsymbol{S_2}$ to recognition. Another part of the contour could be described by detected interpolation nodes monotonic in coordinates y_i: $\boldsymbol{S}' = \{(\underline{x}_0',\underline{y}_0'), (\underline{x}_1',\underline{y}_1'),\ldots, (\underline{x}_l',\underline{y}_l')\}$ for $\underline{y}_j' < \underline{y}'_{j+1}$. Number of models $\boldsymbol{S_i}$ is determined for $i = 1, 2, 3, 4...$ Let there is given the set $\boldsymbol{S}$= {(-0.1;0.2), (0.9;0.8), (2;1), (3.1;0.85), (4.1;0.1)} for unknown object.

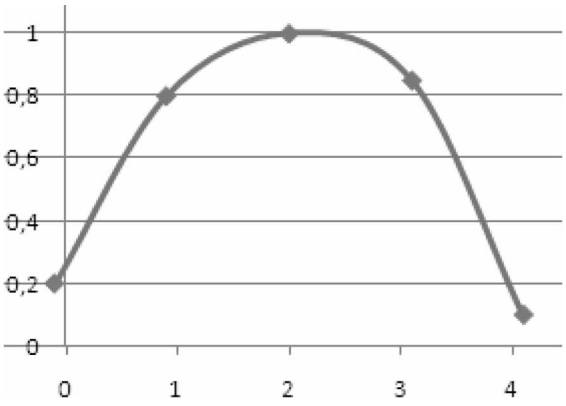
How the unknown object from Figure 3 can be recognized as *OBJECT1*, *OBJECT2* or *OBJECT3* from Figure 1?

Figure 2. The set of nodes S_3' and contour fragment of the object



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. The set of nodes S and contour fragment of unknown object



Recognition

Contour points of unknown object have to be calculated by MHR method: (x_0,y_0), (x_1,y_1),…, (x_M,y_M). The criterion of object recognition for models $\boldsymbol{S}_j$, $j = 1,2,3\ldots$ is:

$$\sum_{i=0}^{M}\left|y_i - y_i^{(j)}\right| \to \min .$$

How does this criterion work for models $\boldsymbol{S}_1$, $\boldsymbol{S}_2$, $\boldsymbol{S}_3$ and for the set of contour points $\boldsymbol{S}$? According to characteristic points (x_i,y_i) of unknown object from Figure 3 have to be computed for $x_i = 0, 1, 3, 4$. MHR calculations are done as follows:

1. For nodes (-0.1;0.2), (2;1) OHR operator (1)

$$M_0 = \begin{bmatrix} 0.494 & 0.125 \\ -0.125 & 0.494 \end{bmatrix},$$

for nodes (0.9;0.8), (3.1;0.85) OHR operator

$$M_1 = \begin{bmatrix} 0.322 & 0.165 \\ -0.165 & 0.322 \end{bmatrix},$$

for $c_1 = 0$:

$$\alpha = \frac{0.9 - 0}{0.9 - (-0.1)} = 0.9 ,$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$c_2 = \alpha\bullet 2 + (1-\alpha)\bullet 3.1 = 2.11,$

$$M_2 = \alpha \cdot M_0 + (1-\alpha)\cdot M_1 = \begin{bmatrix} 0.477 & 0.129 \\ -0.129 & 0.477 \end{bmatrix},$$

$$M_2 \cdot \begin{bmatrix} 0 \\ 2.11 \end{bmatrix} = \begin{bmatrix} 0.272 \\ 1.006 \end{bmatrix},$$

for $c_2 = 3$:

$$\alpha = \frac{3.1-3}{3.1-2} = 0.091,$$

$c_1 = \alpha\bullet(-0.1) + (1-\alpha)\bullet 0.9 = 0.809,$

$$M_2 = \alpha \cdot M_0 + (1-\alpha)\cdot M_1 = \begin{bmatrix} 0.338 & 0.161 \\ -0.161 & 0.338 \end{bmatrix},$$

$$M_2 \cdot \begin{bmatrix} 0.809 \\ 3 \end{bmatrix} = \begin{bmatrix} 0.756 \\ 0.883 \end{bmatrix};$$

2. For nodes (0.9;0.8), (3.1;0.85) OHR operator

$$M_0 = \begin{bmatrix} 0.322 & 0.165 \\ -0.165 & 0.322 \end{bmatrix},$$

for nodes (2;1), (4.1;0.1) OHR operator

$$M_1 = \begin{bmatrix} 0.116 & 0.187 \\ -0.187 & 0.116 \end{bmatrix},$$

for $c_1 = 1$:

$$\alpha = \frac{2-1}{2-0.9} = 0.909,$$

$c_2 = \alpha\bullet 3.1 + (1-\alpha)\bullet 4.1 = 3.191,$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$M_2 = \alpha \cdot M_0 + (1-\alpha) \cdot M_1 = \begin{bmatrix} 0.303 & 0.167 \\ -0.167 & 0.303 \end{bmatrix},$$

$$M_2 \cdot \begin{bmatrix} 1 \\ 3.191 \end{bmatrix} = \begin{bmatrix} 0.835 \\ 0.801 \end{bmatrix},$$

for $c_2 = 4$:

$$\alpha = \frac{4.1-4}{4.1-3.1} = 0.1\,,$$

$c_1 = \alpha \bullet 0.9 + (1-\alpha) \bullet 2 = 1.89,$

$$M_2 = \alpha \cdot M_0 + (1-\alpha) \cdot M_1 = \begin{bmatrix} 0.136 & 0.185 \\ -0.185 & 0.136 \end{bmatrix},$$

$$M_2 \cdot \begin{bmatrix} 1.89 \\ 4 \end{bmatrix} = \begin{bmatrix} 0.998 \\ 0.196 \end{bmatrix}.$$

So here are computed points to use in criterion: (0;0.272), (1;0.835), (3;0.883), (4;0.196).
The criterion gives the results:

1. For *OBJECT1* and set $\boldsymbol{S_1}$: $\sum_{i=0}^{4} \left| y_i - y_i^{(1)} \right| = 1.814$;
2. For *OBJECT2* and set $\boldsymbol{S_2}$: $\sum_{i=0}^{4} \left| y_i - y_i^{(2)} \right| = 1.186$;
3. For *OBJECT3* and set $\boldsymbol{S_3}$: $\sum_{i=0}^{4} \left| y_i - y_i^{(3)} \right| = 0.686$ = minimum.

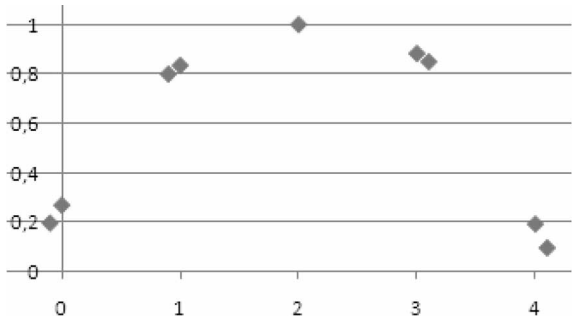
It means that unknown object from Figure 3 with the set of nodes $\boldsymbol{S}$ is recognized as *OBJECT3* from Figure 3 with the model $\boldsymbol{S_3}$.

Discussion of the Example

This example is very simple and this chapter goes to first step of application MHR method in object recognition (Ballard, 1982). Models of known objects possess special features: appropriate contour points with characteristic (identical) one coordinate. Also unknown object must have the same one coordinate. And geometrical transformations ought to be considered (translations, rotations, scaling): if unknown object is transformed, then models of the object (points) have to be also transformed and then MHR recognition is possible.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. The set of nodes **S** *and four calculated points used in the criterion*



OFFLINE TEXT-INEPENDENT HANDWRITING IDENTIFICATION AND SHAPE MODELING VIA PROBABILISTIC NODES COMBINATION

Handwriting identification and writer verification are still the open questions in artificial intelligence (Choraś, 2005) and computer vision (Tadeusiewicz & Flasiński, 1991). Handwriting based author recognition offers a huge number of significant implementations which make it an important research area in pattern recognition (Lorton, Fouladirad & Grall, 2013). There are so many possibilities and applications of the recognition algorithms that implemented methods have to be concerned on a single problem. Handwriting and signature identification represents such a significant problem. In the case of writer recognition, described in this chapter, each person is represented by the set of modeled letters or symbols. The sketch of proposed method consists of three steps: first handwritten letter or symbol must be modeled by a curve (Sebastian & Klein, 2003), then compared with unknown letter and finally there is a decision of identification (Liu & Geiger, 1999). Author recognition of handwriting and signature is based on the choice of key points and curve modeling. Reconstructed curve does not have to be smooth in the nodes because a writer does not think about smoothing during the handwriting. Curve interpolation in handwriting identification is not only a pure mathematical problem but important task in pattern recognition and artificial intelligence (Saber, Yaowu & Murat Tekalp, 2005) such as: biometric recognition (Pergler & Freeman, 2008), personalized handwriting recognition, automatic forensic document examination, classification of ancient manuscripts. Also writer recognition in monolingual handwritten texts is an extensive area of study and the methods independent from the language are well-seen. Proposed method represents language-independent and text-independent approach because it identifies the author via a single letter or symbol from the sample. This novel method is also applicable to short handwritten text.

Writer recognition methods in the recent years are going to various directions: writer recognition using multi-script handwritten texts, introduction of new features, combining different types of features, studying the sensitivity of character size on writer identification, investigating writer identification in multi-script environments, impact of ruling lines on writer identification, model perturbed handwriting, methods based on run-length features, the edge-direction and edge-hinge features, a combination of codebook and visual features extracted from chain code and polygonized representation of contours, the autoregressive coefficients, codebook and efficient code extraction methods, texture analysis with Gabor filters and extracting features, using Hidden Markov Model (Cocozza-Thivent, Eymard, Mercier & Roussignol, 2006) or Gaussian Mixture Model. But no method is dealing with writer identification via curve modeling (Dyn, Levin & Gregory, 1987) or interpolation (Kozera, 2004) and points comparing as it is presented in this chapter.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Handwriting Modeling and Author Identification

PNC method enables signature and handwriting recognition. This process of recognition consists of three parts:

1. **Modeling:** Choice of nodes combination and probabilistic distribution function for known signature or handwritten letters;
2. **Unknown Writer:** Choice of characteristic points (nodes) for unknown signature or handwritten word and the coefficients of points between nodes;
3. **Decision of Recognition:** Comparing the results of PNC interpolation for known models with coordinates of unknown text.

Modeling: The Basis of Patterns

Known letters or symbols ought to be modeled by the choice of nodes, determining specific nodes combination and characteristic probabilistic distribution function. For example a handwritten word or signature "*rw*" may look different for persons A, B or others (Dejdumrong, 2007). How to model "*rw*" for some persons via PNC method? Each model has to be described by the set of nodes for letters "*r*" and "*w*", nodes combination h and a function $\gamma=F(\alpha)$ for each letter. Less complicated models can take $h(p_1,p_2,\ldots,p_m) = 0$ and then the formula of interpolation looks as follows:

$$y(c) = \gamma \bullet y_i + (1 - \gamma) y_{i+1}.$$

It is linear interpolation for basic probability distribution ($\gamma = \alpha$). How first letter "*r*" is modeled in three versions for nodes combination $h = 0$ and α=0.1,0.2…0.9? Of course α is a random variable and $\alpha \in [0;1]$.

Person A: Nodes (1;3), (3;1), (5;3), (7;3) and $\gamma = F(\alpha) = \alpha^2$ (see Figure 5).
Person B: Nodes (1;3), (3;1), (5;3), (7;2) and $\gamma = F(\alpha) = \alpha^2$: (see Figure 6)
Person C: Nodes (1;3), (3;1), (5;3), (7;4) and $\gamma = F(\alpha) = \alpha^3$: (see Figure 7)

These three versions of letter "*r*" (Figures 5-7) with nodes combination $h = 0$ differ at fourth node and probability distribution functions $\gamma = F(\alpha)$. Much more possibilities of modeling are connected with a choice of nodes combination $h(p_1,p_2,\ldots,p_m)$. MHR method (Jakóbczak, 2007) uses the combination h with good features because of orthogonal rows and columns at Hurwitz-Radon family of matrices (Jakóbczak, 2010):

$$h(p_i, p_{i+1}) = \frac{y_i}{x_i} x_{i+1} + \frac{y_{i+1}}{x_{i+1}} x_i$$

and then

$$y(c) = \gamma \cdot y_i + (1 - \gamma) y_{i+1} + \gamma(1 - \gamma) \cdot h(p_i, p_{i+1},).$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. PNC modeling for nine reconstructed points between nodes

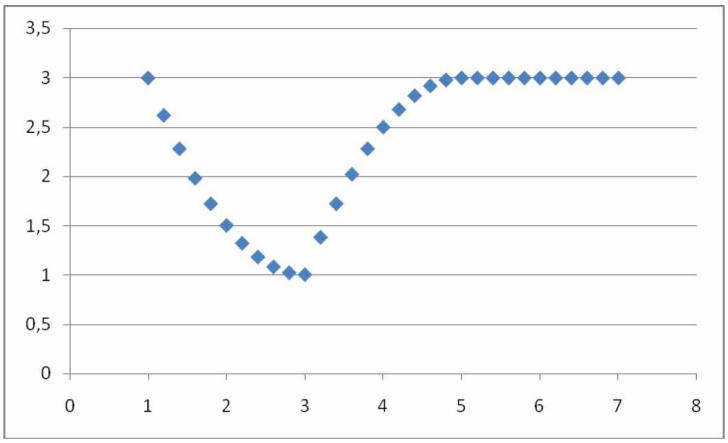


Figure 6. PNC modeling of letter "r" with four nodes

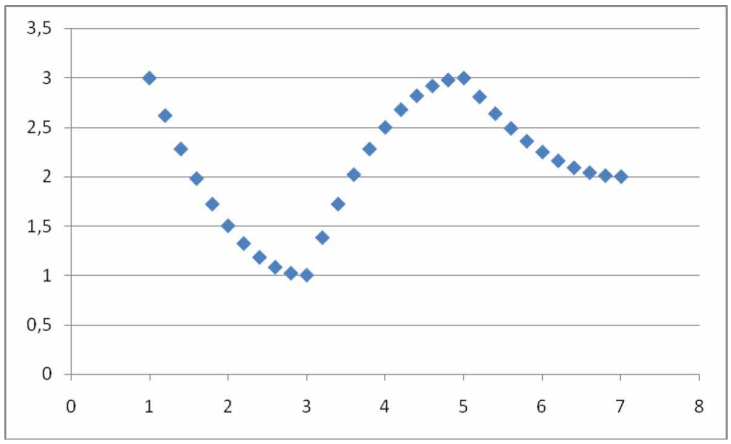
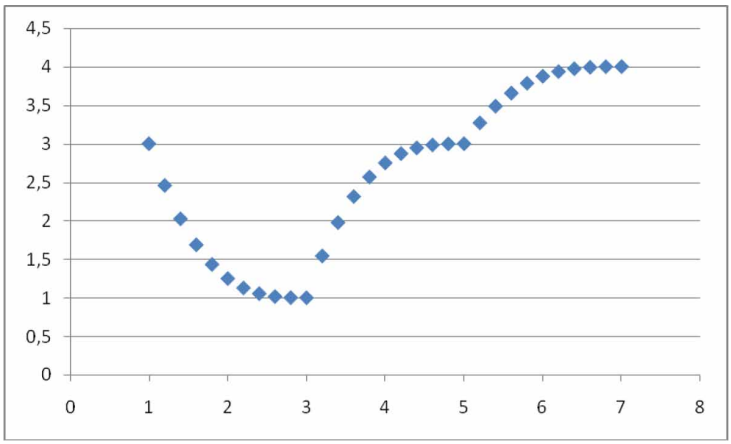


Figure 7. PNC modeling of handwritten letter "r"



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Here are two examples of PNC modeling (Jakóbczak, 2011) with MHR combination.

Person D: Nodes (1;3), (3;1), (5;3) and $\gamma = F(\alpha) = \alpha^2$: (see Figure 8)
Person E: Nodes (1;3), (3;1), (5;3) and $\gamma = F(\alpha) = \alpha^{1.5}$: (see Figure 9)

Figures 5- 9 show modeling of letter "*r*". Now let us consider a letter "*w*" with nodes combination $h = 0$.

Person A: Nodes (2;2), (3;1), (4;2), (5;1), (6;2) and $\gamma = F(\alpha) = (5^{\alpha}-1)/4$: (see Figure 10)
Person B: Nodes (2;2), (3;1), (4;2), (5;1), (6;2) and $\gamma = F(\alpha) = sin(\alpha \cdot \pi/2)$: (see Figure 11)
Person C: Nodes (2;2), (3;1), (4;2), (5;1), (6;2) and $\gamma = F(\alpha) = sin^{3.5}(\alpha \cdot \pi/2)$: (see Figure 12)

Figure 8. PNC modeling of letter "r" with three nodes

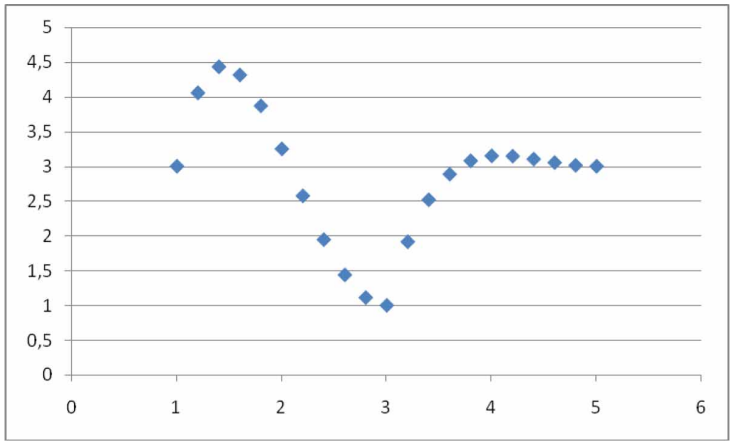
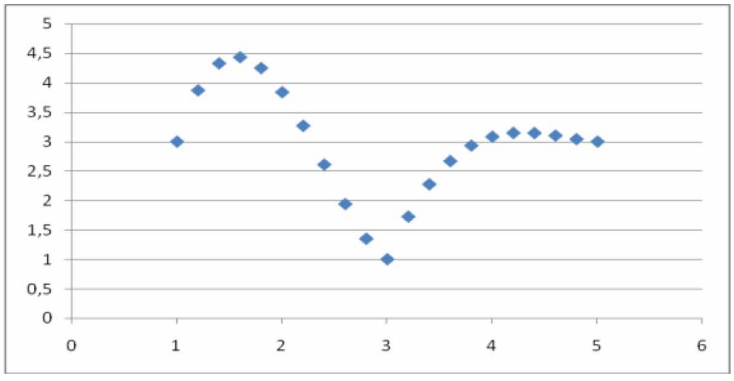


Figure 9. PNC modeling of handwritten letter "r"



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. PNC modeling for nine reconstructed points between nodes

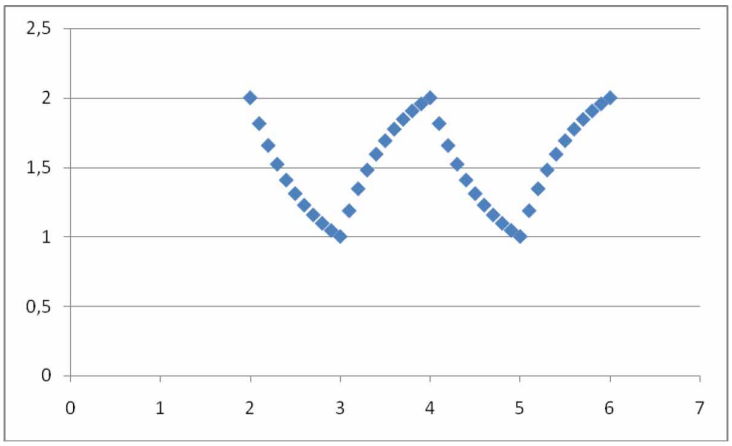


Figure 11. PNC modeling of letter "w" with five nodes

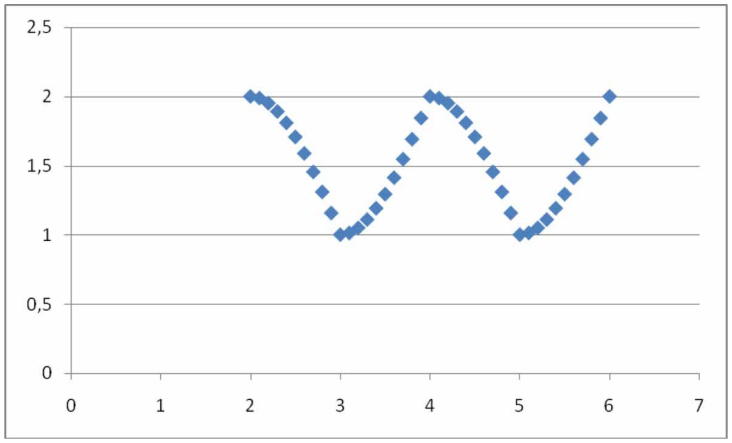
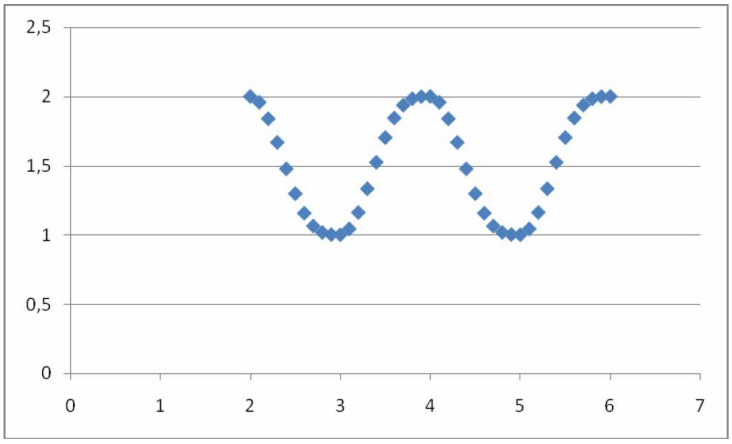


Figure 12. PNC modeling of handwritten letter "w"



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

These three versions of letter "*w*" (Figure 10-12) with nodes combination $h = 0$ and the same nodes differ only at probability distribution functions $\gamma = F(\alpha)$. Figure 13 is the example of nodes combination h formula from MHR method.

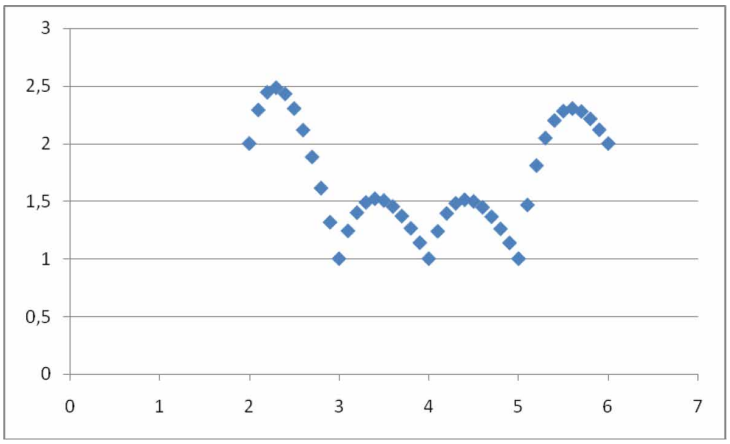
Person D: Nodes (2;2), (3;1), (4;1), (5;1), (6;2) and $\gamma = F(\alpha) = 2^{\alpha}$-1: (see Figure 13)

Examples above have one function $\gamma = F(\alpha)$ and one combination h for all ranges between nodes. But it is possible to create a model with functions $\gamma_i = F_i(\alpha)$ and combinations h_i individually for a range of nodes $(p_i;p_{i+1})$. It enables very precise modeling of handwritten symbol between each successive pair of nodes. Each person has its own characteristic and individual handwritten letters, numbers or other marks. The range of coefficients x has to be the same for all models because of comparing appropriate coordinates y. Every letter is modeled by PNC via three factors: the set of nodes, probability distribution function $\gamma = F(\alpha)$ and nodes combination h. These three factors are chosen individually for each letter, therefore this information about modeled letters seems to be enough for specific PNC curve interpolation, comparing and handwriting identification. Function γ is selected via the analysis of points between nodes and we may assume $h = 0$ at the beginning. What is very important-PNC modeling is independent of the language or a kind of symbol (letters, numbers or others). One person may have several patterns for one handwritten letter. Summarize: every person has the basis of patterns for each handwritten letter or symbol, described by the set of nodes, probability distribution function $\gamma = F(\alpha)$ and nodes combination h. Whole basis of patterns consists of models $\boldsymbol{S}_j$ for $j = 0,1,2,3\ldots K$.

Unknown Writer: Points of Handwritten Symbol

Choice of characteristic points (nodes) for unknown letter or handwritten symbol is a crucial factor in object recognition. The range of coefficients x has to be the same like the x range in the basis of patterns. Knots of the curve (opened or closed) ought to be settled at key points (Schumaker, 2007), for example local minimum or maximum (the highest point of the curve in a particular orientation), convexity changing or curvature maximum and at least one node between two successive key points (Rogers, 2001). When

Figure 13. PNC modeling for nine reconstructed points between nodes



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

the nodes are fixed, each coordinate of every chosen point on the curve (x_0^c,y_0^c), $(x_1^c,y_1^c),\ldots,(x_M^c,y_M^c)$ is accessible to be used for comparing with the models. Then probability distribution function $\gamma = F(\alpha)$ and nodes combination h have to be taken from the basis of modeled letters to calculate appropriate second coordinates $y_i^{(j)}$ of the pattern $\boldsymbol{S}_j$ for first coordinates x_i^c, $i = 0,1,\ldots,M$.

Figures 14-16 present a reconstructed signature or a handwritten symbol of unknown writer for 17 nodes (red dots). These three interpolations (Figure 14-16) are calculated for PNC method in MHR version where probability distribution function is $\gamma = F(\alpha) = \alpha^2$. Figure 14 is computed for Hurwitz-Radon matrices of dimension $N = 2$ (MHR-2) and nodes combination h with four ($2N$) points:

$$h(p_1,p_2,p_3,p_4) = \frac{1}{x_1^{\,2}+x_3^{\,2}}(x_1x_2y_1 + x_2x_3y_3 + x_3x_4y_1 - x_1x_4y_3) + \frac{1}{x_2^{\,2}+x_4^{\,2}}(x_1x_2y_2 + x_1x_4y_4 + x_3x_4y_2 - x_2x_3y_4)$$

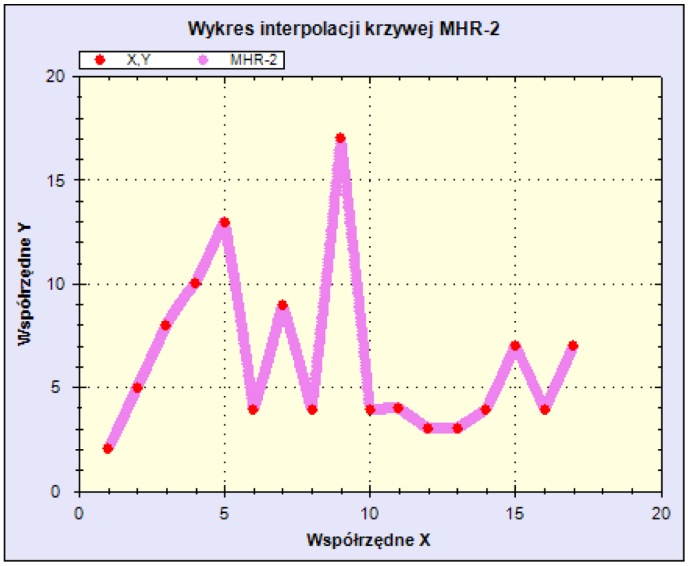
Figure 15 is reconstructed for unknown author with $h(p_1,p_2,\ldots,p_8)$ build similarly for Hurwitz-Radon matrices of dimension $N = 4$ (MHR-4) and eight ($2N$) successive nodes.

Figure 16 is reconstructed with $h(p_1,p_2,\ldots,p_{16})$ build similarly for Hurwitz-Radon matrices of dimension N=8 (MHR-8) and sixteen ($2N$) successive nodes.

Figure 17 represents the comparison of three curves from Figures 14-16. There are minimal differences between MHR-2, MHR-4 and MHR-8 because of orthogonal Hurwitz-Radon matrices. Important feature of MHR and then PNC method is the stability of calculations.

Figures 18-20 consist of Lagrange (Ralston & Rabinowitz, 2001) classical polynomial interpolations (Collins, 2003) for the same 17 nodes as Figures 14-17. Figure 18 represents four splined polynomials (Chapra, 2012) of degree 4, each polynomial is built for five successive nodes.

Figure 14. MHR-2 curve modeling for 99 reconstructed points between nodes



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 15. MHR-4 interpolation with 99 reconstructed points between nodes

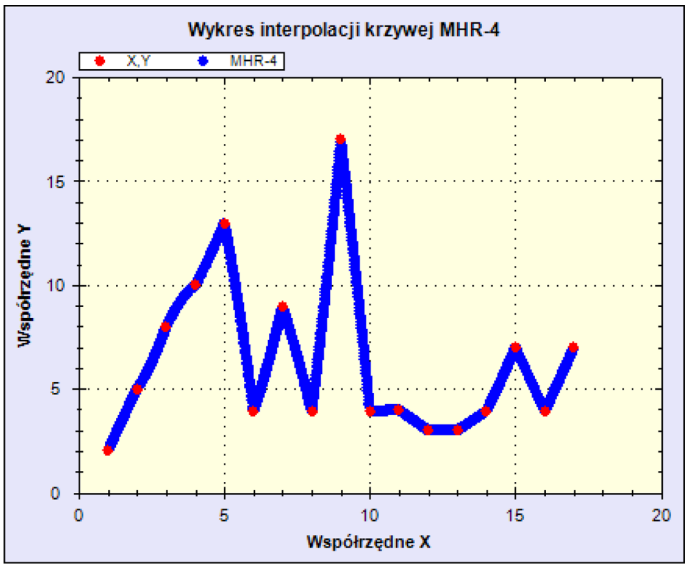


Figure 16. MHR-8 method for 99 reconstructed points between nodes

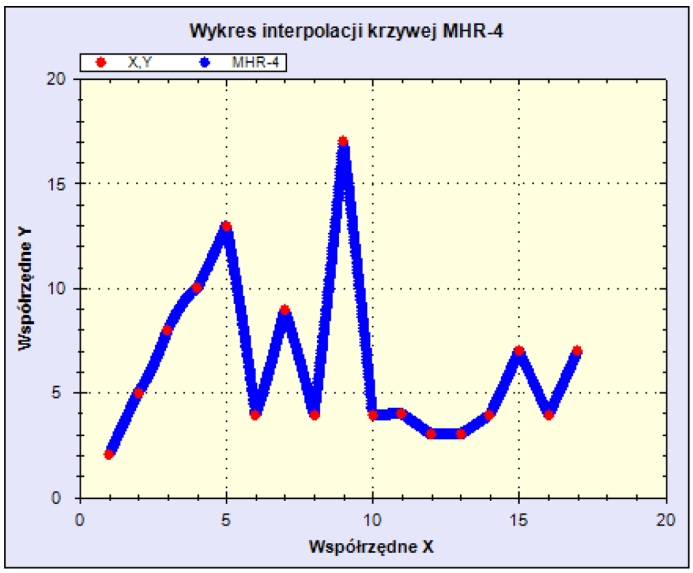


Figure 19 represents two splined polynomials of degree 8, each polynomial is built for nine successive nodes.

Figure 20 represents the polynomial of degree 16, which is built for all 17 nodes.

Polynomials of high degree (more then 4) are often useless because of some bad features (Dahlquist & Bjoerck, 1974): roots (zeros), Runge phenomenon and others. PNC modeling is more accurate then polynomial or spline interpolations. After interpolation it is possible to compare given handwritten symbol with a letter in the basis of patterns.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 17. Three versions of MHR method for the same nodes are almost identical

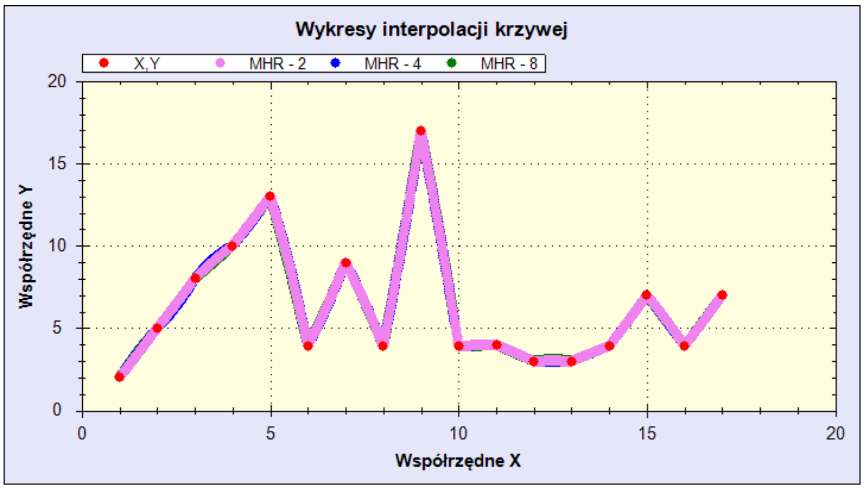
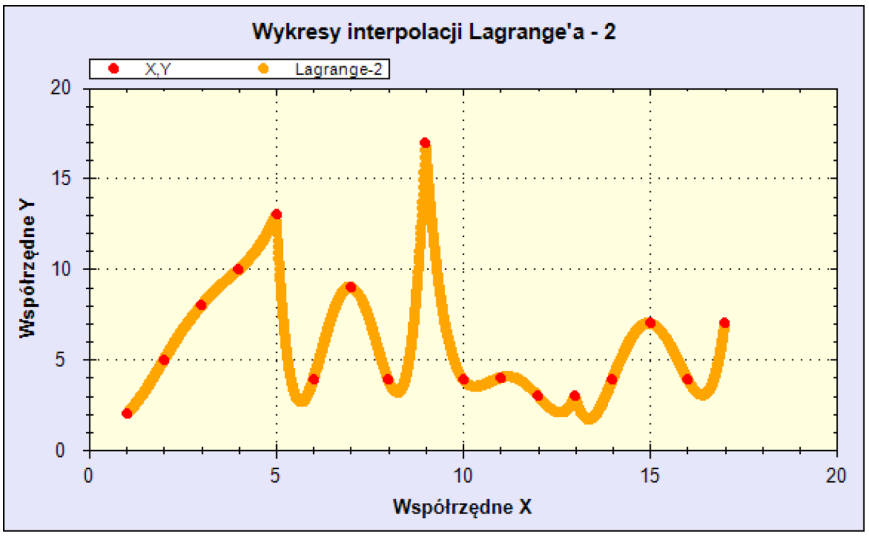


Figure 18. Four splined interpolation polynomials of degree 4



Decision of Recognition: The Author

Comparing the results of PNC interpolation for required second coordinates of a model in the basis of patterns with points on the curve (x_0^c,y_0^c), (x_1^c,y_1^c),…, (x_M^c,y_M^c), we can say if the letter or symbol is written by person A, B or another. The comparison and decision of recognition is done via minimal distance criterion. Curve points of unknown handwritten symbol are: (x_0^c,y_0^c), (x_1^c,y_1^c),…, (x_M^c,y_M^c). The criterion of recognition for models $\boldsymbol{S}_j = \{(x_0^c,y_0^{(j)}), (x_1^c,y_1^{(j)}),\ldots, (x_M^c,y_M^{(j)})\}$, j=0,1,2,3…K is given as:

$$\sum_{i=0}^{M}\left|y_i^{\,c} - y_i^{\,(j)}\right| \to \min .$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 19. Two splined interpolation polynomials of degree 8

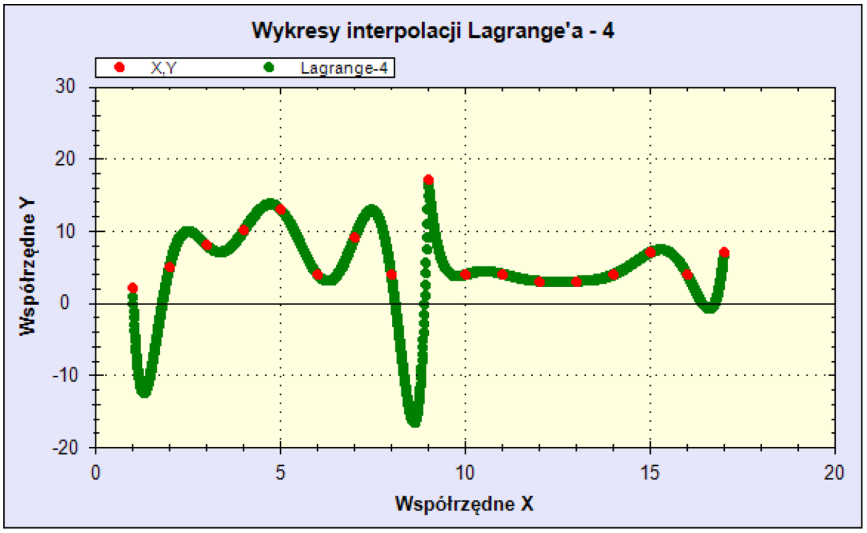
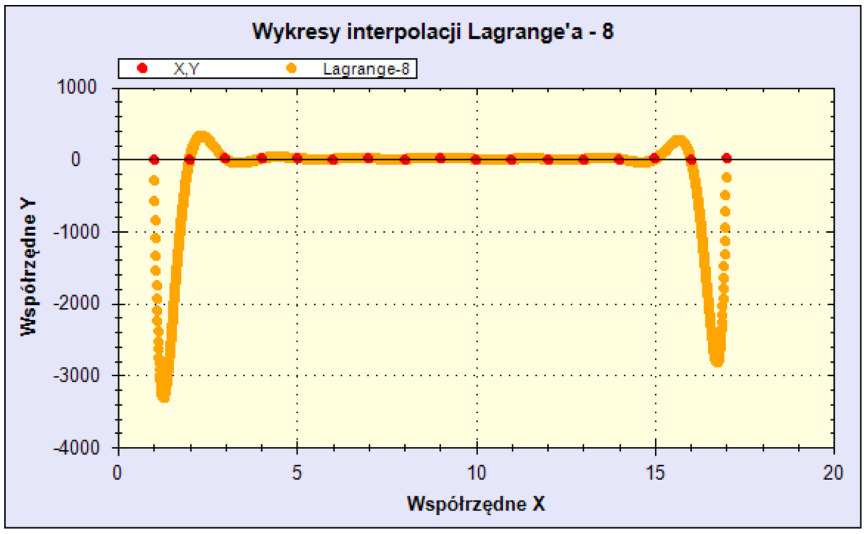


Figure 20. Interpolation polynomial of degree 16



Figures 21 and 22 show how the criterion is used for $M = 1600$ (number of reconstructed points together with nodes in Figure 14-16 is 1601) and $j = 0{,}1$ (an easy example with two persons S_0 and S_1 in the basis of patterns). Figure 21 with the model S_0 presents less differences between graphs (or between coordinates) then Figure 22 with the model S_1. So using two persons from the basis of patterns, our unknown writer (pink curve) is recognized as the pattern S_0 from Figure 21 (orange curve).

Minimal distance criterion helps us to fix a candidate for unknown writer as a person from the model $\boldsymbol{S_0}$.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 21. The comparison between unknown writer and first person S_0 from the basis

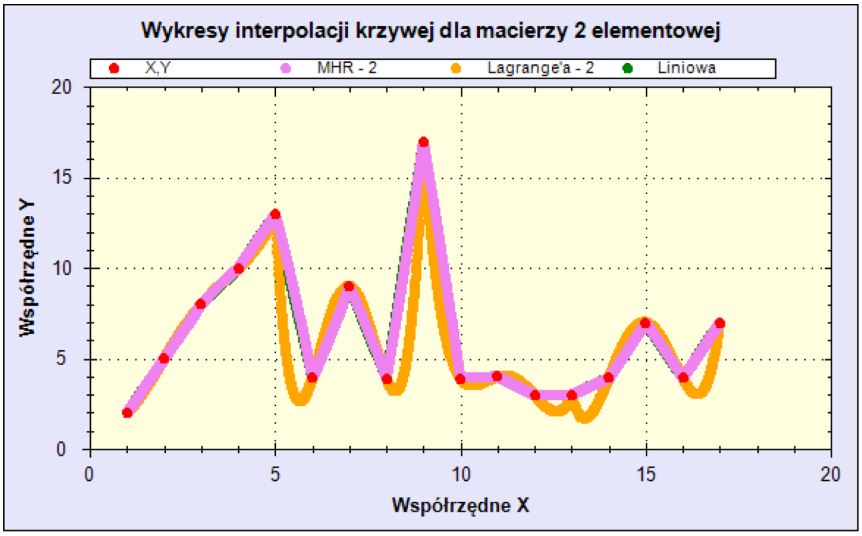
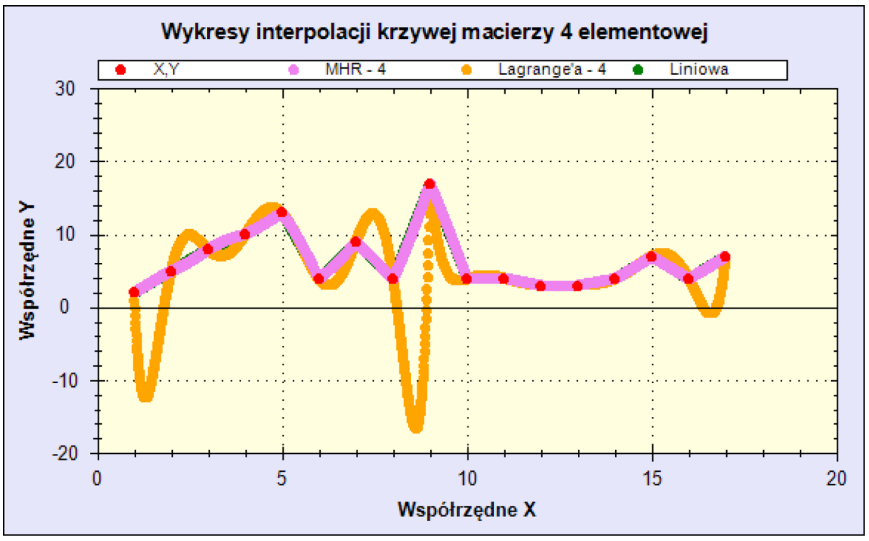


Figure 22. The comparison between unknown writer and second person S_1 from the basis



SOLUTIONS AND RECOMMENDATIONS

Proposed method, called Probabilistic Features Combination (PFC), is the method of 2D curve interpolation and extrapolation using the set of key points (knots or nodes). Nodes can be treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PFC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. Two-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

dimensional curve is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function.

FUTURE RESEARCH DIRECTIONS

Future trends will go to various directions: how to fix the best probability distribution function for the nodes, how to calculate the most appropriate nodes combination and what extrapolation is the most valuable in decision making and risk analysis.

FINAL CONCLUSION

Providing relevant and current research, this book and its author'individual publications would be useful for academics, researchers, scholars, and practitioners interested in improving decision making models and business functions. Probabilistic modeling represents a subject arising in many branches of mathematics, economics and computer science. Such modeling connects pure mathematics with applied sciences. Operations research similarly is situated on the border between pure mathematics and applied sciences. So when probabilistic modeling meets operations research, it is very interesting occasion. Our life and work are impossible without planning, time-tabling, scheduling, decision making, optimization, simulation, data analysis, risk analysis and process modeling. Thus, it is a part of management science or decision science. This book looks to discuss and address the difficulties and challenges that occur during the process of planning or decision making. The editors have found the chapters that address different aspects of probabilistic modeling, stochastic methods, probabilistic distributions, data analysis, optimization methods, probabilistic methods in risk analysis, and related topics. Additionally, the book explores the impact of such probabilistic modeling with other approaches.

This comprehensive and timely publication aims to be an essential reference source, building on the available literature in the field of probabilistic modeling, operational research, planning and scheduling, data extrapolation in decision making, probabilistic interpolation and extrapolation in simulation, stochastic processes, and decision analysis. It is hoped that this text will provide the resources necessary for economics and management sciences, also for mathematics and computer sciences. Decision makers, academicians, researchers, advanced-level students, technology developers, and government officials will find this text useful in furthering their research exposure to pertinent topics in operations research and assisting in furthering their own research efforts in this field. Proposed method, called Probabilistic Features Combination (PFC), is the method of 2D curve interpolation and extrapolation using the set of key points (knots or nodes). Nodes can be treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PFC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. Two-dimensional curve is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Risk analysis needs suitable methods of data extrapolation and decision making. Proposed method of Hurwitz-Radon Matrices (MHR) can be used in extrapolation and interpolation of curves in the plane. For example quotations from the Stock Exchange, the market prices or rate of a currency form a curve. This chapter contains the way of data anticipation and extrapolation via MHR method and decision making: to buy or not, to sell or not. Proposed method is based on a family of Hurwitz-Radon (HR) matrices. The matrices are skew-symmetric and possess columns composed of orthogonal vectors. The operator of Hurwitz-Radon (OHR), built from these matrices, is described. Two-dimensional data are represented by the set of curve points. It is shown how to create the orthogonal and discrete OHR and how to use it in a process of data foreseeing and extrapolation. MHR method is interpolating and extrapolating the curve point by point without using any formula or function.

Proposed method, called Probabilistic Nodes Combination (PNC), is the method of 2D curve interpolation and extrapolation using the set of key points (knots or nodes). Nodes can be treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PNC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. Two-dimensional curve is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function. Novelty of this book consists of two generalizations: generalization of previous MHR method with various nodes combinations and generalization of linear interpolation with different (no basic) probability distribution functions and nodes combinations. Computer vision needs suitable methods of shape representation and contour reconstruction. One of them, invented by the author and called method of Hurwitz-Radon Matrices (MHR), can be used in representation and reconstruction of shapes of the objects in the plane. Proposed method is based on a family of Hurwitz-Radon (HR) matrices. The matrices are skew-symmetric and possess columns composed of orthogonal vectors. 2D shape is represented by the set of successive nodes. It is shown how to create the orthogonal and discrete OHR operator and how to use it in a process of shape representation and reconstruction. Then MHR method is generalized to Probabilistic Nodes Combination (PNC) method.

This work clarifies the significance and novelty of the proposed method compared to existing methods (for example polynomial interpolations and Bézier curves). Previous published papers of the author were dealing with the method of Hurwitz-Radon Matrices (MHR method). Novelty of this monograph and proposed method consists in the fact that calculations are free from the family of Hurwitz-Radon Matrices. Problem statement of this book is: how to reconstruct (interpolate) missing points of 2D curve having the set of interpolation nodes (key points) and using the information about probabilistic distribution of unknown points. For example the simplest basic distribution leads to the easiest interpolation – linear interpolation. Apart from probability distribution, additionally there is the second factor of proposed interpolation method: nodes combination. The simplest nodes combination is zero. Thus proposed curve modeling is based on two agents: probability distribution and nodes combination. Significance of this chapter consists in generalization for MHR method: the computations are done without matrices in curve fitting and shape modeling, with clear point interpolation formula based on probability distribution function (continuous or discrete) and nodes combination. This book also consists of generalization for linear interpolation with different (no basic) probability distribution functions and nodes combinations. So this chapter answers the question: “Why and when should we use PNC method?”.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Curve interpolation represents one of the most important problems in mathematics and computer science: how to model the curve via discrete set of two-dimensional points? Also the matter of shape representation (as closed curve-contour) and curve parameterization is still opened. For example pattern recognition, signature verification or handwriting identification problems are based on curve modeling via the choice of key points. So interpolation is not only a pure mathematical problem but important task in computer vision and artificial intelligence. The book wants to approach a problem of curve modeling by characteristic points. Proposed method relies on nodes combination and functional modeling of curve points situated between the basic set of key points. The functions that are used in calculations represent whole family of elementary functions with inverse functions: polynomials, trigonometric, cyclometric, logarithmic, exponential and power function. These functions are treated as probability distribution functions in the range [0;1]. Significant problem in machine vision and computer vision is that of appropriate 2D shape representation and reconstruction. Classical discussion about shape representation is based on the problem: contour versus skeleton. This monograph is voting for contour which forms boundary of the object. Contour of the object, represented by successive contour points, consists of information which allows us to describe many important features of the object as shape coefficients. 2D curve modeling and generation is a basic subject in many branches of industry and computer science, for example in the cad/cam software. The representation of shape can have a great impact on the accuracy and effectiveness of object recognition. In the literature, shape has been represented by many options including curves, graph-based algorithms and medial axis to enable shape-based object recognition. Digital 2D curve (open or closed) can be represented by chain code (Freeman's code). Chain code depends on selection of the started point and transformations of the object. So Freeman's code is one of the method how to describe and to find contour of the object. Analog (continuous) version of Freeman's code is the curve α-*s*. Another contour representation and reconstruction is based on Fourier coefficients calculated in discrete Fourier transformation (DFT). These coefficients are used to fix similarity of the contours with different sizes or directions. If we assume that contour is built from segments of a line and fragments of circles or ellipses, hough transformation is applied to detect contour lines. Also geometrical moments of the object are used during the process of object shape representation. Contour is also applied in shape decomposition. Many branches of medicine, industry and manufacturing are looking for methods connected with geometry of the contour.

Why and when should we use MHR and PNC methods? Interpolation methods and curve fitting represent so huge problem that each individual interpolation is exceptional and requires specific solutions. PNC method is such a novel tool with its all pros and cons. The user has to decide which interpolation method is the best in a single situation. The choice is yours if you have any choice. Presented method is such a new possibility for curve fitting and interpolation when specific data (for example handwritten symbol or character) starts up with no rules for polynomial interpolation. This chapter consists of two generalizations: generalization of previous MHR method with various nodes combinations and generalization of linear interpolation with different (no basic) probability distribution functions and nodes combinations. The method of Probabilistic Nodes Combination (PNC) enables interpolation and modeling of two-dimensional curves using nodes combinations and different coefficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function, also inverse functions. This probabilistic view is novel approach a problem of modeling and interpolation. Computer vision and pattern recognition are interested in appropriate methods of shape representation and curve modeling. PNC method represents the possibilities of shape reconstruction and curve interpolation via the choice of nodes combination and probability distribution function

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

for interpolated points. It seems to be quite new look at the problem of contour representation and curve modeling in artificial intelligence and computer vision.

Function for γ calculations is chosen individually at each curve modeling and it is treated as probability distribution function: γ depends on initial requirements and curve specifications. PNC method leads to curve interpolation as handwriting modeling via discrete set of fixed knots. So PNC makes possible the combination of two important problems: interpolation and modeling. Main features of PNC method are:

1. The smaller distance between knots the better;
2. Calculations for coordinates close to zero and near by extremum require more attention because of importance of these points;
3. PNC interpolation develops a linear interpolation into other functions as probability distribution functions;
4. PNC is a generalization of MHR method via different nodes combinations;
5. Interpolation of L points is connected with the computational cost of rank $O(L)$ as in MHR method;
6. Nodes combination and coefficient γ are crucial in the process of curve probabilistic parameterization and interpolation: they are computed individually for a single curve.

What is the most important feature of MHR and PNC methods? Interpolation methods and curve fitting represent so huge problem that each individual interpolation is exceptional and requires specific solutions. PNC method is such a novel tool with its all pros and cons. The user has to decide which interpolation method is the best in a single situation. The choice is yours if you have any choice. Presented method is such a new possibility for curve fitting and interpolation when specific data (for example handwritten symbol or character) starts up with no rules for polynomial interpolation. This chapter consists of two generalizations: generalization of previous MHR method with various nodes combinations and generalization of linear interpolation with different (no basic) probability distribution functions and nodes combinations. The method of Probabilistic Nodes Combination (PNC) enables interpolation and modeling of two-dimensional curves using nodes combinations and different coefficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function, also inverse functions. This probabilistic view is novel approach a problem of modeling and interpolation. Computer vision and pattern recognition are interested in appropriate methods of shape representation and curve modeling. PNC method represents the possibilities of shape reconstruction and curve interpolation via the choice of nodes combination and probability distribution function for interpolated points. It seems to be quite new look at the problem of contour representation and curve modeling in artificial intelligence and computer vision. Function for γ calculations is chosen individually at each curve modeling and it is treated as probability distribution function: γ depends on initial requirements and curve specifications. PNC method leads to curve interpolation as handwriting modeling via discrete set of fixed knots. So PNC makes possible the combination of two important problems: interpolation and modeling. The method of Probabilistic Features Combination (PFC) enables interpolation and modeling of high-dimensional N data using features' combinations and different coefficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function. Functions for γ calculations are chosen individually at each data modeling and it is treated as N-dimensional probability distribution function: γ depends on initial requirements and features' specifications. PFC method leads to data interpolation as handwriting or signature identification and image retrieval via discrete set of feature vectors in N-dimensional feature

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

space. So PFC method makes possible the combination of two important problems: interpolation and modeling in a matter of image retrieval or writer identification. Main features of PFC method are: PFC interpolation develops a linear interpolation in multidimensional feature spaces into other functions as *N*-dimensional probability distribution functions; PFC is a generalization of MHR method and PNC method via different nodes combinations; interpolation of *L* points is connected with the computational cost of rank $O(L)$ as in MHR and PNC method; nodes combination and coefficients γ are crucial in the process of data probabilistic parameterization and interpolation: they are computed individually for a single feature. Future works are going to applications of PFC method in signature and handwriting biometric recognition: choice and features of nodes combinations *h* and coefficients γ.

Proposed method, called Probabilistic Features Combination (PFC), is the method of *N*-dimensional data interpolation and extrapolation using the set of key points (knots or nodes). Nodes can be treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PFC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. *N*-dimensional object is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function. The method of Probabilistic Features Combination (PFC) enables interpolation and modeling of high-dimensional data using features' combinations and different coefficients γ as modeling function. Functions for γ calculations are chosen individually at each data modeling and it is treated as *N*-dimensional probability distribution function: γ depends on initial requirements and features' specifications. PFC method leads to data interpolation as handwriting or signature identification and image retrieval via discrete set of feature vectors in N-dimensional feature space. So PFC method makes possible the combination of two important problems: interpolation and modeling in a matter of image retrieval or writer identification. PFC interpolation develops a linear interpolation in multidimensional feature spaces into other functions as *N*-dimensional probability distribution functions. Future works are going to applications of PFC method in biometric recognition, computer vision and artificial intelligence. Nodes are treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PFC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. Two-dimensional object is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function. Functions for γ calculations are chosen individually at each data modeling and it is treated as 2-dimensional probability distribution function: γ depends on initial requirements and features' specifications. PFC method leads to data interpolation as handwriting or signature identification and image retrieval via discrete set of feature vectors in 2-dimensional feature space. So PFC method makes possible the combination of two important problems: interpolation and modeling in a matter of image retrieval or writer identification. PFC interpolation develops a linear interpolation in multidimensional feature spaces into other functions as two-dimensional probability distribution functions. Future works are going to applications of PFC method in biometric recognition, computer vision and artificial intelligence.

Future works are going to: application of PNC method in signature and handwriting recognition, choice and features of nodes combinations and coefficient γ, implementation of PNC in computer vision and artificial intelligence: shape geometry, contour modelling, object recognition and curve parameterization.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Ballard, D. H. (1982). *Computer Vision*. New York: Prentice Hall.

Chapra, S. C. (2012). *Applied Numerical Methods*. McGraw-Hill.

Chorać, R. S. (2005). *Computer Vision*. Warsaw, Poland: Exit.

Cocozza-Thivent, C., Eymard, R., Mercier, S., & Roussignol, M. (2006). Characterization of the Marginal Distributions of Markov Processes Used in Dynamic Reliability. *Journal of Applied Mathematics and Stochastic Analysis*, 1–18.

Collins, G. W. II. (2003). *Fundamental Numerical Methods and Data Analysis*. Case Western Reserve University.

Dahlquist, G., & Bjoerck, A. (1974). *Numerical Methods*. New York: Prentice Hall.

Dejdumrong, N. (2007). A Shape Preserving Verification Techniques for Parametric Curves. *Computer Graphics, Imaging and Visualization*, 163–168.

Dyn, N., Levin, D., & Gregory, J. A. (1987). A 4-Point Interpolatory Subdivision Scheme for Curve Design. *Computer Aided Geometric Design*, *4*(4), 257–268. doi:10.1016/0167-8396(87)90001-X

Jakóbczak, D. (2007). 2D and 3D Image Modeling Using Hurwitz-Radon Matrices. *Polish Journal of Environmental Studies*, *4A*(16), 104–107.

Jakóbczak, D. (2009). Curve Interpolation Using Hurwitz-Radon Matrices. *Polish Journal of Environmental Studies*, *3B*(18), 126–130.

Jakóbczak, D. (2010a). Shape Representation and Shape Coefficients via Method of Hurwitz-Radon Matrices. *Lecture Notes in Computer Science*, *6374*, 411–419. doi:10.1007/978-3-642-15910-7_47

Jakóbczak, D. (2010b). Object Modeling Using Method of Hurwitz-Radon Matrices of Rank k. In W. Wolski & M. Borawski (Eds.), *Computer Graphics: Selected Issues* (pp. 79–90). Szczecin, Poland: University of Szczecin Press.

Jakóbczak, D. (2011a). Curve Parameterization and Curvature via Method of Hurwitz-Radon Matrices. *Image Processing & Communications- International Journal (Toronto, Ont.)*, *1-2*(16), 49–56.

Jakóbczak, D. (2011b). Data Extrapolation and Decision Making via Method of Hurwitz-Radon Matrices. *Lecture Notes in Computer Science*, *6922*, 173–182. doi:10.1007/978-3-642-23935-9_17

Jakóbczak, D. (2011c). Curve Extrapolation and Data Analysis using the Method of Hurwitz-Radon Matrices. *Folia Oeconomica Stetinensia*, *9*(17), 121-138.

Jakóbczak, D. (2013). Probabilistic Modeling of Signature using the Method of Hurwitz-Radon Matrices. *Global Perspectives on Artificial Intelligence*, *1*(1), 1–7.

Kozera, R. (2004). *Curve Modeling via Interpolation Based on Multidimensional Reduced Data*. Gliwice, Poland: Silesian University of Technology Press.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Liu, T., & Geiger, D. (1999). Approximate tree matching and shape similarity. *Int. Conf. Computer Vision,* Corfu, Greece.

Lorton, A., Fouladirad, M., & Grall, A. (2013). A Methodology for Probabilistic Model-based Prognosis. *European Journal of Operational Research*, *225*(3), 443–454. doi:10.1016/j.ejor.2012.10.025

Pergler, M., & Freeman, A. (2008). Probabilistic Modeling as an Exploratory Decision-Making Tool. *McKinsey Working Papers on Risk, 6*, 1-18.

Ralston, A., & Rabinowitz, P. (2001). *A First Course in Numerical Analysis* (2nd ed.). New York: Dover Publications.

Rogers, D. F. (2001). *An Introduction to NURBS with Historical Perspective*. Morgan Kaufmann Publishers.

Saber, E., Xu, Y., & Murat Tekalp, A. (2005). Partial shape recognition by sub-matrix matching for partial matching guided image labeling. *Pattern Recognition*, *38*(10), 1560–1573. doi:10.1016/j.patcog.2005.03.027

Schumaker, L. L. (2007). *Spline Functions: Basic Theory*. Cambridge Mathematical Library. doi:10.1017/CBO9780511618994

Sebastian, T. B., Klein, P. N., & Kimia, B. B. (2003). On aligning curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(1), 116–124. doi:10.1109/TPAMI.2003.1159951

Tadeusiewicz, R., & Flasiński, M. (1991). *Image Recognition*. Warsaw, Poland: PWN.

Zhang, D., & Lu, G. (2004). Review of Shape Representation and Description Techniques. *Pattern Recognition*, *1*(37), 1–19. doi:10.1016/j.patcog.2003.07.008

ADDITIONAL READING

Basu, S., & Bresler, Y. (2000). $O(N^2 \log_2 N)$ filtered backprojection reconstruction algorithm for tomography. *IEEE Transactions on Image Processing*, *9*(10), 1760–1773. doi:10.1109/83.869187 PMID:18262914

Brankov, J. G., Yang, Y., & Wernick, M. N. (2004). Tomographic image reconstruction based on a Content – Adaptive Mesh Model. *IEEE Transactions on Medical Imaging*, *2*(23), 202–212. doi:10.1109/TMI.2003.822822 PMID:14964565

Brasse, D., & Defrise, M. (2004). Fast fully 3-D image reconstruction in PET using planograms. *IEEE Transactions on Medical Imaging*, *4*(23), 413–425. doi:10.1109/TMI.2004.824231 PMID:15084067

Bulacu, M., & Schomaker, L. (2007). Text-independent writer identification and verification using textural and allographic features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(4), 701–717. doi:10.1109/TPAMI.2007.1009 PMID:17299226

Bulacu, M., Schomaker, L., & Brink, A. (2007). *Text-independent writer identification and verification on off-line Arabic handwriting*. In: *International Conference on Document Analysis and Recognition*. 769–773.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Cetin, M., Karl, W. C., & Willsky, A. S. (2002, September). *Edge – preserving image reconstruction for coherent imaging application.* Paper presented at the IEEE International Conference on Image Processsing, Rochester, NY, USA. 10.1109/ICIP.2002.1039992

Chen, J., Cheng, W., & Lopresti, D. (2011). Using perturbed handwriting to support writer identification in the presence of severe data constraints. In: Document Recognition and Retrieval. 1–10. doi:10.1117/12.876497

Chen, J., Lopresti, D., & Kavallieratou, E. (2010). *The impact of ruling lines on writer identification*. In: *International Conference on Frontiers in Handwriting Recognition*. 439–444. 10.1109/ICFHR.2010.75

Chlebus, E., & Cholewa, M. (1999). Rapid prototyping – rapid tooling. *CADCAM Forum, 11*, 23-28.

Cormen, T. H., Leiserson, C. E., & Rivest, R. L. (1996). *Introduction to algorithms*. Massachusetts, USA: the Massachusetts Institute of Technology Press and McGraw-Hill.

Defrise, M. (2001). A short readers guide to 3D tomographic reconstruction. *Computerized Medical Imaging and Graphics*, *25*(2), 113–116. doi:10.1016/S0895-6111(00)00061-6 PMID:11137787

Djeddi, C., & Souici-Meslati, L. (2010). *A texture based approach for Arabic writer identification and verification*. In: *International Conference on Machine and Web Intelligence*. 115–120.

Djeddi, C., & Souici-Meslati, L. (2011). *Artificial immune recognition system for Arabic writer identification*. In: *International Symposium on Innovation in Information and Communication Technology*. 159–165. 10.1109/ISIICT.2011.6149612

Dryja, M., Jankowska, J., & Jankowski, M. (1982). *Survey of numerical methods and algorithms. Part II*. Warsaw, Poland: WNT.

Eldar, Y. C. (2001). *Quantum Signal Processing.* (Unpublished doctoral dissertation). Massachusetts Institute of Technology, USA.

Eldar, Y. C., & Oppenheim, A. V. (2002). Quantum Signal Processing. *IEEE Signal Processing Magazine*, *6*(19), 12–32. doi:10.1109/MSP.2002.1043298

Fortuna, Z., Macukow, B., & Wąsowski, J. (1982). *Numerical methods*. Warsaw, Poland: WNT.

Galloway, M. M. (1975). Texture analysis using gray level run lengths. *Computer Graphics and Image Processing*, *4*(2), 172–179. doi:10.1016/S0146-664X(75)80008-6

Garain, U., & Paquet, T. (2009). *Off-line multi-script writer identification using AR coefficients*. In: *International Conference on Document Analysis and Recognition*. 991–995. 10.1109/ICDAR.2009.222

Ghiasi, G., & Safabakhsh, R. (2013). Offline text-independent writer identification using codebook and efficient code extraction methods. *Image and Vision Computing*, *31*(5), 379–391. doi:10.1016/j.imavis.2013.03.002

Jakóbczak, D. (2005). Hurwitz-Radon matrices and their children. *Computer Science*, *5*(8), 29–38.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Jakóbczak, D. J. (2014). *2D Curve Modeling via the Method of Probabilistic Nodes Combination-Shape Representation, Object Modeling and Curve Interpolation-Extrapolation with the Applications*. Saarbrucken: LAP Lambert Academic Publishing.

Jankowska, J., & Jankowski, M. (1981). *Survey of numerical methods and algorithms. Part I*. Warsaw, Poland: WNT.

Kontaxakis, G., & Strauss, L. G. (1998). Maximum likelihood algorithms for image reconstruction in Positron Emission Tomography. *Radionuclides for Oncology – Current Status and Future Aspects, 1998,* 73-106.

Kowalczuk, Z., & Wiszniewski, B. (Eds.). (2007). *Intelligent data mining in diagnostic purposes: Automatics and informatics*. Gdansk, Poland: PWNT.

Kundur, D., & Hatzinakos, D. (1998). A novel blind deconvolution scheme for image restoration using recursive filtering. *IEEE Transactions on Signal Processing*, *2*(46), 375–390. doi:10.1109/78.655423

Laine, A., & Zong, X. (1996). *Border identification of echocardiograms via multiscale edge detection and shape modeling*. Paper presented at the IEEE International Conference on Image Processsing, Lausanne, Switzerland. 10.1109/ICIP.1996.560486

Lang, S. (1970). *Algebra*. Reading, Massachusetts, USA: Addison-Wesley Publishing Company.

Le Buhan Jordan, C., Bossen, F., & Ebrahimi, T. (1997). *Scalable shape representation for content based visual data compression*. Paper presented at the International Conference on Image Processing, Santa Barbara, CA, USA. 10.1109/ICIP.1997.647962

Marker, J., Braude, I., Museth, K., & Breen, D. (2006). Contour-based surface reconstruction using implicit curve fitting, and distance field filtering and interpolation. *Volume Graphics*, *2006*, 1–9.

Marti, U.-V., & Bunke, H. (2002). The IAM-database: An English sentence database for offline handwriting recognition. *Int. J. Doc. Anal. Recognit.*, *5*(1), 39–46. doi:10.1007100320200071

Meyer, Y. (1993). *Wavelets: algorithms & applications*. Philadelphia, USA: Society for Industrial and Applied Mathematics.

Nosary, A., Heutte, L., & Paquet, T. (2004). Unsupervised writer adaption applied to handwritten text recognition. *Pattern Recognition Letters*, *37*(2), 385–388. doi:10.1016/S0031-3203(03)00185-7

Ozaki, M., Adachi, Y., & Ishii, N. (2006). *Examination of effects of character size on accuracy of writer recognition by new local arc method*. In: International Conference on Knowledge- Based Intelligent Information and Engineering Systems. 1170–1175. 10.1007/11893004_148

Poggio, T., & Smale, S. (2003). The mathematics of learning: Dealing with data. *Notices of the American Mathematical Society*, *5*(50), 537–544.

Przelaskowski, A. (2005). *Data compression*. Warsaw, Poland: BTC.

Rutkowski, L., Siekmann, J., Tadeusiewicz, R., & Zadeh, A. (Eds.). (2004). *Lecture notes on artificial intelligence: Artificial intelligence and soft computing*. Berlin-Heidelberg, Germany: Springer-Verlag.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Schlapbach, A., & Bunke, H. (2004). *Using HMM based recognizers for writer identification and verification. 9th Int. Workshop on Frontiers in Handwriting Recognition*. 167–172. 10.1109/IWFHR.2004.107

Schlapbach, A., & Bunke, H. (2006). *Off-line writer identification using Gaussian mixture models*. In: *International Conference on Pattern Recognition*. 992–995.

Schlapbach, A., & Bunke, H. (2007). A writer identification and verification system using HMM based recognizers. *Pattern Analysis & Applications*, *10*(1), 33–43. doi:10.100710044-006-0047-5

Schomaker, L., Franke, K., & Bulacu, M. (2007). Using codebooks of fragmented connected- component contours in forensic and historic writer identification. *Pattern Recognition Letters*, *28*(6), 719–727. doi:10.1016/j.patrec.2006.08.005

Shahabinejad, F., & Rahmati, M. (2007). A new method for writer identification and verification based on Farsi/Arabic handwritten texts, *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*. 829–833.

Siddiqi, I., Cloppet, F., & Vincent, N. (2009). Contour based features for the classification of ancient manuscripts. In: Conference of the International Graphonomics Society. 226–229.

Siddiqi, I., & Vincent, N. (2010). Text independent writer recognition using redundant writing patterns with contour-based orientation and curvature features. *Pattern Recognition Letters*, *43*(11), 3853–3865. doi:10.1016/j.patcog.2010.05.019

Vakhania, N. (1993). Orthogonal random vectors and the Hurwitz – Radon-Eckmann theorem. *Proc. of the Georgian Academy of Sciences-Mathematics, 1(1)*, 109-125.

Van, E. M., Vuurpijl, L., Franke, K., & Schomaker, L. (2005). The WANDA measurement tool for forensic document examination. *J. Forensic Doc. Exam.*, *16*, 103–118.

Willis, M. (2000). *Algebraic reconstruction algorithms for remote sensing image enhancement*. Unpublished doctoral dissertation, Department of Electrical and Computer Engineering, Brigham Young University.

Xu, Fang, & Mueller, K. (2005). Accelerating popular tomographic reconstruction algorithms on commodity PC graphics hardware. *IEEE Transactions on Nuclear Science*, *3*(52), 654–661.

Zaletelj, J., & Tasic, J. F. (2003). *Optimization and tracking of polygon vertices for shape coding*. Berlin-Heidelberg, Germany: Springer-Verlag. doi:10.1007/978-3-540-45179-2_52

Zhang, J. K., Davidson, T., & Wong, K. M. (2004). Efficient design of orthonormal wavelet bases for signal representation. *IEEE Transactions on Signal Processing*, *7*(52), 1983–1996. doi:10.1109/TSP.2004.828923

KEY TERMS AND DEFINITIONS

Artificial Intelligence: Intelligence of machines and computers, as a connection of algorithms and hardware, which makes that a man – human being can be simulated by the machines in analyzing risk,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

decision making, reasoning, knowledge, planning, learning, communication, perception and the ability to move and manipulate objects.

Contour Modeling: Calculation of unknown points of the object contour having information about some points of the object contour.

Curve Interpolation: Computing new and unknown points of a curve and creating a graph of a curve using existing data points – interpolation nodes.

Data Extrapolation: Calculation of unknown values for the points situated outside the ranges of nodes.

Hurwitz – Radon Matrices: A family of skew – symmetric and orthogonal matrices with columns and rows that create, together with identical matrix, the base in vector spaces of dimensions $N = 2, 4$ or 8.

MHR Method: The method of curve interpolation and extrapolation using linear (convex) combinations of OHR operators.

OHR Operator: Matrix operator of Hurwitz – Radon built from coordinates of interpolation nodes.

This research was previously published in Probabilistic Nodes Combination (PNC) for Object Modeling and Contour Reconstruction edited by Dariusz Jacek Jakóbczak; pages 269-308; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 27
Fair Use Defences During Copyright Litigation:
Is the Success of a Fair Use Defence Strategy Predictable?

Michael D'Rosario
Deakin University, Australia

ABSTRACT

The prediction of legal outcomes and other legal domain related variables has served as the basis of a number of recent studies. While recent studies have estimated standardised variables and dichotomous outcomes such as the outcome of a judicial decision process, few studies have employed dichotomous data and categorical data to predict the basis of a legal defense strategy or the likelihood of trial success. Empirical research within the judicial sciences continues to employ a limited subset of empirical methods. This article reasserts the benefits of several artificial intelligence based non-parametric techniques that are better suited to the discipline than many of the common methods employed within the literature. The article considers the predictability of fair use defense within the U.S. during copyright infringement proceedings, and the likelihood of trial success.

INTRODUCTION

The prediction of judicial processes has been the subject of much research in recent decades. There is a general acceptance of traditional economic methods within the sphere of legal research[1] (Barker, 1996). The application of economic and econometric methods in the legal domain is ever increasing. Artificial intelligence (AI) methods are generating substantial interest with the legal community. Some might argue that the interest has been a long time coming given the advent of such technologies over three decades ago. Indeed the more practical advent of new methods, such as AI technologies in recent years has made such technologies more accessible to those within the legal domain[2].

DOI: 10.4018/978-1-7998-0951-7.ch027

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The advent of such technologies has not met with positivity from all practitioners. But the potential has resulted in a heightened sense of importance and a desire to develop greater familiarity that is palpable[3]. This occurrence is perhaps warranting some moderation of the evidenced euphoria about such methods in legal practice[4]. While some contend that AI methods have the potential to replace practitioners, the present article takes a contrarian view and argues that AI technologies shall serve to supplement traditional legal information sources, and inform trial strategy. There is also scope for the emergence of a secondary market for data-driven legal services.

Professor Art Cockfield moderated a recent discussion where participant innovator and legal graduand Addison Cameron Fuff offered some insightful comments. "I think the change is people being more proactive. Right now lawyers are very reactive; somebody has an issue, and a lawyer researches it using books and databases. There is an opportunity for software to make that first pass, to highlight new issues. When a new case comes out, you shouldn't have to wait weeks for a newsletter. It should come into your inbox. That proactive aspect is something computers can deliver, because no lawyer on the planet could possibly read all of the cases, laws and regulations that come out. Depending on your scope, you could be talking municipal, provincial, federal, international…"[5]

Jordan Furlong in contributing to the same recent AI and Law dialogue asserts that; "we're going to see the adoption of AI in the legal market, more broadly speaking, rather than in the legal profession for quite some time to come. Lawyers are sort of naturally disinclined, for cultural reasons, to disrupt the way they work and go about their jobs. Technology tends to generate that aversion"[6]. The present article offers an example of just how effective such methods and technologies might be in support of legal practitioners.

Within the legal domain the regulatory framework and its support structure are evolving, capturing more data and enhancing courts administration and judicial accountability.

Richardson (1989) posited the advantageousness of optimisation methodology, specifically of the economic flavour when discussing the role of analysis within the courts, the author's claims remain similarly valid to Artificial intelligence methods. Essentially courts and concerned with the allocation of resources and the behaviour of individuals. While somewhat reductionist there truth in the claim.

It is therefore logical to assume that the rules and models of sanction should be framed while having regard for the potential incentives and disincentives these rules and pronouncements create, and their likely impact on future resource allocations (Richardson, 1989). The current article considers the extant empirical research employed within the literature positing an alternative to the common logit methods employed within legal research.

Advances of empiricism within the law and economics sub-disciplines serve as the starting point for the subsequent precis on research methods. Arguably, the most acceptance of empiricism has been observed in Intellectual property research and judicial decision-making research. This is noteworthy given the present article's focus on both I.P. law and judicial process outcomes. For brevity's sake, these spheres of research shall be surveyed as they represent emergent bodies of research employing empirical methodologies to legal challenges.

EMPIRICISM WITHIN INTELLECTUAL PROPERTY LAW

There is limited literature on the success of fair use defenses and empirical fair use research generally (Sag, 2012). Consequently we review the broader fair use and empirical Intellectual Property literature.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Siebeck et al (1990) offers unique insight into the economic benefits of strengthen IPRs within non-developed countries. The study reviewed the extant literature both encompassing both theoretical and empirical studies. The study claims that the body of literature pertaining to industrialized economies suggests that increases in IPR protections increase R&D activities to the extent that the social cost of having IRP protections is offset.[7] Notably, all surveyed research employed traditional regression methodologies, none of the reviewed studies employed artificial intelligence methods.

Cotropia & Gibson (2015) In "Copyright's Topography: An Empirical Study of Copyright Litigation" claim that one of the most important ways to measure the impact of copyright law is through empirical examination of actual copyright infringement cases (Ibid, 2015). The authors respond to the dearth through the analysis of differences between copyright and non-copyright litigation. The study employs traditional regression methods.

Menell (2016) considers the laws of fair use as they relate to mashups. The author contends that "the appropriate inquiry is whether an allocation mechanism achieves the best overall resolution of the trade-offs among authors' rights, cumulative creativity, freedom of expression, and overall functioning of the copyright system. By adapting the long-standing cover license for the mashup genre, Congress can support a charismatic new genre while affording fairer compensation to owners of sampled works, engaging the next generations, and channeling disaffected music fans into authorized markets." The author analyses fair use employing a conventional economics approach.

Chen (2005) considers the complex trade-off that exists between developing nations pursuing active innovation or choosing to imitate or mimic foreign rivals. The author shows that innovation within a deloping country context gives rise to increases in domestic intellectual property protections. Moreover, the countries development statas influences its level of IPR protection. The author employs a panel of data for 64 developing countries to "confirms both the positive impact of IPRs on innovations in developing countries and the presence of a U-shaped relationship between IPRs and economic development"(Chen, 2005).

Richardson (1991) in building upon general economic development studies asserts the worthwhile contribution associated with protection of IPRs within the framework of a closed economy; and considering the complexity deriving from the international nature of IPRs and innovation. The author provides a short exposition on the nature of the current U.S. practices and framework for IPR protection and enforcement particularly in relation to the interactions and model of sanctions employed on the basis of the national origin of an infringing service or product. The study contends that it is inapt to an employ a subsidy based approach to providing incentives for R&D to promote shifting profits to U.S. firms (Richardson, 1991). The noteworthy study employs traditional economic methodology within the analysis.

McCalman (2005) titled International Diffusion and Intellectual Property Rights: An Empirical Analysis is a noteworthy study rejecting "…traditional thinking about intellectual property rights (IPR) suggests a monotonically increasing relationship between property rights and the speed of diffusion of new products and technology. Our analysis of data on the international release pattern of Hollywood movies suggests a more complex story: although moderate standards of IPR encourage the spread of movies, either weaker or stronger property rights tend to decrease the speed with which American movies are released abroad. This empirical finding is consistent with a variety of specifications, including controlling for countries' self-selection of IPR standards. Overall, it appears that while some IPR recognition may encourage diffusion, very strong IPR may actually retard the speed of diffusion…" (McCalman, 2005). The study employs traditional regression methodology consistent with the broader literature.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Glass & Saggi (2002) explore the role of FDI and IPR's in moderating innovation and imitation. The authors develop a model based on product life cycle with endogenous innovation while accounting for FDI and imitation. The authors "find that stronger IPR protection keeps multinationals safer from imitation, but no more so than Northern firms. Instead, the increased difficulty of imitation generates resource wasting and imitation disincentive effects that reduce both FDI and innovation. The greater resources absorbed in imitation crowd out FDI. Reduced FDI then transmits resource scarcity in the South back to the North and consequently contracts innovation." (Glass & Saggi, 2002).

McCalman (2002) conducts an analysis of TRIP's agreement implementation focusing on the costs, benefits, and incentives associated with the agreement. The authors determined that "the static costs of raising the standards of patent protection are captured by the transfers of income between countries, with the majority of countries estimated to make net payments abroad, the United States being a major beneficiary. To offset these transfers the model provides estimates of the dynamic benefits from the greater incentive to innovate, revealing that there is potential for all countries to benefit from the TRIPs agreement in the long run. However, the distribution of these benefits is highly skewed towards developed countries." (McCalman, 2002).

The study of Gould & Gruben (1994) considers the theoretical relationship between IPRs and economic growth. This matter is frequently considered within econometrics research. Employing panel data, the study considers this relationship while accounting for country-specific considerations, such as trade, patent protections, and other country-specific factors, determining that rights and protections are strongly and positively associated with economic growth, which is consistent with the broader literature.

A further study by Yang and Maskus (2001) sought to respond to the shortage of research in intellectual property rights literature employing contemporary empirical methods such as panel data approaches and GMM estimation. The study sought to assess the linkage between licensing agreements and support for Intellectual Property rights IPRs. The study identifies the countries that are employing stronger patent protections and those that value patent rights generally attract a larger arm's length volume of licensed technology.

Employing traditional regression methods, Landes & Posner (2003) claimed that the extension of copyright had nominally positive but statistically insignificant effect on copyright registrations. Utilising pooled panel data the study is a noteworthy contribution, but due to the due to the somewhat restrictive methods utilised the study cannot account for interaction effects.

The present study response to the dearth of research employing non-parametric estimation techniques in analysis the interplay between intellectual property regulation, judicial decision making, and fair use litigation.

EMPIRICISM WITHIN JUDICIAL DECISION MAKING

The literature is replete with exploratory studies conducted considering the predictability of judicial decision making employing traditional linear OLS, GLS and logistic regression methods. Burton (1976) offers what is arguably the first systematic and extensive application of cross-judicial methodology. The authors examine the seminal accounts of the members of the Rehnquist Court (1986–94 terms) contrasting them with "…the prior appellate court experience to discern any correlation with their Supreme Court behaviour in terms of non-consensual opinion writing and voting…"

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The authors make the remarkable discover that judges change markedly. The authors assert that the judges … become less consensual as justices than they were as judges in the lower court. Importantly, this finding holds after controlling for such institutional differences between the two court levels as size, ideology, case types, stare decisis, and norms. Consistent with the neo-institutional perspective, we surmise that this behaviour change is due to the modern Supreme Court being unique, a court on which the members feel it is desirable, necessary, and possible to express policy disagreements with the majority via separate opinions and votes.[8] The study authors employ conventional regression methods in the analyses. The underlying complexity of the decision process implies that it may be non-linear in the opinion of the author.

Billica (1997) is an example of a quantitative study considering the judicial decision-making process. Specifically, Billica considered the process in the United States, by which 'states are pushed into the courts for review and interpretation.'[9] The study considered statutes across a 28-year period through the 86th and 99th congresses from 1959-86. Through the use of a randomly selected subset of statutes from the main dataset, the study tracks statutes that have been challenged by the courts, considering the issues addressed by the courts in interpreting statutes.[10]

Sisk el al (1998) considers the quandary faced by judges in 1998 when hundreds of federal district judges were suddenly confronted with the need to render a decision on the constitutionality of the Sentencing Reform Act and the newly promulgated criminal Sentencing Guidelines. Never before has a question of such importance and involving such significant issues of had constitutional law mandated the immediate and simultaneous attention of such a large segment of the federal trial bench. The authors assert that the event serves as an archetypal model for understanding the influence of key factors such as social background, ideology, judicial role and institution, and other factors on judicial decision making[11].

The authors have produced what is described as "…unprecedented empirical study of judicial reasoning in action. By exploiting this treasure trove of data, the authors have looked deeper into the judicial mind and observed the emergence of influences upon the manner in which a judge examined the constitutional issues, adopted a constitutional theory, and engaged in legal reasoning." The seminal account employs traditional regression methodology.

Yoon (2006) offers a unifying account of judicial tenure. The author claims that; Analysing judicial turnover, year by year, this article finds that judges have increasingly synchronized their departure from active service with qualifying for their judicial pension. By comparison, political and institutional factors appear to have little influence on turnover rates. These findings contradict much of the existing scholarship on judicial turnover and also offer more viable alternatives for judicial reform. The study is a further example of the style of research necessary within the legal domain[12].

Martin, Quinn, and Epstein consider the matter of median justice identification.[13] Adapting the method framed by Martin and Quinn,[14] the study develops an improved quantitative method for median justice identification based on conventional regression techniques.[15] 'The Median Justice' is an exemplary study of interdisciplinary law and decision sciences research. Nonetheless, there remain relatively few predictive studies within law and economics, particularly in relation to judicial decision-making.

Epstein et al. consider the pertinent matter of how judicial decision making is impacted by threats to national security.[16] The study employs a large dataset, incorporating every civil rights and liberties case from between 1945 and 2005. The study indicates that during periods of conflict, when national security is threatened the justices are 'substantially more likely' to curtail rights.[17]

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The study is a further example of the benefits of quantitative methods in judicial decision-making research. Like much of the research surveyed in this paper, the extant literature employs one of several popular estimation techniques. In the noted studies, logit and probit-based parametric estimation techniques are employed. The present study, as such, is a worthwhile departure from the norm, adopting a perceptron models to robustly identify associations between judicial experience and judicial outcomes and develop potentially more accurate predictive models.

Calderia and Wright consider the participation of the various parties that participate in judicial processes as Amici curiae within the U.S. Supreme courts, pursuant to petitions for writs of certiorari and jurisdictional statements, in addition to decisions based on merits. The authors employ a large dataset constituted by all cases involving amici since 1982. The study employed traditional regression methodologies.

Collins and Martinek contribute to the judicial decision-making literature by considering whether ideology mediates the impact of amicus curiae briefs on the decision making in the U.S. courts of appeals. The authors contend that amicus briefs afford the judges with critical information that influences their voting decisions. Employing conventional econometric methodology the authors assert that the amicus briefing influences judicial decision-making, but the influence is contingent on the judge's ideology.

The noted studies are broadly representative of the scope of works within the sphere of judicial decision-making research, from a methodological perspective, all employing traditional regression methodologies. [18] Herein an alternative is introduced, the multi-layer perceptron, detailing potential applications within empirical legal research.

The noted studies all employ logistic or probit modelling techniques. The most commonly used method employed in empirical judicial sciences research is logistic regression, largely because it is well suited to the dichotomous dependent variables that serve as the subject of many studies within the judicial sciences. While the contributions to empirical judicial research have been apt, many complex nonlinear problems have not been considered empirically. The present article proffers an alternative method asserted to be parsimonious for research in the judicial sciences, representing a viable alternative to the commonly employed logit model.

THE NEXT WAVE OF LEGAL EMPIRICISM, ARTIFICIAL INTELLIGENCE WITHIN THE LEGAL DOMAIN

Richardson & Hatfield (1989) describes this renewed focus on empiricism as the second wave of law and economics, with the first wave as claimed by the author evidencing only limited use of empirical methods[19]. While there remains the need for further acknowledgment of the benefits associated with parametric estimation methodology within the law and economics field of research amidst the so-called second wave of law and economics, it is apparent that novel non-parametric methodologies made possible by modern computing technologies represent an equally important sphere of inquiry (D'Rosario, 2016). These non-parametric methods offer insights not possible when employing traditional parametric regression. In particular, the advantages of multi-layer perceptron models in comparison to logit models when modelling complex nonlinear processes are pertinent to consider (D'Rosario, 2017).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

A Survey of Artificial Intelligence, Specifically Perceptron Models, and Expert System Methodology Within the Legal Domain

The Multi-layer perceptron model is asserted to be a viable supplement, and where data is limited, or relationships are complex, bi-directional and nonlinear, a superior alternative to the logit model because it offers a more accurate framework for prediction. Consider some of the challenges associated with alternative parametric techniques. Firstly, there are obvious issues associated with predictive analysis where the process is nonlinear. Logit and Probit models are relatively less capable of modelling nonlinear decision processes. MP-ANN overcomes this challenge through the application of weighted, aggregative nonlinear values based estimation. Neural networks have been advocated as an alternative modelling technique to logistic regression within the applied sciences. However there is a dearth of research considering the viability of such methods within law and economics research.

Within law and economics research, logistic regression is by far the most commonly employed method, given its suitability to dichotomous dependent variables. Within the applied sciences many more complex alternatives are being adopted. These methods present as viable alternatives to logit and probit modelling.

As asserted by Tu (1996) neural networks offer a number of advantages over logistic regression, including the ability to implicitly detect complex nonlinear relationships between dependent and independent variables, the ability to detect all possible interactions between the specified predictor variables, and the availability of multiple training algorithms. Driesietl and Ohno-Machado (2002) offer a formal methodological review of each modelling technique, noting the formal association between the techniques. The multi-layer perceptron model can be seen as an abstraction of the logistic model, employing a sigmoid function. A more apt description, however, would be that the logistic model is a simple single layer perceptron model, that does not employ a second (layer) weighted aggregate input set.[20] The following are studies employing artificial intelligence and expert systems methodology within the legal domain. The summation is not exhaustive but rather an outline of indicative studies worthwhile for legal scholars.

Zeleznikow (1995) contend that argument structures proposed by Toulmin can be used to represent legal knowledge in a manner that enables rule-based reasoning to be integrated with neural networks. The authors employ this approach for the construction of a system known as Split-up which predicts the outcome of property disputes in the domain of Australian family law.

Bellucci, Lodder and Zelznikow (1994) contend that current research in developing negotiation support systems focuses upon argumentation, artificial intelligence and game theory[21]. These techniques are rarely used in tandem. The authors argue that truly intelligent negotiation support systems require the integration of such techniques. The authors integrate the argumentation techniques of Lodder and the combined artificial intelligence/game theory approach of Bellucci and Zeleznikow to develop an online system of negotiation.

Oatley et al. (2005) summarises research within the sphere of crime prediction outlining both the authors own work and experiences as an employee of the United Kingdom police force[22]. The study incorporates detailed explanatory notes from the spheres of spacial statistics and forensic psychology that serve to inform the case based reasoning and logic programing within the presented models[23]. The crime matching techniques used are case-based reasoning, logic programming and ontologies, and naïve Bayes augmented with spatio-temporal features (Ibid, 2005). The crime prediction techniques are survival analysis and Bayesian networks (Ibid, 2005). The techniques present favourably against traditional methods in terms of ease of interpretations and enable predictive estimation not possible with traditional methods.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bellucci (2001) conducts expert systems research centered upon the sphere of Australian Family Law[24]. The authors analyse, detail and critically appraise four systems that have built in their specialized laboratory, specifically *Family_Negotiator, Split_Up, AdjustWinner* and *DEUS*. In doing so, we learn the complexities of the negotiation domain and investigate modelling issues for construction of an intelligent negotiation support tool.

The research has centered upon the study of negotiation strategies, specifically how to construct decision support technologies to help support human actor decision making and strategy formulation. The authors hope that their discoveries will result in future research involving multi-criteria decision making, the application of genetic algorithms and graph theory to build more efficacious Negotiation Decision Support Systems (NDSS).

D'Rosario (2017) seeks to identify whether trial representation experience possessed by lawyers within the U.S. Supreme Court impacts upon judicial decision making through the formulation of a predictive model. The author notes that a number of relatively recent studies within the extant literature have considered the role of experience in influencing judicial decision-making[25]. The D'Rosario study responds directly to the seminal work of McGuire[26]. McGuire is a significant study into the impact of the experience of competing trial lawyers on judicial decision making, employing traditional logistic regression methodologies[27]. The McGuire study found that trial experience possessed by trial lawyers was associated with favourable trial outcomes. The D'Rosario study extends upon McGuire, assessing the robustness of the original study employing a series of more advanced parametric estimation techniques (D'Rosario, 2017). The study then uses the McGuire logistic model framework to develop a model of prediction, employing a backward propagation, multilayer perceptron network model (D'Rosario, 2017).

D'Rosario (2016) asserts that significant methodological and theoretical contributions have been made to the extant intellectual property law and economics literature. The author asserts that 'in recent years, a number of studies have considered the impact of IPRs on software piracy, specifically TRIPS and more recently U.S. USTR 301 reporting, pursuant to the Trade Act' (D'Rosario, 2017). The noteworthy study of Shadlen (2005) is broadly supportive of the notion that IPR reforms directly influence rates of copyright infringement. The D'Rosario (2017) study extends upon Shadlen (2005), contrasting the traditional pooled panel model framework evidenced to a novel alternative model, a backward propagation, multilayer perceptron network model. The analysis conducted within the study focuses directly on ASEAN member countries. The study identifies the importance of IPRs and governance variables as well as presenting substantially improved predictive accuracies and goodness of fit measures.

Each of the noted studies incorporates the MP-ANN method in place of traditional estimation, to elucidate understanding in legal analysis and judicial sciences research. Pertinently each paper employs different data frameworks and variable structures. The studies incorporate binary, continuous, categorical and ordinal data, and as such introduce scholars to the treatment of each variable type. Each is written in a manner that is accessible to a non-technical audience.

STUDY DATA, HYPOTHESES, AND METHODOLOGY

The data was procured from Sag (2012).[28] Sag's seminal study compiled trial data and litigant data from the U.S. district courts, consider the main battle ground for copyright disputes, over a 33 year period between 1978 and 2011. The present study employs consolidated court determination data as derived by

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Variable descriptions

Variable	Variable Description	Type
Win_fu	The success of the defendant employing a fair use defense	Binary
reform	The level of advantage of the petitioner relative to the defendant (employee pool size)	Binary
UseShift	The transformative use status of the copyrighted work	Binary
Unpublished	The published status of the work	Binary
Commuse	The commercial status of the defendant	Binary
Partial	Whether the usage was partial or full	Binary
Year	The year of the case	Binary
Natural_cr	The status of the claimant as a natural person	Binary
Natural	The status of the defendant as a natural person	Binary
Repeatplayer_cr	The repeat litigation status of the claimant	3 point ordinal
Repeatplayer	The repeated litigation status of the defendant	3 point ordinal

Sag (2012), and from the United States Supreme Court Judicial database, as presented in the Sag (2014) seminal study. A total of 222 cases *(n=222)* were included in the study.

Win_fu is the dependent variable denoting a successful or unsuccessful fair use defense. Rel_firm denotes the relative firm sizes of each of the litigants. Repeatplayer is the critical repeat action variable denoting whether a defendant has appeared previously. Useshift accounts for a shift in the usage of the media, essentially transformative usage. Unpublished is a binary/Boolean variable denoting the publication status of the media. Commuse denotes the usage status of the media, either commercially on non-commercially. Partial is a variable denoting the degree of replication either full or partial. Year denotes the year of the trial. While natural denotes the natural person's status of the defendant (see Table 1).

A number of factors are likely to have influence, aside from those considered herein, the model specifications seek to determine the critical association between litigant advantage, amicus curie and, critically, judicial experience with petitioner success. The manner with which this relationship is explored appears to be the first unique contribution of the current study, extending upon the existing work of Sag (2012).[29]

It is postulated herein that the relative firm size of litigants is a key determinant of judicial outcomes when accounting for litigant natural person status. A further postulate is that transformative use is positively and significantly associated with copyright fair use defense failure. The latter postulate is consistent with Sag (2012).[30] Consequently, we test the following hypotheses:

H1: Where a defendant is larger relative to the claimant in firm size terms this position shall be positively and significantly associated with litigant fair use defense.

H2: Where the defendant engages in transformative use this action shall be positively and significantly associated with fair use defense success.

The Sag (2014) study adopted a traditional logistic regression framework to build a predictive model. Herein, three predictive models are constructed using in the first instance the same methodology as employed by Sag, and then two alternative non-parametric estimation technique capable of more accurate predictive outcomes based on neural networks theory. Firstly, a multilayer perceptron network structure

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

based on artificial neural networks theory is adopted to predict fair use defense success. Then a hybrid logit-MPANN is employed, utilising a backward propagation technique for model training. The following hypotheses pertain to the different methods of estimation:

H3: It is posited that the Perceptron shall achieve higher predictive accuracy than the traditional logit regression.

H4: The Hybrid Logit-MPANN shall achieve higher predictive accuracy than traditional logit regression and the standalone MPANN.

The MPANN approaches lend themselves well to the nature and structure of the available data. The dearth of research considering these matters is addressed by the present study which seeks to respond to omissions in the extant literature by framing a potentially more robust and accurate collection of the predictive model. A brief introduction to each of the techniques noting their benefits is detailed in the subsequent section.

This study adopts a logit model, a perceptron model and a hybrid model with the intention of understanding the impact of Repeat player status and Firm advantage on trial success and with the intent of predicting the likelihood of defence success where a fair use defence is successful, this latter task is where the multilayer perceptron model may afford the most significant benefit.

The employed data is binary, as such the excepted approach for statistical robustness all things being equal is logit regression or multiple discriminant analysis. It is posited within a number of recent studies that that such approaches are relatively less capable of accommodating dynamic and non-linear relationships. It is asserted herein that threshold effects relating to independent variables and interaction effects between independent variables cannot be estimated without knowledge of the relationship in logit specifications. A brief summary or precis on methods is provided for those unfamiliar with any of the methods, and particularly because of the new hybrid model proposed herein.

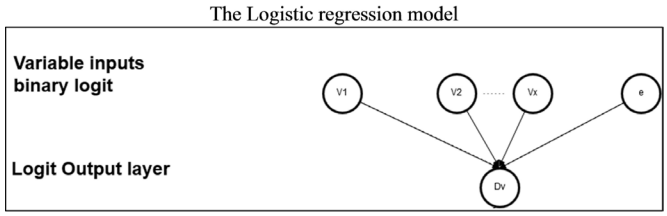
The Logit Model

Logistic regression is a common method in empirical research, sometimes called logit modeling. The method is employed where an outcome is dichotomous in nature. When logit is employed the log odds of the dichotomous outcome is modeled as a linear combination of the predictor variables. Logistic regression seeks to measure the relationship between the dichotomous dependent variable and a vector of one or more independent variables, through the estimation of coefficients as with linear regression but by estimating probabilities using a logistic function. The logistic function has a cumulative logistic distribution.

The logistic regression model is essentially a special instance of the generalised linear model and consequently, shares a level of association with linear estimation. The underlying assumptions of the logistic regression model differ markedly from those of linear regression. The most pertinent difference relates to underlying distribution structures. The logistic regression employs a Bernoulli distribution rather than a Gaussian because the dependant variable is a dichotomous binary variable. Additionally, the predicted values of the logistic specification are probabilities and therefore seek to determine outcome values of either 0 or 1 because the output variable is binary. The logit may also be conceptually viewed as a single layer perceptron. Essentially a special instance of the perceptron without a hidden layer (see Figure 1).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Conceptual drawing of the logit model (Note: Here V1…Vx are the independent variables (input variables), and e denotes the error term (referred to as bias traditionally within the machine learning). Dv denotes the Dependent variable Winfu. The solid lines denote the sigmoid/logistic regression process. Dv evidences a Bi-nominal distribution).



Model Structure

Equation (1) summarises the econometric structure of the estimation. Recall that Win_fu is the dependent variable denoting a successful or unsuccessful fair use defense. Rel_firm denotes the relative firm sizes of each of the litigants. Repeat player is the critical repeat action variable denoting whether a defendant has appeared previously. Use shift accounts for a shift in the usage of the media, essentially transformative usage. Unpublished is a binary variable denoting the publication status of the media. Commuse denotes the usage status of the media, either commercially on non-commercially. Partial is a variable denoting the degree of replication or the subject matter, copyrighted work either full or partial. Year denotes the year of the trial. While natural denotes the natural person status of the defendant, where it indexes each data item across time:

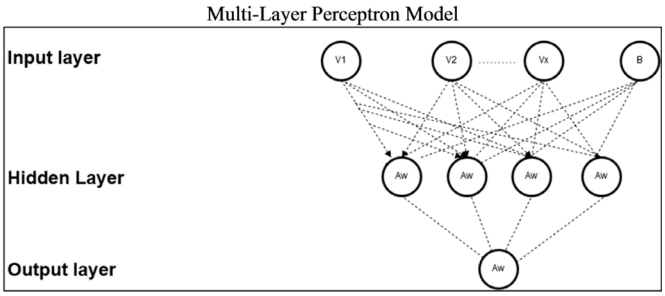
$$Winfu_{it} = a + \beta.Relfirm_{it} + \beta.Partial_{it} + \beta.Transform + \beta.Commuse_{it} + \beta.Unpublished_{it} + \beta.Repeatplayer_{it} + \beta.Natural_{it} + \beta.Repeat_cr_{it} + \beta.Natural_cr_{it} + \varepsilon_{it} \quad (1)$$

The Multilayer Perceptron Network Model

The article given both its interdisciplinary nature and focus introduces a number of technical concepts that are likely to be unfamiliar to non-technical audiences and in particular those with a largely jurisprudential background. In the interest of engaging with the broadest audience, the following precis on neural network methodology[31] is provided. The multilayer perceptron model is simply a form of the artificial neural network model (herewith MP-ANN) [32], a nonparametric predictive model[33] that seeks to replicate the structure of a biological neuron as it occurs in nature[34]. More plainly, it is a series of weighted, aggregative, non-linear values[35]. Neural network models emerged out of the body of research on artificial intelligence[36] that should attempt to model the human process of learning through the development of a modeling framework resembling the structure of the human brain.[37] The MP-ANN model has the potential to provide more accurate predictive outcomes than traditional parametric estimation techniques and is therefore well suited to the present research questions, any criticisms of legal realist logic notwithstanding[38]. Moreover, the methods appear more suited even accepting any legal realist underpinnings (see Figure 2).[39]

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Conceptual drawing of the multi-layer perceptron model (Note: Here V1…Vx are the independent variables (input variables), x denotes the number of variables and b denotes the error term (referred to as bias traditionally within the machine learning). Dv denotes the Dependent variable Winfu. Aw denote the hidden layer nodes, and the latter Aw denotes the aggregate weight at the output layer).



Model Structure

The sigmoidal function was chosen given the data structure, and its dichotomous nature. Hidden layer size was determined based on Heaton (2011) recommendations. Optimality generally exists between the number of input layers and output layers.[40] Hidden layer structure was based on optimisation. Two hidden layers offered no additional optimisation benefit. Moreover, two hidden layers may be useful when employed in situations where modeling data evidences discontinuities. Within the present dataset, this is not the case.

The use of an additional hidden layer increased computational resources required and offered no improvement to the model. A single hidden layer model is employed to avoid a greater risk of converging to an incorrect local minima. One potential shortcoming of the selected methodology relates to the efficiency of ANN models in accommodating non-numeric variables, commonly referred to as dummy variables in economics. This is, however, only the case where non-numeric variables can take a large number of values, which is not the case in the present study.

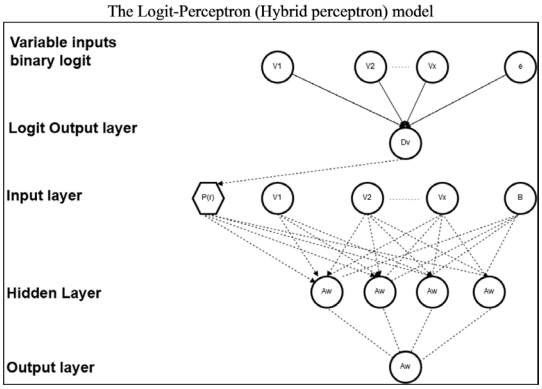
Data reduction and data pruning techniques were not necessary to improve model resolution, and the model performed efficiently and robustly absent of any node reduction. As such, the input values remain the same as those employed in Sag (2014). This is pertinent as it ensures that each predictive modeling framework can be compared on equal terms employing the same parameter set.

The Hybrid Perceptron Network Model

The hybrid multilayer perceptron model is a new form of the artificial neural network model (herewith MP-ANN), incorporating initial logit pre-processing. The logic behind the hybrid approach is rather simple, where the perceptron has a tendency to overfitting, pre-processing may be beneficial. Herein the proposed logit-perceptron derives initial probabilities from the logit model that serve as an input for the perceptron[41]. Employing this method the series of weighted, aggregative values derive input from the existing vector set of variables as well as the logistic regression probability estimates. While neural network models emerged out of the body of research on artificial intelligence that should attempt to model the human process of learning through the development of a modeling framework resembling the structure of the human brain.[42] The MP-ANN model has the potential to provide more accurate predictive

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Conceptual drawing of the Hybrid perceptron model (Note: Here V1…Vx are the independent variables (input variables), and e denotes the error term (referred to as bias traditionally within the machine learning). P(r) denotes the probabilities associated with the logit specification outputs. B denotes the bias function at the input layer. Dv denotes the Dependent variable Winfu. Aw denote the hidden layer nodes, and the latter Aw denotes the aggregate weight at the output layer).



outcomes than traditional parametric estimation techniques and is therefore well suited to the present research questions[43]. While this article may not be the first instance where hybrid methods such as this have been employed, it appears to be quite plausibly the first within the legal domain (see Figure 3).

The MP-ANN is selected as a viable alternative to the Sag (2014) logit [44] model because it offers a more accurate framework for prediction. Consider some of the challenges associated with alternative parametric techniques. Firstly, there are obvious issues associated with predictive analysis where the process is non-linear. Logit and probit models are relatively less capable of modeling non-linear decision processes robustly. MP-ANN overcomes this challenge through the application of weighted, aggregative non-linear values-based estimation. The MP-ANN model appears superior when modeling complex relationships, such as the relationship between judicial experience variables and judicial outcomes.

Neural models appear more consistent with judicial decision making[45]. Members of the judicial system carefully weigh alternative criteria, balancing competing criteria and implicitly employ these weighted considerations in their decision making. While the process may not be explicit, it is nonetheless essential to judicial decision making. Neural networks are more capable of replicating complex decision criteria and decision weighting than the commonly employed logistic regression.

The Layers of a Hybrid Neural Network

The first layer is the logit input layer. Here standardised and unstandardized logit inputs are employed in a logit specification. They are employed to specify a logit function representation of the relationship between the variables and litigation success. The output is a non-Gaussian distributed, Bernoulli distribution of litigation success. The logit specification classification probabilities are then captured and specified as an input in the next layer. The next layer is the first layer is the perceptron, an input layer. The perceptron input layer processes and standardises a vector of the predictor variables. This process loosely resembles the process of "saltatory conduction" with a biological neuron.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The standardised vector values of the predictor variables, reach the perceptron hidden layer. At the hidden layer, their standardised values are multiplied by a weight, with the weighted values aggregated and passed forward to the transfer function. The values from the transfer function are passed to the output layer.

The output layer accepts the values from the hidden layer, multiplying the values by weights, with the weighted values aggregated and passed forward to the transfer function. The values from the transfer function are the model outputs. As the target variable is the outcome for the petitioner, there are *2* neurons in the output layer producing *2* values, one for each of the categories of the target variable.

The hidden layer size was determined based on model optimisation, and Heaton (2011), noting optimality usually exists between the number of input layers and output layers.[46] Moreover, two hidden layers may be employed when modeling data that evidence discontinuities, such as a saw-tooth wave pattern (D'Rosario, 2016). Given the nature of the underlying dataset, this is not the case. The use of an additional hidden layer reduces estimation efficiency and offers no improvement to the model in this instance. Herein, a single hidden layer model is employed to avoid a greater risk of converging to an incorrect local minima.

Data pruning techniques were not employed to improve model resolution, and the model performed efficiently and robustly absent of any node reduction. As such, the input values remain the same as those employed in Sag (2014). This is pertinent as it ensures that each predictive modeling framework can be compared on equal terms employing the same parameter set.

MODEL FINDINGS

Considering first the logit specification, the findings indicate that relative firm size as measured by employees, the claimant's status as a natural person and whether the defendant was a repeat player are significant at the 1% level (relsize, cl_natural) and 10% level (df_repeat). These findings are in contrast with the earlier Sag (2014) study that identified both the amount copied, transformative use, the plaintiff status and the defendant's underdog status as all significant at either the five or one percent level. The contrary findings are likely due to the inclusion of the relsize variable which captures both underdog status and arguably more accurately represents firm level advantage. The substantially improved goodness of fit measures strongly supports this claim.

The multilayer perceptron findings offer some insight into independent variable importance. Remarkably, usage shifting and unpublished status are found to be relatively less important than the status of a defendant as a repeat player, the claimant status as a natural person, whether the infringing act involved partial or full copying and most pertinently the relative size variable which is asserted to be the most important determinant of fair use claim litigation success, employing the methods of Garson (1991) and Goh (1995) (see Table 2). These findings are in stark contrast with those of Sag (2012), the differences are likely due to the inclusion of the relsize variable and the method of estimation. The perceptron is capable of modeling complex non-linear processes and capturing variable interaction effects (Garson 1991, Goh 1995).

The hybrid model findings are largely consistent with the multi-layer perceptron findings, however the logit probabilities largely ranked first across all iterations of the hybrid model indicating that logit pre-processing was indeed worthwhile the relsize variable and the claimant and defendant status as repeat players were also found to be important employing the methods of Garson (1991) and Goh (1995).

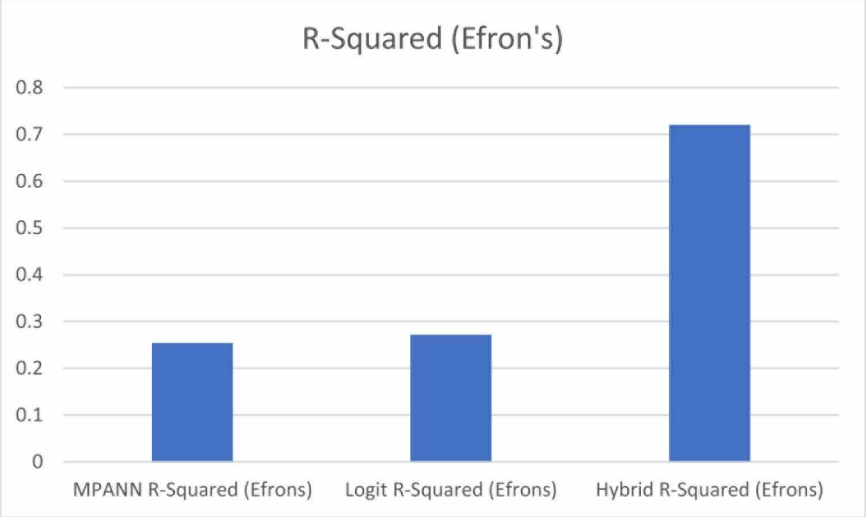
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Summary of logit model

	B	S.E.	Wald	df	Sig.	Exp(B)
case_id	0.010	0.016	0.389	1	0.533	1.010
year	-0.128	0.194	0.437	1	0.509	0.879
cr_repeat			1.346	2	0.510	
cr_repeat(1)	-1.056	0.990	1.137	1	0.286	0.348
cr_repeat(2)	0.079	0.446	0.031	1	0.859	1.082
df_repeat			3.428	2	0.180	
df_repeat(1)	-0.284	1.041	0.074	1	0.785	0.753
df_repeat(2)	-0.898	0.492	3.325	1	0.068	0.407
useshift2	0.385	0.459	0.703	1	0.402	1.470
unpublished	0.640	0.609	1.101	1	0.294	1.896
relsize	-0.009	0.003	8.859	1	0.003	0.991
cr_natural(1)	-1.783	0.469	14.464	1	0.000	0.168
df_natural(1)	-0.867	0.601	2.083	1	0.149	0.420
Constant	256.471	385.212	0.443	1	0.506	2.42+111

Note – This table summarises the co-efficient, t statistics, and associated probabilities for all variables within the specification

Figure 4. Efron's R-squared goodness of fit measures



Goodness of Fit and Predictive Accuracy

We employ Efron's method of determining goodness of fit because of its ease of interpretation (see Figure 4). Unlike other measures of goodness of fit Efron's method enables the resultant R-squared to be interpreted as a measure of explained variability within the specification. The ratio is the proportion of variability explained by the model specification. It is also mathematically the square of the correla-

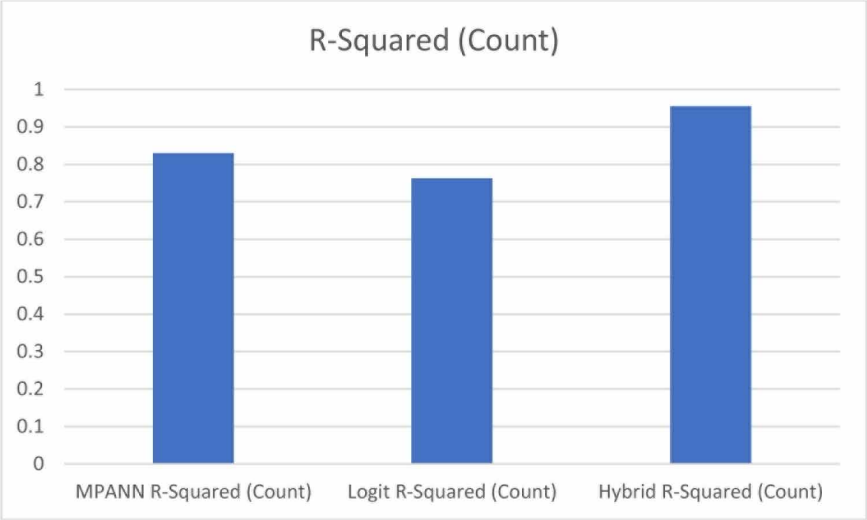
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

tion between the actual values and those predicted by the model. Consequently, this measure observes a range from minus one to plus one. This is consistent with the equivalent non-pseudo R-squared values employed in OLS regression.

The initial multi-layer perceptron model evidence an Efron's R-squared of 0.2401. This is substantially higher than the R-squared values observed within the extant literature. In Sag (2012), R-squared values vary between 0.095 and 0.215. These logit specifications were replicated to ensure consistency when framing the alternative logit, MPANN and hybrid logit/MPANN. This represents an 11.67% increase in goodness of fit between the full form logit model in Sag (2012) and the equivalent MPANN. Considering again the models presented above, we observe a shift in predictive accuracy and goodness of fit through the inclusion of a re-specified variable relsize denoting relative firm size. The inclusion of relsize, is in the logit model gives rise to substantial improvements in the logit model goodness of fit measures. Efron's R-squared, increases from 0.215 to 0.2714 a 26.25%. However, the increase in goodness of fit is not as substantial within the MPANN specification, increasing from 0.2401 to 0.2540.

The application of the hybrid method results in some remarkable efficiency gains in terms of computational time as well as goodness of fit and predictive accuracy. Through the use of the logit model as a pre-processing method, the MPANN incorporating the logit probabilities evidences an Efron's R-squared of 0.7208 representing a 201% increase in goodness of fit, when compared with the MPANN in isolation, and 184% increase in goodness of fit when compared with the logit used in isolation. While there is no perfect measure to assess model goodness of fit, all of the indicia suggests the hybrid method (logit-MPANN) offers the most robust specification. Considering now the central question of predictive accuracy, the findings indicate that the multi-layer perceptron and the hybrid method results in far greater predictive accuracy. Consider the count R-squared values for each of the models, essentially the percentage of observations predicted correctly by each model specification (see Figure 5). The traditional logit gives rise to a count R-squared of 0.763, while the multi-layer perceptron evidences a count R-squared of 0.83 equating to approximately 9.1% increase in predictive accuracy through the use of multi-layer perceptron. Through the application of the logit model as a pre-processing technique, the hybrid multi-

Figure 5.Count R-squared values for each model specification



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

layer perceptron achieves a predictive accuracy level of 0.9553, a substantial increase in predictive accuracy of 15.1 to 25.2% when compared with the logit and multi-layer perceptron respectively.

Considering the posited hypotheses, the findings are broadly consistent with the stated hypotheses. Each of the posited hypotheses is noted below:

H1: Where a defendant is larger relative to the claimant in firm size terms this position shall be positively and significantly associated with litigant fair use defense.

The findings across all specifications suggest that relative firm size is one of the most significant determinants of litigation success. Within the logit, perceptron and hybrid-perceptron models, the coefficient and independent variable importance analysis indicate that relsize ranks first in importance terms and is the most statistically significant within the logit specification.

H2: Where the defendant engages in transformative use this action shall be positively and significantly associated with fair use defense success.

In relation to transformative use the findings are mixed. Useshift is found to be nominally important within the perceptron models and positively associated with fair use defense success during litigation. However, it is found to be statistically insignificant.

H3: It is posited that the Perceptron shall achieve higher predictive accuracy than the traditional logit regression.

H4: The Hybrid Logit-MPANN shall achieve higher predictive accuracy than traditional logit regression and the standalone MPANN

Hypotheses 3 and 4 shall be considered concurrently. The perceptron and hybrid-perceptron achieved vastly superior predictive accuracy than the logit model. The hybrid-perceptron achieved the highest predicative accuracy overall, consistent with the stated hypotheses.

As such, in relation to the central consideration, the framing of a predictive model of fair use litigation success, the findings strongly favour the use of non-traditional methods specifically the multilayer perceptron and the hybrid multilayer perceptron. This is consistent with a number of studies incorporating discussion and/or analysis of logistic regression and machine learning methods (Tu, 1996; D'Rosario, 2016). There is however, a genuine dearth of research incorporating an analysis of hybrid methods and traditional methods within the legal domain. The present study appears to be the first within the area of litigation analysis.

CONCLUSION, LIMITATIONS, AND SCOPE FOR FUTURE RESEARCH

A number of studies have evidenced the benefits of artificial intelligence and machine learning methods within the legal domain. The current study extends upon this literature offering a unifying account, and detailing a new method of estimation not utilised within the legal domain previously. Through the use of a new specified variable capturing relative firm size it was identified that the status of the claimant as a natural person, defendant repeated status and relative firm size are identified as of significance to

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

litigation success. The perceptron and logit-perceptron offered deeper insights. The logit-perceptron is shown to achieve better predictive accuracy and better goodness of fit outcomes than the conventional regression methods employed within the extant literature.

Moreover, through the use of the perceptron and logit perceptron the variable relsize was found to be the most significant determinant of litigation success. The defendant status as repeat players was also found to be important employing the methods of Garson (1991) and Goh (1995). The findings appear to validate and further endorse the use of these contemporary non-parametric methods. The study does have some limitations worth noting specifically the limited available data and the absence of a singularly accepted method for determining robustness. Clearly, as more data becomes available, subsequent studies shall respond to the former issue in relation to the latter the use of Efron's R-Squared and the count R-Squared method appear to be the most valid and allow for the most coherent comparison of the parametric and non-parametric methods.

While economic methods are employed consistently within legal research, the use of non-parametric methods such as those within the domains of artificial intelligence and machine learning are relatively under utilised within the legal domain. The present study seeks to advise the next wave of interdisciplinary legal research, typified by applications of machine learning within the legal domain. We refer to this next wave as the third wave of legal empiricism in honour of Richardson & Hatfield (1989). We offer the conjecture that legal empiricism shall be strongly influenced by machine learning methodology, particularly as such technologies become more accessible to those within the legal domain. As this study makes evident, there is a dearth of research considering the relationship between law and machine learning. While this study responds directly to this shortage, further research is necessary. It would be worthwhile to consider the importance of judiciary member characteristics on judicial decision making as well as considering the legal domain beyond intellectual property. Such studies would serve to inform legal practice regulatory policy and our understanding of judicial decision making. Such studies would serve to inform the development of even more accurate and robust machine learning methods; and it is hoped that such studies are inspired and informed by the present study.

REFERENCES

Baker, L.A. (1996). Interdisciplinary Due Diligence: The Case for Common Sense in the Search for the Swing Justice. *Southern California Law Review, 70,* 187.

Epstein, L., Ho, D. E., King, G., & Segal, J. A. (2005). The Supreme Court During Crisis: How War Affects Only Nonwar Cases. *New York University Law Review, 80.*

Garson, G.D. (1991). Interpreting Neural-Network Connection Weights.

Goh, A.T.C. (1995). Back-Propagation Neural Networks for Modeling Complex Systems. *Artificial Intelligence in Engineering*, *9*(3), 143-151.

Gressman, E., & Stern, R. (2002). *Supreme Court Practice: For Practice in the Supreme Court of the United States* (8th ed.). Bureau of National Affairs.

Hunter, D. (1994). Looking for law in all the wrong places: Legal theory and legal neural networks. In *The relation with legal theory* (pp. 55-64).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hunter, D. (1999). Out of their minds: Legal theory in neural networks. *Artificial Intelligence and Law*, *7*(2-3), 129-151.

Enns, K. (1998). Can a California Litigant Prevail in an Action for Legal Malpractice Based on an Attorney's Oral Argument before the United States Supreme Court? *Duke Law Journal, 48*(1), 111-146.

Karnin, E.D. (1990). A Simple Procedure for Pruning Back-Propagation Trained Neural Networks. *IEEE Transactions on Neural Networks*, *1*(2), 239-242.

Martin, A.D., Quinn, K.M., & Epstein, L. (2005). The Median Justice on the US Supreme Court. *North Carolina L. Rev.*, *83*, 1275.

Maskus, K.E. (2002, August 6) Intellectual Property Issues for the United States and Japan: Disputes and Common Interests. In *Proceedings of the Issues and Options for the Multilateral, Regional, and Bilateral Trade Policies of the United States and Japan Conference.*

McAtee, A., & McGuire, K. T. (2007). Lawyers, Justices, and Issue Salience: When and How Do Legal Arguments Affect the US Supreme Court?. *Law & Society Review*, *41*(2), 259-278.

McGuire, K. T. (1995). Repeat players in the Supreme Court: The role of experienced lawyers in litigation success. *The Journal of Politics*, *57*(1), 187-196.

Minsky, M., & Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry*. Cambridge, Mass.: MIT Press.

Olden, J. D., & Jackson, D. A. (2002). Illuminating the "black box": a randomization approach for understanding variable contributions in artificial neural networks. *Ecological modelling*, *154*(1), 135-150.

Patterson, D., & Patterson, A. N. N. (1998). Theory and Applications. Singapore: Prentice Hall.

Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological review*, *65*(6), 386.

Rumelhart, D.E., Hinton, G.E., & Williams, R.J. (1988). Learning Internal Representations by Error Propagation. In A. Collins & E.E. Smith (Eds.), *Readings in Cognitive Science* (p. 399). San Mateo, CA: Morgan Kaufmann Publishers.

Sankar, A., & Mammone, R. J. (1991, July). Optimal pruning of neural tree networks for improved generalization. In *Proceedings of the Seattle International Joint Conference on Neural Networks IJCNN '91* (Vol. 2, pp. 219-224). IEEE.

Sheehan, R. S., Mishler, W., & Songer, D. R. (1992). Ideology, status, and the differential success of direct parties before the Supreme Court. *American Political Science Review*, *86*(2), 464-471.

Songer, D. R., & Sheehan, R. S. (1992). Who wins on appeal? Upperdogs and underdogs in the United States courts of appeals. *American Journal of Political Science*, *36*(1), 235-258.

Epstein, L., & Knight, J. (1997). The Choices Justices Make. Congressional Quarterly Inc.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Weigend, A. S., Rumelhart, D. E., & Huberman, B. A. (1990). Back-propagation, weight-elimination and time series prediction. In *Proceedings of the 1990 Connectionist Models Summer School* (Vol. 105). Morgan Kaufmann.

ENDNOTES

1 See Barker (1996)

2 See the recent CNBC article titled "Lawyers could be replaced by artificial intelligence"; http://www.cnbc.com/2017/02/17/lawyers-could-be-replaced-by-artificial-intelligence.html

3 Ibid

4 Ibid

5 For the full plenary session transcript visit; http://law.queensu.ca/how-will-artificial-intelligence-affect-legal-profession-next-decade

6 Ibid

7 W E Siebeck et al, 'Strengthening Protection of Intellectual Property in Developing Countries: A Survey of the Literature' (1990), World Bank discussion paper No.112, Washington D.C. The World Bank.

8 Burton M. Atkins, and Justin J. Green . 1976. "Consensus on the United States Courts of Appeals: Illusion or Reality?" American Journal of Political Science 20(11):735–48.

9 Nancy Ellen Billica, 'Just Leave it to the Courts: How, When, and Why Congress Abdicates the Legislative Power' (PhD Thesis, Harvard University, 1997).

10 Ibid.

11 Sisk, Gregory C. and Heise, Michael and Morriss, Andrew P., Charting the Influences on the Judicial Mind: An Empirical Study of Judicial Reasoning (November 1, 1998). New York University Law Review, Vol. 73, No. 5, 1998. Available at SSRN: https://ssrn.com/abstract=2447727

12 Albert Yoon Pensions, Politics, and Judicial Tenure: An Empirical Study of Federal Judges, 1869–2002, Am Law Econ Rev (2006) 8 (1): 143-180

13 Martin, Quinn, & Epstein (2005).

14 Martin, A.D., & Quinn, K.M. (2002). Dynamic Ideal Point Estimation via Markov Chain Monte Carlo for the US Supreme Court, 1953–1999. *Political Analysis, 10*(2), 134-153.

15 Martin et al, above n 11.

16 Epstein, Ho, King et al., 2005

17 Ibid.

18 Clearly there are numerous regression based methods employed within empirical judicial sciences research, and empirical legal research generally. The author does not assert that all studies employ the noted methods but rather the noted methods are broadly representative of the most commonly employed methods.

19 Megan Richardson and Gillian Hadfield, The second wave of law and economics, Federation press, 1999

20 For a full methodological review, see Dreiseitl, S. & Ohno-Machado, L. (2002, October). Logistic Regression and Artificial Neural Network Classification Models: A Methodology Review. *Journal of Biomedical Informatics*, *35*(5–6), 352-359.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

[21] Bellucci, Lodder and Zelznikow (1994)

[22] Oakley et. al. (1995)

[23] Ibid (1995)

[24] Bellucci (2001)

[25] D'Rosario (2016)

[26] D'Rosario (2016)

[27] D'Rosario (2017)

[28] The author is grateful for Matthew Sag's generosity in making this dataset available, and thus making the present study possible.

[29] Sag (2012), above n 2.

[30] Sag (2012), above n 2.

[31] A multilayer perceptron model is selected as an alternative to the radial basis function due to its superiority in estimation. Any potential gains in estimation efficiency associated with the radial basis method are offset by the losses in estimation accuracy.

[32] McCulloch, W.S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, *5*(4), 115-133.

[33] D'Rosario (2016) "Rosenblatt posited the perceptron model, framing the concept of the retina layer with distributing input values. The work of Rosenblatt owed much to the initial work of Hebb, and also McCulloch, and Pitts. The work of Marvin Minsky and Seymour Papert asserted a number of the limitations of the perceptron model, and their work brought about a diminution of interest in perceptron modelling. The key challenge evidenced in the perceptron model as posited by Rosenblatt was the step function approach that made neural network models problematic to train. However interest in the field was renewed with the advent of techniques to moderate the issues associated with neural network training. A revival of sorts occurred as a result of the work of Rumelhart, Hinton and Williams, whose seminal paper 'Learning Internal Representations by Error Propagation' offered resolution. Their work enabled the effective training for a multilayer neural network. The backpropagation training method and a viable algorithm was first asserted in Rumelhart, Hinton and Williams, and was the first practical method for training neural networks. Herein, a gradient descent backpropagation training method is used for efficiency."

[34] Within a biological neuron, the synaptic signals are received by the soma (cell body) and dendrites and transmitted through the axon. Within an artificial neuron the vector of the variable is passed from one, neuron in the input layer to another neuron with the assignment of synaptic weights (weighted values). Transfer occurs through the assistance of a transfer function, the synthetic axon.

[35] Minsky, M., & Papert, S. (1969). *Perceptrons: An Introduction to Computational Geometry.* MIT Press.

[36] Rumelhart, D.E., Hinton, G.E., & Williams, R.J. (1986). Learning internal representations by error propagation. In D.E. Rumelhart, & J.L. McClelland (Eds), *Parallel distributed processing: explorations in the microstructure of cognition* (Vol. 1, pp. 318-362). MIT Press

[37] Patterson, D.W. (1996). *Artificial neural networks: theory and applications*. Prentice Hall.

[38] It is necessary to discuss the criticisms of neural network methodology within legal research of which there appear to be few of a unique nature. Dan Hunter details the problems of using statistical sub-symbolic techniques in a symbolic field of research (D'Rosario, 2017a). At the outset it is apparent that such criticisms are relevant to all quantitative methods not merely machine learning. Hunter, is highly critical legal realist philosophy insofar as machine learning methods such as

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

perceptron models and artificial intelligence methods derive from them (D'Rosario,). It is evident that Hunters assertions are based on what Hunter considers to be inapt realist claims about the focus of network applications in his view and not the methods directly (D'Rosario, 2017). Hunter rightly and notably claims that useful applications are likely to be observed in latter studies (Hunter, 1994). It is hoped that the present study is such a study. It is also worth acknowledging that at the time that Hunter's paper article was authored, computing technologies did not enable the use of complex multi-layer perceptron models, nor was it possible to account for independent variable importance and as such the methods available now are far more useful in accommodating symbolic and sub symbolic languages as well as the open texture within legal research. Herein, as with D'Rosario (2016, 2017a) the author agrees with the claims of Hunter only to the extent that legal realists espouse that neural networks can provide normative assessments of legal outcomes and jurisprudence. No such claim is made herein. However, Neural networks methodology does enable a fuller analysis of judicial outcomes, jurisprudence and the importance of different variables in relation to such determinations.

[39] Consider the statements of Warner (1989), plausibly the most ardent proponent of this position, "While our language dictates a sequential description of the (legal reasoning) process, the process is in fact parallel. Many aspects of the problem resolution process are carried out simultaneously. The problem domain is defined by the initial statement of the problem. That initial problem is then resolved into a number of issues... the solution to which will be sought within the problem domain utilizing a sub symbolic paradigm that is not rule based…"
When accounting for even this staunch position, or when ascribing to the Hunter (1994) view position, the view of Warner does not coincide with the correct interpretation of legal reasoning as a process, neural networks nonetheless remain superior in methodological terms within legal research pertaining to legal reasoning; when compared with traditional regression methodology. This is because traditional regression models such as logit models are incapable of modelling dynamic nonlinear functions and as previously noted accounting for the interaction between variables (D'Rosario, 2016, 2017).

[40] Heaton, J. (2011). *Introduction to the Math of Neural Networks.* Heaton Research, Inc.

[41] Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, *65*(6), 386- 408 .

[42] Patterson, D.W. (1996). *Artificial neural networks: theory and applications*. Prentice Hall.

[43] For the initial theoretical propositions that served as the basis for the methodology see Hebb, D.O. (1952). *The organisation of behaviour: a neuropsychological theory*. Wiley.

[44] The multi-layer perceptron model shares a level of association with the logistic model, employing a logistic/sigmoid function. A more apt description however would be that the logistic model is a simple single layer perceptron model, that does not employ a second (layer) weighted aggregate input set.

[45] See Lothar & Philipps (1981), Hobson & Slee (1994) and Rose & Belew (1989)

[46] Heaton, J. (2011). *Introduction to the Math of Neural Networks.* Heaton Research, Inc.

This research was previously published in the International Journal of Strategic Decision Sciences (IJSDS), 8(2); edited by Madjid Tavana; pages 31-51; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 28
A Graphical Tool for the Creation of Behaviors in Virtual Worlds

Andrea Corradini
Design School Kolding, Denmark

Manish Mehta
Accenture Technology Lab, USA

ABSTRACT

Typically, the creation of AI-based behaviors for Non-Playing Characters (NPC) in computer games is carried out by people with specialized skills in the broad areas of design and programming. This book chapter presents Second Mind (SM), a digital solution that makes it possible for novice users to successfully author behaviors. In SM, authors can use a graphical interface to define the behaviors of virtual characters in response to interactions with players. Authors can easily endow NPCs with behavior capabilities that make them act in different possible roles such as e.g. shopkeepers, museum hosts, etc. A series of user tests with human subjects to evaluate the behavior authoring process shows that Second Mind is easy to understand and simplifies the process of behavior production.

INTRODUCTION

In order to increase player experience, the virtual characters that populate video games and virtual worlds should exhibit believable behaviors (Umarov, Mozgovoy, & Rogers, 2012). This does not apply only to the player characters, which are the characters that are controlled by the player, and as such are meant as more of a player's extension. This also applies to any other character that is controlled by the game, usually through artificial intelligence, such as townsfolk, vendors, enemies, allies, historical figures, bystanders, quest givers, etc. All these latter characters are referred to as non-player characters (NPCs) and are an integral part of the game. They exist to interact with the player characters, other NPCs, and the environment in a way that reflects their own distinct personalities and dialog capabilities.

DOI: 10.4018/978-1-7998-0951-7.ch028

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In the context of synthetic characters, believability is all about making the users believe they are interacting with a living agent, whose existence is consistent with the virtual world the character is situated in. In essence, believability is the ability to suspend the users' disbelief, by providing an illusion of life (Bates 1994).

Creating behaviors for NPC usually requires specialized skills in both design and programming well beyond the skills of the average person (Magerko, Laird, Assanie, Kerfoot, & Stokes, 2004). Hence, the intrinsic nature of the behavior generation process is restricted to the individuals who have a certain expertise in this specific area. To date, there is little understanding of how individuals without programming and design experience (whom we will refer to as "novice users" throughout this paper) can carry out the behavior authoring process.

The behaviors of NPCs are usually scripted and are automatically triggered by certain actions, events, or dialogue. In that sense, player characters play the game, while NPCs display some facet of the game and help to further the storyline making the game what it is (Gillies & Ballin, 2004). In early digital games, NPCs display only basic non-interactive dialogue capabilities. Monologues realized with screens of text, floating text, voiceovers, cutscenes, text clouds and other non-branching dialogue techniques made it possible for NPCs to convey only an immediate impression of their personality in reaction to or to initiate interaction with the player. However, this lack of interactivity turns the player into a mere passive consumer of content. More recent, advanced games allow for interactive dialogue where the player can engage in a compelling conversation with NPCs that results in witty and dramatic storylines that fit the context at hand. In those games, a conversation is typically modelled using dialogue trees (Adams, 2012) where the player is presented with and can choose from a limited list of interaction options in response to an NPC's action. Each choice affects the course of the game in a different way since each of them may result in a different response from the NPC. Moreover, the possibility to choose between different options makes it necessary for the game designer to customize the NPCs' behaviors by modifying their default scripts or by creating entirely new ones.

Branching the overall storyline puts the player in charge of learning about the game by engaging him in conversation with NPCs. In this way, uncovering the storyline becomes part of an increased overall player experience (Li, Thakkar, Wang, & Riedl, 2014). Besides engaging interactive narrative experiences (Yu & Riedl, 2014) with unfolding storylines, it is also necessary to have the tools that make it possible to design and create the appropriate behaviors. This is not just about making them funnier. It is also about creating more surprising and unpredictable behaviors within the game constraints so that two interactions with an NPC never turn out to be the same.

Digital games have much to gain from adopting artificial intelligence (AI) techniques to author complex behaviors. This is however, a difficult task for it requires a great deal of competence and resources. Usually, only the game creators or the people with specialized skills in both design and programming can carry out such a task. As a result, authoring behavior can even become a bottleneck in the process of creating the content displayed within digital worlds.

In this paper, we present Second Mind, a digital authoring tool solution that helps any players to populate virtual worlds with NPCs that can exhibit believable interactive behaviors. Our solution is based on a study that we carried out using paper prototypes to investigate how individuals without programming and design experience carry out the behavior authoring process. That study reveals some major limitations that we account for in the development of Second Mind. In Second Mind, novice users can easily design the behavior of graphical characters that impersonate different tradesmen such as shopkeepers, museum hosts, etc. in virtual worlds.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The rest of the paper is organized as follows. In the next section, we discuss related works and the challenges that behavior authoring poses to users, with particular focus on novice users. Thereafter, we present the results of the user study that we ran with 40 novice users to evaluate the feasibility of a paper-based prototype authoring system in a restricted context. We then detail the architecture of the digital authoring solution that we developed for novice users based on the findings discovered in running the user study. Eventually, we conclude with an outline of the work in process and our ideas for further improvement.

BACKGROUND AND RELATED WORK

Believable characters play an important role in many interactive stories and digital games (Rickel, Marsella, Gratch, Hill, Traum, & Swartout, 2002) because they can contribute to the overall narrative. Thus, the improvement and the control of their behaviors is propaedeutic to an increased game experience for players.

To ensure believability, virtual agents usually integrate multiple cognitive abilities of people such as language, gesture, emotion, and the problems associated with real-time navigation in and interaction with a simulated virtual world. Moreover, they also simulate many of the embodied and social aspects of human behavior as well as the need of autonomous behaviors, consistent with the character's personality, motivational state, and world context. Such an autonomy is usually implemented with an internal structure that allows the agent to have its own needs, motivation and interests. Based on these needs and interests the character can then dynamically select and generate goals that eventually result in autonomous behavior. An architecture to represent and manage internal driving factors in intelligent virtual character, using the concept of motivation, has been presented in (Avradinis, Panayiotopoulos, & Anastassakis, 2013). Because of the integration of human cognitive capabilities, virtual characters are gaining interest also as a methodological tool for studying and validating theories in human cognition (Gratch, Hartholt, Dehghani, & Marsella, 2013). At the same time several cognitive architectures have been put forward also for applications with virtual agents (Langley, Laird, & Rogers, 2009, Laird, 2008, Langley & Choi, 2006, Nuxoll & Laird, 2007, Sun 2007, Gluck 2010).

Common methods for authoring behaviors rely on heavy scripting, where a blend of dialogue scripts, hand-coded animation scripts, motion-captured characters, and hard-coded behaviors is used to exhibit a character's features. Usually, artists, designers and programmers must work together to detail a character-attitude and to portray the appropriate character's movements given a set of constraints. However, scripting is labor intensive, rigid because it does not adapt to all variations induced by interaction, and requires programming and design skills. An alternative approach is to use artificial intelligent (AI) techniques to adapt character behaviors in response to the different contexts that are induced by different interaction sessions. Despite its own inherent limitations, AI technology has been increasingly applied with success to digital game design and interactive narrative both to enhance interaction and to create novel gameplay scenarios (Riedl & Zook, 2013). AI techniques have been explored with diverse focus on character believability and expressivity in terms of e.g. synchronization of verbal and non-verbal modalities (Cassell, 2001, Corradini, Mehta, Bernsen, & Charfuelan, 2005), portrayal of emotion and personality (Allbeck & Badler, 2004, Aylett, Louchart, Dias, Paiva, & Vala, 2005), and so on.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The ease of the actual process of behavior creation has been given less relevance and explored by fewer researchers. The language ABL (Mateas & Stern, 2004) is an early effort in this direction. ABL is a Java-like programming language that can be used to manually script behaviors (Mateas & Stern, 2003, Gomes & Jhala, 2013). The production costs are high and the use of ABL requires advanced programming skills. Finite State Machines (FSMs) have also been used in digital games for many years as an alternative to scripting. However, FSMs can quickly become difficult to maintain and do not scale well with increasing game complexity (Schwab, 2009). Various FSMs modifications have been proposed to bypass part of these problems (Isla, 2005, Champandard, 2007). The use of AI techniques such as Layered Statechart-based AI (Dragert, Kienzle, & Verbrugge, 2012), Behavior Trees (Tomai & Flores, 2014), automated planning and scheduling strategies (Coman & Munoz-Avila, 2012, Kadlec, Toth, Cerny, Bartak, & Brom 2012, Kelly, Botea, & Koenig, 2008), has been explored as well. All these works require human intervention and a great deal of effort and skills to create and to debug the scripted behaviors. A recent attempt at eliminating manual scripting in favor of automatic script generation is presented in (Zhao & Szafron, 2014) where the content creation for cyclic behaviors is facilitated using a tiered behavior architecture model. This model features a scheduling algorithm to determine the objectives of virtual characters and to specify the roles that satisfy these objectives dynamically during game play. The results of a user study that compares the creation of manually scripted behaviors with the behaviors generated by an implementation of their model is also presented. Three metrics, and notably behavior completion time, behavior completeness, and behavior correctness, are used as a measure for assessing behavior reliability and designer efficiency.

Differently to most of the works presented previously, our solution provides a graphical user interface with different constraint types. While a few graphical interfaces have also been put forward to help people create scripts, these tools are aimed at story designers and at people who want to learn to program. Our implementation targets any person who wishes to create behaviors with no prerequisites in terms of skills and experience.

THE AUTHORING PROCESS AS A MENTAL ACTIVITY

Illustrative Example

Steven is an enthusiastic digital game player without any programming and design background. Steven is essentially an average end user and game player. He likes to buy interactive story games. Lately he has been playing an online game recently released that is based on the story "*And then there were none*" (Figure 1).

The story context is that the player has been invited to have dinner at a manor house. When the dinner is about to start, the player discovers that one of the guests has been murdered. After the introduction, the player is free to move around and interact with the other characters in the house. The goal of the game is to find out who committed the murder among the guests in the house.

After one of his many playing sessions, Steven wants to show off his creative skills by adding a set of simple scenes to the game in order to add more variability for other online players. Steven has a few scenes in mind that involve only characters with a particular personality, personal traits, emotions, and background. Authoring such experience is an example of the creative process that we wish to understand and enhance.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Using the traditional approach (i.e., hand-authoring scripts) would require a tremendous effort in authoring each one of the individual characters. Moreover, game companies usually provide toolkits to create only scenes/maps. The behavior creation is commonly carried out through scripts in a programming language, which restricts the process to people with a certain degree of programming skill. Steven does not have such a programming experience and thus he cannot add this kind of new content and experience to the game. If there was an environment that would allow more people like Steven to be involved in extending the game, it would provide them an opportunity to exercise their creative skills and make them available for other players. Consequently, the gameplay experience would improve given a larger base of users creating content and sharing it with each other.

AI Behavior Construction by Novice Users

Novice users bring little or no programming experience to the behavior construction task. However, they can employ a combination of real-world experiences, experience with other software, computers and electronic gadgets in general to understand the construction task. One of the fundamental capabilities of human cognition is the ability to relate something new to something similar that was already encountered and/or experienced.

Familiar interaction design approaches that allow users to create correspondences with a more familiar domain (Heckel, 1984) have been used in product design and have been shown to make a product or interface easier to understand. The key question, therefore boils down to identifying the right interaction design approach that can be employed as part of the presentation of the authoring process that can help novice users understand the construction process.

An AI behavior construction process requires an underlying representation formalism and vocabulary that is used to represent behaviors. Representation formalisms have been a keenly studied topic within the AI field for quite a long time. In fact, several representational formalisms have been put forward, ranging from predicate calculus through frames and scripts to semantic networks and more (Brachman

Figure 1. A snapshot of the commercial interactive story "And then there were none"



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

& Levesque, 2004, Russell & Norvig, 2010). An AI scientist, who is familiar with the vocabulary of the underlying representational formalism, is adept at translation between the domain that is being represented and the formal language used to describe it. A novice user, however, would not have the necessary background to use the formal representational language. Several criteria can be utilized to classify different AI representation formalisms: expressivity is one of the most common among them. Clearly, the more expressive the formalism is, the easier it is to represent something with it. In our context though, the key requirements for the representation formalism and the corresponding vocabulary, are understandability and ease of use. Our idea is to forsake some level of expressivity for better understandability and ease of use. We contend that the representation vocabulary should speak the user's language and use words, phrases, and concepts familiar to the user, rather than terms more familiar to an AI scientist.

Assuming that a design solution that will enable novice authors to construct AI behaviors should provide the right interaction design approach and representation vocabulary to support them in the authoring activity, the main question that we have to address is what is a design solution that enables novice users to easily create AI behaviors for NPCs in a virtual world?

For the sake of simplicity, we focus on the construction of particular kinds of AI behaviors situated in a particular domain. Namely, we confine the scope of AI behavior creation as sequence of actions to NPCs impersonating tradesmen in a virtual world.

Representation Vocabulary

As a starting point in the process to identify the basic vocabulary and terminologies, we used an existing reactive architecture that we presented in (Corradini, Mehta, Bernsen, & Charfuelan, 2005, Zang, Mehta, Mateas, & Ram, 2007). The key property of such kind of architectures is their reactivity and responsiveness response to time-critical events, such as emergencies in a robot assistance domain, or dodging and avoiding unexpected obstacles in real-world navigation domains. This reactivity to time critical events makes them ideally suited for creating interactive avatars. Examples of reactive architectures include the Reactive Action Packages (Firby, 1987), the Gapps formalism (Kaelbling & Rosenschein, 1990), the Procedural Reasoning System (Georgeff & Lansky, 1987), the Soar architecture (Laird, 2012), and many more. A notable example of reactive architecture specifically developed for interactive believable characters is the A Behavior Language (ABL) (Mateas & Stern, 2004). ABL has been successfully used to author the central characters Trip and Grace for the interactive drama Facade (Knickmeyer & Mateas, 2005).

Based on our existing architecture and taking inspiration from the architectures previously mentioned, we identified the following terms:

- **Scene:** Provides a social context under which the behaviors of the virtual characters can be described. An example of a scene might be "a boy meeting a girl at the park for a date". Scenes may contain one or more characters, and can be of arbitrary length.
- **Behavior:** Represents the activities that the avatar can accomplish. Behaviors can be considered similar to plans in a traditional AI planning paradigm. An example of a behavior within the scene "a boy meeting a girl at the park for a date", the boy avatar could perform a "flirt with a girl" behavior. In general, behaviors consist of a sequence of other behaviors and actions so that "flirt with a girl" may contain e.g. the behavior "smile to a girl".

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- **Action:** Represents the lowest level behaviors that can be performed by the avatar in the world and cannot be further decomposed. They represent the elementary units on which all behaviors are further built on.
- **Personality:** Represents a character's overall way of conducting himself. The personality is manifested in the way a specific avatar goes about performing the scenes through the kind of behaviors he can pursue, the way he carries out these behaviors and the way he conducts the basic actions.
- **Emotion:** Represents the avatar feelings in response to things happening in the environment.
- **Percept:** Represents basic events that the avatar detects from the environment.
- **Like/Dislike:** Represents things that the virtual characters likes or dislikes.

Behaviors are further associated with the following set of conditions:

- **Pre-Condition:** Detects when the behavior needs to be activated and carried out.
- **Success-Condition:** Detects when a behavior has been successful. This information is needed to identify when a behavior needs or not to be pursued again.
- **Non-Progress-Condition:** Detects when a behavior is not progressing well. If a behavior is not progressing well, the character would try to repair the situation by different means such as pursuing a different behavior to achieve the same or a similar behavior and/or have an emotional reaction in response.
- **Failure-Condition:** Detects if the behavior has failed to achieve its purpose.
- **Advancing-Condition:** Keep a check on whether the behavior is progressing well or not.

Success-condition and failure-condition are mutually exclusive. As such, they can be thought of as opposite to each other. They are used separately to provide more flexibility to the user to use either of them as needed. A similar reasoning went into providing the Non-progress-condition and the Advancing-condition as separate entities. As part of the authoring process, these conditions can be created by using basic sensory information through mathematical and logical operations that can be detected from the virtual world.

In our framework, the behaviors are defined bottom-up meaning that each behavior must resolve into a series of basic actions. Hence, during the authoring activity, a novice user must define each behavior in terms of its constituent actions. Likewise, in order for the framework to detect the conditions associated with a behavior, these conditions need to be defined in terms of the basic set of events that can be detected within the virtual world.

Next, we discuss the authoring process that we envisioned for the authors to carry out.

The Envisioned Authoring Process

The envisioned authoring activity is based on our previous works on the generation of interactive characters using reactive planning languages (Corradini, Mehta, Bernsen, & Charfuelan, 2005, Corradini, Mehta, & Charfuelan, 2005, Zang, Mehta, Mateas, & Ram, 2007). It can be carried out following three main steps.

The first step involves the creation of scenes with a hierarchical behavior structure. Scenes may contain one or more characters, and be of arbitrary length. As an example, we can consider a scene called "boy flirting with a girl in a bar". The given scene consists of two avatars: one representing a girl and

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

the other representing a boy. The boy might be assigned a personality trait called "flirtatious" and "enthusiastic". The female character may be assigned a personality trait of being "shy". For such a scene, the behavior tree is shown in Figure 2 (top). The top-down structure linking the lowest level behaviors in terms of basic actions that can be carried out in the virtual world is also defined by the authors as shown in Figure 2 (bottom) for the behavior "Introduce". Such a behavior is made up of three actions in sequence, namely, a parameterized "Talk" action (the parameter being e.g. "I am Max"), followed by a "Smile" action, and concluded with another "Talk" action with parameter "how are you".

The second step in the authoring process consists in defining the various conditions for the behaviors as mentioned in the previous section. Various conditions for two example behaviors are shown in the Figure 3 (left). For example, the behavior "Go to a bar" is associated with an activation condition that the boy wants to chill out.

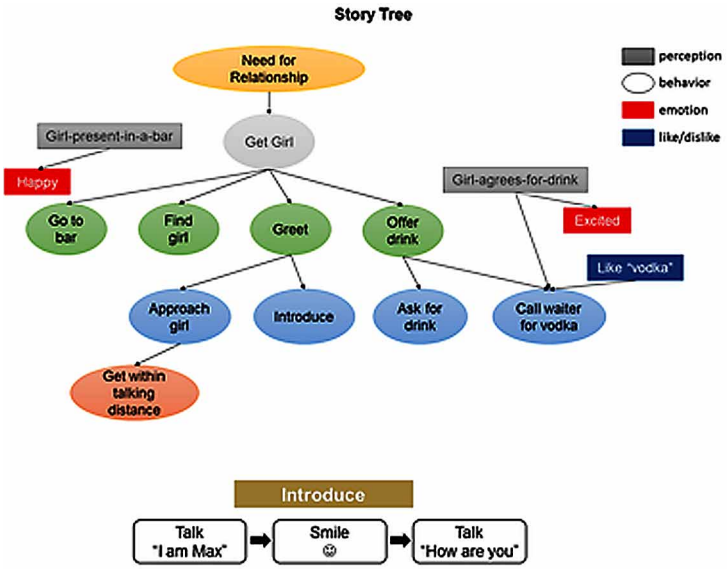
The third and final step involves linking higher level conditions with the basic level percepts available in the virtual world. For example, the condition "Girl is present in a bar" can be detected by performing an inside operation on location of bar and location of the girl as shown in Figure 3 (right). An exhaustive list of basic actions and percepts that are available to authors can be found in (Mehta, 2012).

User Study: A Paper-Based Authoring Prototype

We conducted a user study in order to identify if novice users would be able to understand the representational vocabulary and carry out the authoring process following the steps listed above.

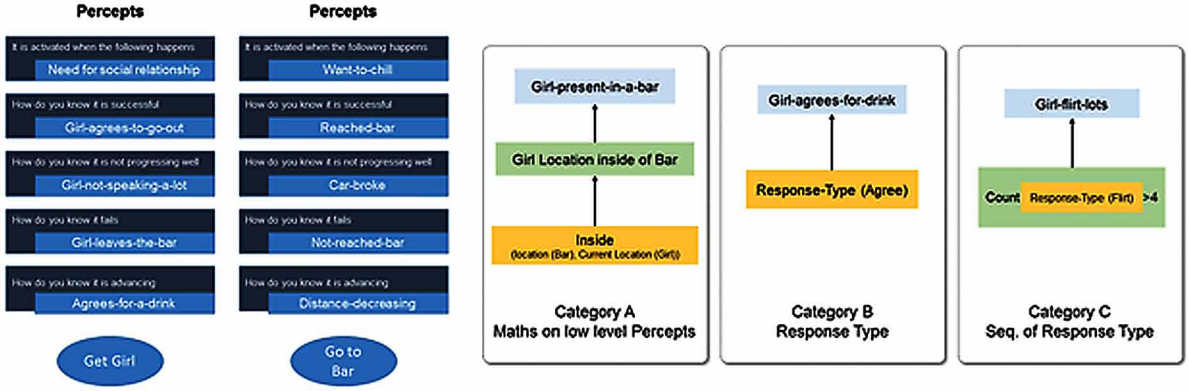
The user study was carried out using paper prototypes. We were particularly interested in figuring out which parts of the representation vocabulary are difficult for users to understand, what parts of the authoring process are difficult to conceptualize, and possibly what are the points at which the authoring process becomes tedious.

Figure 2. (top) Graphical representation of the behavior tree structure created for a scene; (bottom) the behavior "Introduce" as linked sequence of basic actions available in the virtual world



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. (left) Example of conditions for two behaviors; (right) percepts from the virtual world



We recruited forty participants (referred to as P1 through P40): 10 females and 30 males. The participants were pre-screened to make sure they did not have any programming and design background. The participants ranged across multiple races, education levels, and ages (from 22 to 40 with an average age of 28). Each subject was paid 10 USD per hour rounded up to the nearest hour. Players signed a consent form at the beginning of the session. The authoring task was described to participants through example forms (as shown in Figure 2 and Figure 3).

Each participant was given a brief description of the scene setting. Two such scene settings were provided to the participants. The first scene was about a boy taking a girl to prom, while the second scene involved an armed robber who holds someone up. The list of basic actions and basic events that can be detected from the environment was also provided to the user. The participants were then asked to conduct the authoring activity using the paper-based prototype interface forms.

These forms were blanked out versions of the filled forms similar to those shown in Figure 2 and Figure 3.

During the authoring session, a researcher observed the player's interaction with the paper prototype and took notes. This included any unusual things such as "unable to fill a condition", or "having difficulty creating behaviors". On an average, a complete player interaction lasted approximately one hour. At the end of their interaction, subjects were asked to fill out a quantitative questionnaire. Moreover, an open-ended interview about the authoring experience was conducted with each participant immediately thereafter.

To perform the analysis, we transcribed both the user responses from the interviews and the actions observed during the authoring process. We analyzed the data obtained with focus on the qualitative analysis of the authoring process. To that extent, as qualitative analysis method we chose the well-known Grounded Theory (Strauss & Corbin, 1990). Using grounded theory principles, the transcriptions were processed, first in order to take open-ended notes and second to highlight more salient ideas expressed by each subject.

We generated codes, such as "User not able to fill precondition" or "User not able to complete scene creation process", and listed multiple user quotes for each category in an Excel spread-sheet. We then grouped related codes: for instance, all codes that referred to the problem of creating behaviors were initially grouped together into related concepts. We went through an iterative process of describing each concept and conceptually linking them to each other. We formed a single hierarchy of the concepts, and

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

organized the same concept with a paragraph to describe each concept along with the top two to three quotes for it. This hierarchical organization of concepts served as an outline for the qualitative findings.

We also analyzed the results from the forms filled out by the users during the authoring process to identify any further issues during the authoring process and to get a better understanding of issues that emerged from the interview session.

The analysis of data collected during the study provided several key insights that can summarized into the following categories.

Conceptualization of Behaviors

The data analysis indicated that users had difficulty conceptualizing the behavior creation process as a hierarchical, top-down process. To that regards, P16 states:

..When you think of a scenario where X needs to happen, you can think of a few things that would happen, like A would happen and then B could happen and so on. Here (using the authoring interface) I am thinking of things in a different way. I am thinking like A would happen and then in order to accomplish A, I need to do these two other things and then I would have to break them down, that is kind of hard..

When users visualize the creation process, it seems very hard for them to think in a top down fashion and they seem more used to thinking in a sequential fashion. Subject P23 stated:

..When I do things, I do them in a very sequential fashion, it would be hard for me to plan out like I would do this and then it would be doing these other two things which would then involve other activities..

The difficulty of creating hierarchical structures is confirmed by high ratings extrapolated from the quantitative questionnaires. To the question of difficulty level in creating behaviors, our subjects provided an average rating of 4.03 (with 5 indicating "very difficult").

Feedback on Created Scene

Several users reported that they had difficulty seeing the results of their authoring process. Creating the behaviors on a paper prototype did not give them the opportunity to evaluate what they just created. Subjects mentioned that it would have been easier for them to create behaviors if the system could "play back a simulation of the behavior" (P13) so that they could "verify or correct their creations"(P15). Subject P2 mentioned:

..I was imagining how things would go when someone does these things, I could see that there could be some things which might not work exactly, I would say it would have been easier if things played out in a game, that would give good ideas on what works and what doesn't..

Users mentioned that they would want to see how "*the scene they had created would work in actual practice*" (P10). They felt that the scene "*might come out funny or some gesture might not work*" (P24). Looking at the avatar acting out the created behaviors would be very helpful in finding how well things would work in actual practice. Subject P19 mentioned:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

..Seeing my creation being acted out would be great. That would tell me how well it works from the performance perspective..

During the authoring process, it is very hard for users to take into account all the possible situations that might occur in reality during gameplay. This task becomes even harder for novice users as they are new to carrying out the design task. Thus, it is very likely that while designing behaviors, the users would forget to take into account certain situations. It would be helpful if the users were able to observe a simulation of the created behaviors to identify immediately any issues with them.

Behavior Conditions

Some parts of the behavior condition creation process were left empty by the subjects. They had difficulty identifying things like the "progress-condition" (30 out of 40 users) and "failure-condition" (34 out of 40 users), the "advancing condition" (28 out of 40 users), and the "success-condition" (26 out of 40 users). Figure 4 (top) shows an example behavior condition form filled by one user who did not fill some of these conditions and provided only a very vague description for some of the others. Subject P24 mentioned:

..When my behavior needs to be activated was clear. I know when I want my character to perform something. Some of the other things were difficult like success condition. I may not be able to know whether my behavior is successful or not by looking at things in the environment..

On the question of difficulty level in creating advancing condition, users provided an average rating of 4.1 (with 5 indicating "very difficult"). Similarly, for difficulty level in creating failure, success and progress condition users provided an average rating of 4.28, 4.35 and 4.55, respectively (with 5 always indicating "very difficult").

On the one hand, a simple way to address this problem would be to reduce the set of conditions. On the other hand, we are aware that this also takes away some of the expressivity of the system. Having novice users as target group, we decided to reduce the set of conditions in order to make the authoring process more accessible for them.

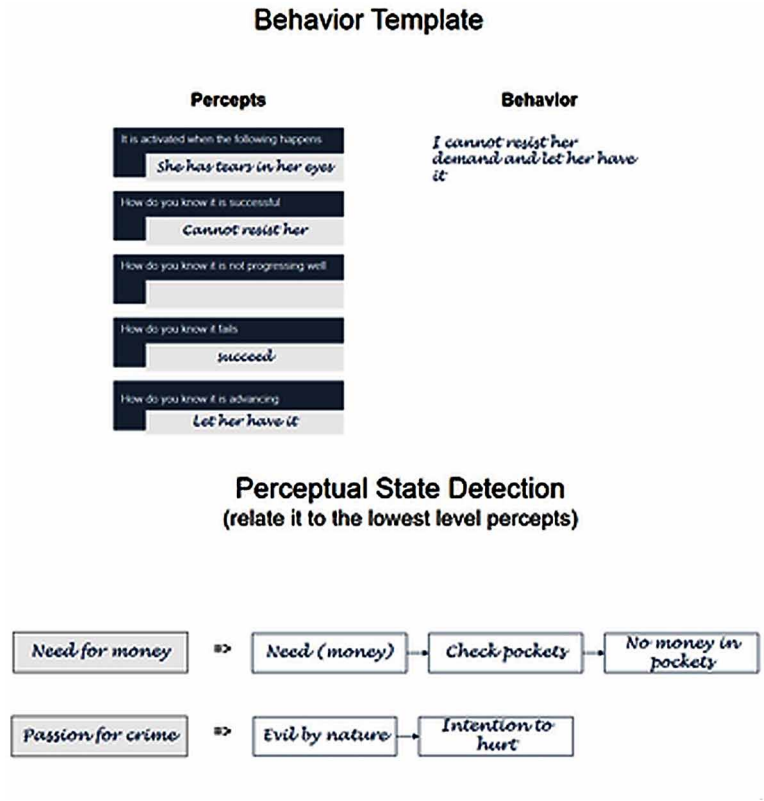
Additionally, subjects had difficulty linking behavior conditions with lowest level sensory information that was available in the environment. As P32 mentioned:

..Sometimes I could say when my behavior fails. But in general I wasn't sure if I could see these things in terms of what that was available in the environment..

The perceptions could become very abstract and users had difficulty relating it to sensory information that can be detected from the environment. Figure 4 (bottom) shows an example where "need for money" is broken down to "money in pocket" which a subject could not relate to the sensory information describing "no money in someone's pocket". Apart from that, in order to concretely connect the sensory information obtained from the world to higher level behavior conditions, mathematical operators needed to be defined in certain situations and contexts. Most of the times, subjects skipped this process. Subject P6 reported that "..creating mathematical formulas was the most difficult part..".

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. (top) A paper form related to behavior conditions as filled out by one user; (bottom) example of behavior condition that users had difficulty in connecting to sensory information from the environment



Similarly, subject P25 explicitly expressed that authors need technical skills in order to carry out this step:

..Defining it mathematically in terms of the lowest level things that are happening in the environment was unclear. I would think someone who has the technical skills would be able to understand it better than me..

One possibility would be to remove the notion of complex conditions defined using mathematical operators over the basic sensory information could be to provide a base set of parameterized conditions that can be used individually or combined together to result in more complex conditions.

SECOND MIND: A SYSTEM FOR BEHAVIOR AUTHORING

Rationale and Context

The previous user study explored a design solution that would enable novice users to successfully author AI behaviors. That study provided, among others, the following key insights and design directions for creating the authoring interface:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

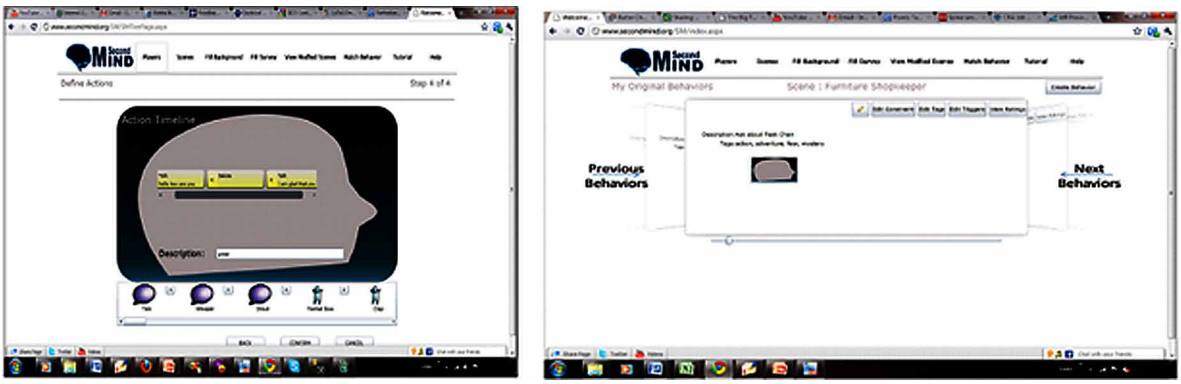
- The process to link behavior constraints to lower level percepts must be as simple as possible; the conditions themselves must be limited, clearly understandable and should be expressed with an easy-to-use formalism.
- The authoring task should be presented as a sequential process and not rely on deep hierarchical structures.
- Authors need continuous feedback on the authoring activity so that they can see the results of their ongoing process and thus promptly identify possible problems.

On one hand, reducing the set of conditions takes away some of the expressivity of the system, but on the other hand, it allows keeping the authoring process simple. Conditions help the avatar identify when, what, and how a certain behavior must be carried out.

We express conditions using what we refer to as triggers. We provide a set of triggers that can be used individually or combined together to define more complex constraints. An example of trigger is "User Clicks on (button label)", which can be used when the player wants the avatar to perform a behavior in response to a button click ("button label" is a placeholder that must be provided during the authoring process). Many other triggers are also provided such as "Avatar Close to (object name)", "Avatar Touches (object name)", "Player Close to (object name)", "Player Touches (object name)", "Behavior Ends (behavior name)", "Pause after Behavior Ends (behavior name, time)", "Run Event (event name)", and so on. Triggers can be composed using logical operators. For instance an "Trigger And(A, B)" can be used to specify a behavior for the avatar to perform when both a given trigger A and trigger B are specified.

In order to present the behavior creation task as a sequential process, we applied the notion of timelines. Timelines indicate a left-to-right sequential sequence, where each event occurs in chronological order. Various storyboard-authoring tools like Adobe Director and Apple's Garageband Timelines as well as AI-based interactive story authoring tools like Scribe (Medler & Magerko, 2006), U-Create (Sauer, Osswald, Wielemans, & Stifter, 2006) and Viper (Agamanolis & Bove, 2003) employ timelines as an interaction design approach. We use timelines to represent the ordering of story segments in a left-to-right sequence while also allowing story segments to be re-ordered on the timeline. In order to help the user conceptualize the task concretely, the authoring activity along the timeline is presented in a mind-like shape (Figure 5).

Figure 5. (left) Authoring of a greeting behavior as a left-to-right sequence of the basic actions "talk (text)", "salute" and "talk (text)" on the timeline; (right) Behaviors are shown in a sequence using the cover-flow approach



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Eventually, in order to comply with the need to provide a continuous feedback to the user, we connect our authoring environment to an online virtual world named Second Life (Rymaszewski, 2009) where avatar's behaviors can be executed and displayed graphically in 3D. In Second Life, users can create, design and build their own clothes, buildings, artifacts, landscape, pets and much more. We chose Second Life as platform to display 3D-based user-generated content because it has a large user base, it allows for easy connection to external controllers through an API, and it provides a rich set of basic actions for the avatar in the form of emotions and physical actions. Using our authoring interface Second Mind, authors can create behaviors, execute them in Second Life, and immediately see the effects of their authored artifacts.

The Authoring Process in Second Mind

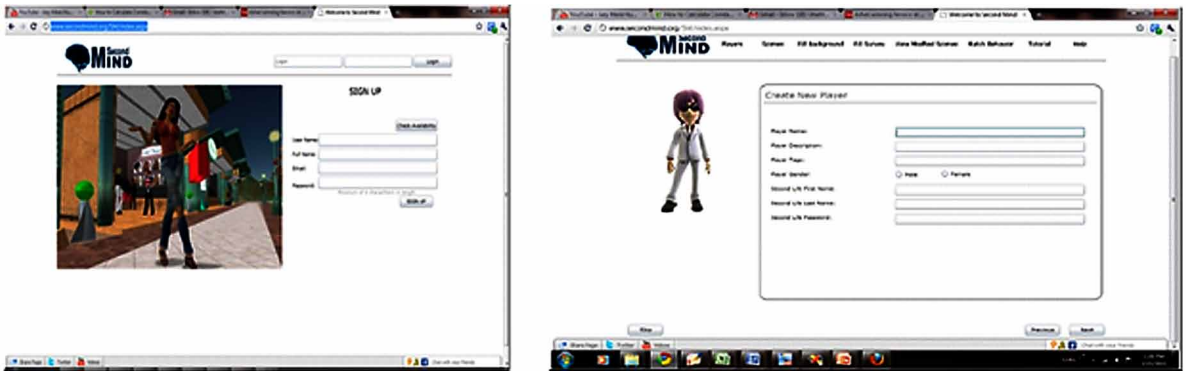
In order to use Second Mind, the user needs to create an account where to store, among other things, his/her Second Life login credentials. Once the account is created, the user is guided through a tutorial, which explains the terminology used in the authoring interface (like triggers, scenes, behaviors, actions, etc.). The user then can define one or more characters in Second Mind and connect each of them to a different avatar in Second Life.

At this stage, the user is ready to create a scene and its corresponding behaviors. A scene provides a social context such as, e.g. a furniture shop where the avatar acts as a shopkeeper whose goal is to sell furniture items. The author creates a description of the scene by entering some information like the name, a description and by selecting one of the characters created in earlier steps (see Figure 6). The user also associate constraints a limited set of constraints. For instance, a constraint to apply to the scene in the shopkeeper context could be "likeable" or "salesman skills".

Beside the contextual constrains, all scenes are associated by default with two additional general constraints, namely "overall experience" and "avatar performance". The author uses a set of sliders to specify matching values on a 0 to 5 point scale for each of the constraints s/he wants to set. Moreover, the user can associate semantic tags to the scene. Now, avatar behaviors can be added within the scene.

Similarly, to the scene creation process, the author can create a complex behavior from a list of simple elementary behaviors. To define a new behavior, the author needs to enter a name that uniquely identifies it followed by the list of behaviors that make it up and by the specification of the constraints that characterize the behavior as a whole.

Figure 6. (left) Screenshot of the web version of Second Mind; (right) initial steps for the creation of a scene



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. The player interaction with an avatar can be initiated by clicking on the device carried by the avatar on his left shoulder; clicking on the device causes a graphical interface to pop up with various scenes to choose from



Using a set of sliders, the author can specify the values for the constraints. For example, in order to have an "enthusiastic" and "pushy" shopkeeper for a "greeting behavior", the user has to select first the corresponding values for those two constraints that give an idea of the avatar personality. These values influence the way (e.g. speed, spatial extension, etc.) a certain behavior is rendered. The user then can assign triggers to the behavior in order to specify the context in which it should be carried out by the avatar. The process is iterative, in the sense that each of the steps can be applied to define an even more complex behavior as well as to each single behaviors that are used to define the complex one. During the creation of a specific behavior, the author can also define a constraint for which it is the avatar that initiates the interaction. Once a behavior is created, the user can immediately play it out in Second Life and test its 3D graphical representation.

In order to interact with an avatar in Second Life, the player must click on an attachment on the avatar's left shoulder (see Figure 7). Upon clicking, the player is presented with a GUI containing buttons labeled with the names of the scenes that have been previously authored for the avatar (see Figure 7). Each of these buttons represents one "User Clicks on (button label)" trigger that is associated with the behaviors of a scene. To interact with the avatar, the player must simply click on a button.

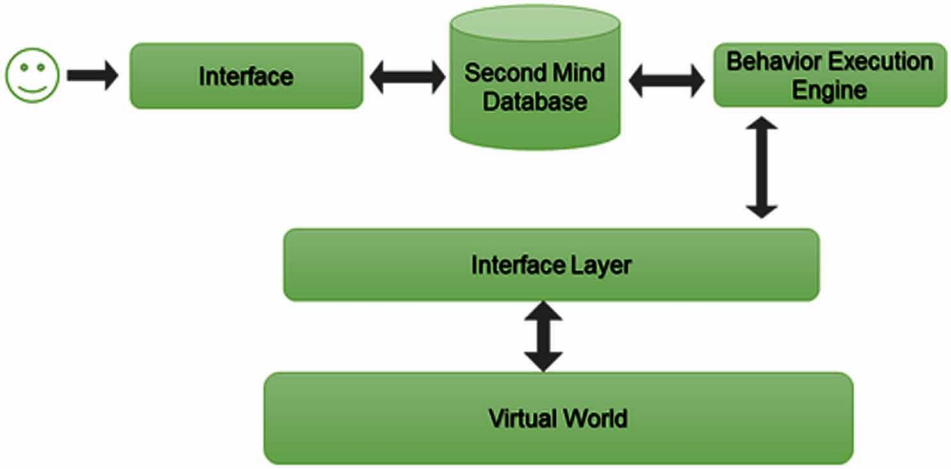
System Architecture

With a basic understanding of the authoring activity, we can delve into the system architecture. As shown in Figure 8, Second Mind architecture has four core components and one application-specific component. The four core components are:

- **Second Mind Database:** A MS-SQL server based database that stores scenes, players, behavior elements such as triggers, basic action, etc.
- **Behavior Execution Engine:** An execution engine that loads the behavior elements and executes them in the external world. The behavior execution engine connects to the external world through the interface layer that is specific to the world that Second Mind is connecting to.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Overview of Second Mind overall architecture



- **Graphical Interfaces:** Provides the interfaces for creating user account, scenes and behaviors as described in the previous section.
- **Behavior Recommendation Engine:** Provides suggestions during the behavior-authoring task.

In addition, in order to apply Second Mind to a given graphical world, one application-specific components is required: are:

- **Interface Layer:** The architecture of Second Mind is domain-independent and can be connected to different virtual world simulations via a custom interface layer. The interface layer provides the ability to connect Second Mind to the simulation platform, such as Second Life or other virtual world. The interface layer requires the definition of percepts, actions, and other application-specific elements along with the APIs to connect Second Mind to the simulation.

Behavior Execution Engine

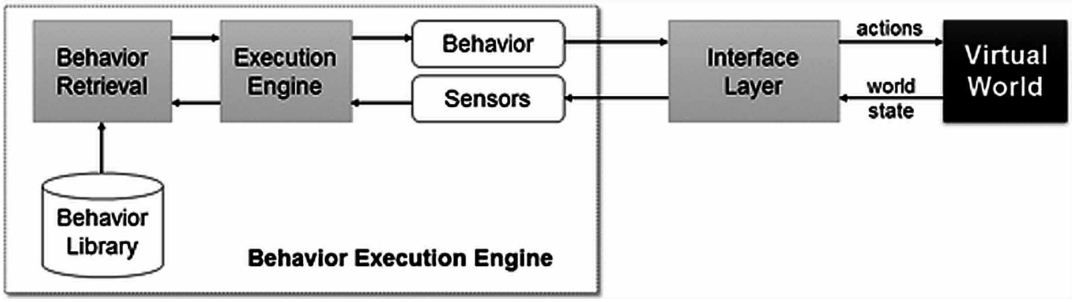
The behavior execution engine (see Figure 9) loads the behaviors corresponding to the currently active scene, identifies the behavior(s) to carry out based on the current perceptual information from the world, and sends the information about these behaviors to the graphical world for execution through the interface layer.

The behavior execution engine can be correspondingly broken down into three main modules: interface:

- **Behavior Library:** It is part of the Second Mind database that holds behavior elements for a specific application domain. For example, if the application domain were Second Mind in an educational setting, the library would contain personalities corresponding to teachers, students, etc. If instead the application domain were a household robot, the library would contain personality elements corresponding to cooking food, cleaning the house, etc.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Overview of behavior execution engine architecture



- **Behavior Retrieval:** It is responsible for loading all the behaviors corresponding to the active scene from the behavior library. The active scene is the scene that the player selected when s/he was presented with all possible authored scenes associated with a certain avatar.
- **Execution Engine:** It receives the currently active domain independent sensors from the interface layer and checks the triggers for the behaviors corresponding to the active scene to see if any of the behaviors were activated. It also takes the current active behavior and sends it through the interface layer to the virtual world for the avatar to perform it.

Interface Layer

The Interface layer parses the domain dependent perceptual state into a set of domain independent sensors (see Figure 10). The domain dependent perceptual state from Second Life contains the following items:

- **Object Info:** Contains the information about various objects in the environment such as their current position, rotation, type (whether an avatar or a passive object), region to which the object belongs, etc.
- **User Button Click Info:** Contains information regarding button clicks made by the player while interacting with the avatar.
- **Event Info:** Contains the information regarding various events happening in Second Life such as the player moving around in the environment, the player touching an object, etc. Objects in the environment can send various events to communicate with the avatar. If there are behaviors that have "Event happens (event name)" as triggers, avatar can react to events with name "event name".

The world specific information is parsed into a set of domain independent sensors that correspond to various triggers. This information is then sent to the execution engine, which checks the sensors against the triggers corresponding to the current set of behaviors to identify if any new behavior(s) has become activated. The interface layer module also receives the action corresponding to the active behavior and sends it to the virtual world. The interface layer provides an abstraction layer and maps the action name received from the execution engine to virtual world specific action name. For example, an action "*laugh*" in Second Mind is mapped to actions that express "*express_laugh*", which is the specific name used by Second Life virtual world. Table 1 shows the mapping of a few action names in Second Mind with the corresponding names in the virtual world Second Life.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. The interface layer parses the world state in second life into domain independent sensors

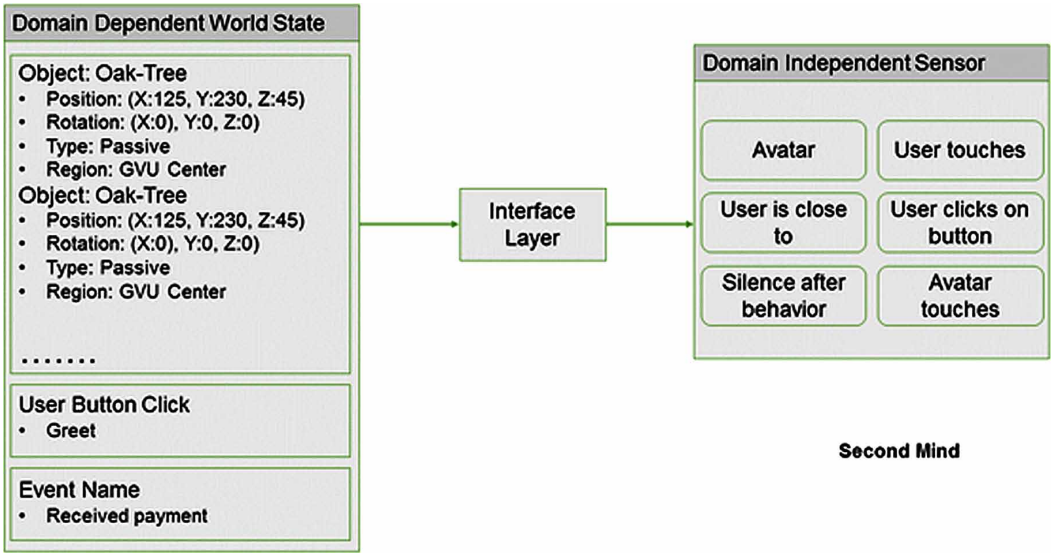


Table 1. Action names in SM and corresponding virtual world action names in second life

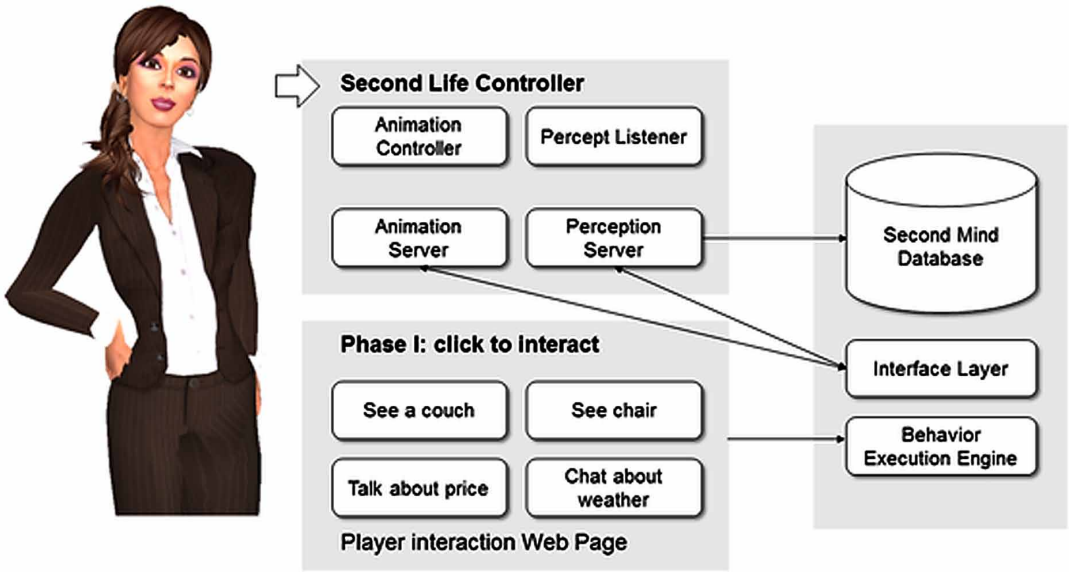
Domain Independent Action Name	Virtual World Action Name
Laugh	express_laugh
Shrug	express_shrug_emote
Bow	avatar_hold_bow

Initially, in order to connect Second Mind to Second Life, we modified an open source version of Second Mind clients. We soon realized though that not all Second Mind users are willing to installed a modified version of Second Life as well. Therefore, we instead created scripts in Second Life's proprietary scripting language named Linden Scripting Language (LSL). Those scripts make it possible to send and to receive information through the Second Mind interface layer from/to Second Life for controlling the avatar with any genuine Second Life client (Figure 11).

The collection of scripts is packaged in a component called Second Mind Controller and distributed through the Second Life marketplace. Second Life marketplace is a social web portal from Second Life where all the items from various vendors are listed and available for purchase. Second Mind controller is listed on the Second Life marketplace and is available for free. In order to have their avatar controllable through Second Mind, players need to download it from the Second Life marketplace. Purchasing items from Second Life marketplace is a common activity performed by Second Life users. Once users have bought the items from Second Life marketplace. Second Mind controller is downloaded automatically onto their avatars, which they can then attach to their avatar in Second Life.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Scripts organization for the connection of second mind and second life



Behavior Recommendation

During the behavior authoring process, once the name of the behavior is entered, the user is presented with a list of recommended behaviors that he can use as a starting point. The behaviors are suggested based on similarity between the behavior name specified for the currently authored behavior and existing behaviors in the database. The semantic similarity is calculated using a WordNet-based algorithm (Simpson & Dao, 2014).

CONCLUSION AND FUTURE DIRECTIONS

In this paper, we have presented the results from a user study conducted to evaluate a paper-based prototype version of the authoring activity. The findings from the user study, made it possible for us to identify and modify early design flaws that we accounted for and incorporated into the current version of the digital authoring tool, which employs Second Life (Rymaszewski, 2009) as the concrete virtual world where the behaviors can be executed. To address the issue of having concrete feedback on the authoring activity, the authoring interface was connected with a concrete virtual world, Second Life, so that author can view the created behaviors and any possible problems with them. In order to simplify the steps required to link behavior conditions to lower level percepts, we provided to the user a base set of triggers that can be used individually or combined together to make more complex triggers. In order to present the authoring task as a sequential one, instead of deep hierarchies, we used a left to right sequential action creation approach represented on a timeline. Eventually, we outlined the authoring steps necessary to create and run a behavior along with the system architecture and its four core components.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

We evaluated the Second Mind approach to check whether it can be understood and easily used by non-programmers. In order to evaluate our approach, we invited 65 subjects to create a certain set of behaviors for an avatar (e.g. act as a shopkeeper) on a Second Life island using the authoring system. We divided the subjects into three broad categories: User Interface designers, non-programmers, and programmers (with experience in Java). We transcribed data from the subjects' responses to interviews as well as from the notes taken while observing subject actions during the authoring process. To analyze this vast quantity of data collected from the user study, we adopted a combination of qualitative and quantitative analysis. Specifically, we employed again a Grounded Theory (Strauss & Corbin, 1990) approach to perform the qualitative analysis. A detailed report on the evaluation of the system is outside the scope of this paper. However, for the sake of completeness, the evaluation results show that non-programmers were able to easily create behaviors using the scaffolding provided by Second Mind. Moreover, no relevant statistical difference was noticed in the quantitative measures used to asses authoring skills among the different subject categories. This seems to indicate that designers, programmers, and non-programmers perform very similarly while authoring behaviors with Second Mind.

There is much work left to do in Second Mind, some of which is research, some of which is in development. There are still numerous open questions that form possible avenues of future directions for this research. Authoring support through demonstrating behaviors is something we are considering. Learning from demonstration is also a point we would like to investigate. Human learning is often accelerated by observing a task being performed or attempted by someone else. These capabilities of the human brain are also evident in computer games where players go through a process of training and imitating experienced players. These results have inspired researchers in artificial intelligence to study learning from imitation techniques. By observing an expert's actions, new behaviors can quickly be learnt and then successfully performed.

ACKNOWLEDGMENT

We would like to thank Ms. Maria Kikidou for the integral role she played in helping us with the editing process and the creation of the images in this book chapter.

REFERENCES

Adams, E. (2012). *Fundamentals of game design* (2nd ed.). New Riders.

Agamanolis, S., & Bove, V. M. (2003). Viper: A Framework for Responsive Television. *IEEE MultiMedia*, *10*(1), 88–98. doi:10.1109/MMUL.2003.1218260

Allbeck, J., & Badler, N. (2004). Representing and parameterizing behaviors. In H. Prendinger & M. Ishizuka (Eds.), *Life-Like Characters: Tools, Affective Functions and Applications*. Springer. doi:10.1007/978-3-662-08373-4_2

Avradinis, N., Panayiotopoulos, T., & Anastassakis, G. (2013). Behavior believability in virtual worlds: Agents acting when they need to. *SpringerPlus*, *2*(1), 1–11. doi:10.1186/2193-1801-2-246 PMID:23853745

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Aylett, R. S., Louchart, S., Dias, J., Paiva, A., & Vala, M. (2005). FearNot! - An experiment in emergent narrative. *Proceeding of the International Conference on Intelligent Virtual Agents (IVA '05)*.

Bates, J. (1994). The role of emotion in believable agents. *Communications of the ACM*, *37*(7), 122–125. doi:10.1145/176789.176803

Brachman, R., & Levesque, H. J. (Eds.). (2004). *Knowledge Representation and Reasoning*. Morgan Kaufmann.

Cassell, J. (2001). Embodied conversational agents: Representation and intelligence in user interfaces. *AI Magazine*, *22*(3), 67–83.

Champandard, A. (2007). Behavior trees for next-gen AI. *Proceedings of the Game Developers Conference*.

Coman, A., & Munoz-Avila, H. (2012). Plan-based character diversity. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE'12)*.

Corradini, A., Mehta, M., Bernsen, N.-O., & Charfuelan, M. (2005). Animating an interactive conversational character for an educational game system. *Proceedings of IUI*. ACM Press. 10.1145/1040830.1040872

Corradini, A., Mehta, M., & Charfuelan, M. (2005). Interacting with an animated conversational agent in a 3d graphical setting. *Proceedings of Workshop on Multimodal Interaction for the Visualization and Exploration of scientific data at ICMI*.

Dragert, C., Kienzle, J., & Verbrugge, C. (2012). Statechart-based game AI in practice. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE'12)*.

Firby, R. J. (1987). An investigation into reactive planning in complex domains. *Proceedings of the AAAI*.

Georgeff, M. P., & Lansky, A. L. (1987). Reactive reasoning and planning. *Proceedings of the AAAI*.

Gillies, M., & Ballin, D. (2004). Integrating Autonomous Behavior and User Control for Believable Agents. *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems*.

Gluck, K. A. (2010). Cognitive architectures for human factors in aviation. In E. Salas & D. Maurino (Eds.), *Human Factors in Aviation* (2nd ed.; pp. 375–400). New York, NY: Elsevier. doi:10.1016/B978-0-12-374518-7.00012-2

Gomes, P. F., & Jhala, A. (2013). AI authoring for virtual characters in conflict, *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE'13)*.

Gratch, J., Hartholt, A., Dehghani, M., & Marsella, S. C. (2013). Virtual Humans: A new Toolkit for Cognitive Science Research. *Cognitive Science*.

Heckel, P. (1984). *The Elements of Friendly Software Design*. Alameda, CA: Sybex Inc.

Isla, D. (2005). Handling complexity in the Halo 2 AI. *Proceedings of the Game Developers Conference*.

Kadlec, R., Toth, C., Cerny, M., Bartak, R., & Brom, C. (2012). Planning is the game: action planning as a design tool and game mechanism. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE'12)*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kaelbling, L. P., & Rosenschein, S. J. (1990). Action and planning in embedded agents. *Robotics and Autonomous Systems*, *6*, 35–48.

Kelly, J. P., Botea, A., & Koenig, S. (2008). Off-line planning with hierarchical task networks in video games. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE'08)*.

Knickmeyer, R., & Mateas, M. (2005). Preliminary Evaluation of the Interactive Drama Façade. *Proceedings of CHI 2005*. ACM Press.

Laird, J. E. (2008). Extending the Soar cognitive architecture. *Proceedings of the Artificial General Intelligence Conference*.

Laird, J. E. (2012). *The Soar Cognitive Architecture*. MIT Press.

Langley, P., & Choi, D. (2006). A unified cognitive architecture for physical agents. *Proceedings of the twenty-first AAAI Conference on Artificial Intelligence*. Boston: AAAI Press.

Langley, P., Laird, J. E., & Rogers, S. (2009). Cognitive architectures: Research issues and challenges. *Journal of Cognitive Systems Research*, *10*(2), 141–160. doi:10.1016/j.cogsys.2006.07.004

Li, B., Thakkar, M., Wang, Y., & Riedl, M. O. (2014). Data-driven alibi storytelling for social believability. *Proceedings of the Foundations of Digital Games Workshop on Social Behavior in Games*.

Magerko, B., Laird, J. E., Assanie, M., Kerfoot, A., & Stokes, D. (2004). AI characters and directors for interactive computer games. *Proceedings of the sixteenth innovative applications of artificial intelligence conference*.

Mateas, M., & Stern, A. (2003). Facade: an experiment in building a fully-realized interactive drama. *Game Developer's Conference: Game Design Track*.

Mateas, M., & Stern, A. (2004). A Behavior Language: Joint Action and Behavioral Idioms. In H. Prendinger & M. Ishizuka (Eds.), Life-like Characters. Tools, Affective Functions and Applications. Springer.

Medler, B., & Magerko, B. (2006). Scribe: A general tool for authoring interactive drama. *Proceedings of the International Conference on Technologies for Interactive Digital Storytelling and Entertainment*. 10.1007/11944577_14

Mehta, M. (2012). *Construction and Adaptation of AI Behaviors for Computer Games*. (PhD thesis). Georgia Tech.

Nuxoll, A. M., & Laird, J. E. (2007). Extending cognitive architecture with episodic memory. *Proceedings of the twenty-second AAAI Conference on Artificial Intelligence*.

Rickel, J., Marsella, S., Gratch, J., Hill, R., Traum, D., & Swartout, W. (2002). Toward a new generation of virtual humans for interactive experiences. *IEEE Intelligent Systems*, *17*.

Riedl, M. O., & Zook, A. (2013). AI for game production. *Proceedings of the IEEE Conference on Computational Intelligence in Games*.

Russell, S. J., & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach* (3rd ed.). Prentice Hall.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Rymaszewski, M. (2009). *Second Life: The Official Guide* (2nd ed.). Sybex Inc.

Sauer, S., Osswald, K., Wielemans, X., & Stifter, M. (2006). U-create: Creative authoring tools for edutainment applications. *Proceedings of the International Conference on Technologies for Interactive Digital Storytelling and Entertainment*. 10.1007/11944577_16

Schwab, B. (2009). *AI Game Engine Programming* (2nd ed.). Course Technology.

Simpson, T., & Dao, T. (2014). *WordNet-based semantic similarity measurement*. Retrieved on March 03, 2015 from http://www.codeproject.com/KB/string/semanticsimilaritywordnet.aspx

Strauss, A., & Corbin, J. (1990). Basics of Qualitative Research: Grounded Theory Procedures and Techniques. *Sage (Atlanta, Ga.)*.

Sun, R. (2007). The importance of cognitive architectures: An analysis based on CLARION. *Journal of Experimental & Theoretical Artificial Intelligence*, *19*(2), 159–193. doi:10.1080/09528130701191560

Tomai, E., & Flores, R. (2014). Adapting in-game agent behavior by observation of players with learning behavior trees. *Proceedings of the International Conference on the Foundations of Digital Games*.

Umarov, I., Mozgovoy, M., & Rogers, P. C. (2012). Believable and effective AI agents in virtual worlds: Current state and future perspectives. *International Journal of Gaming and Computer-Mediated Simulations*, *4*(2), 37–59. doi:10.4018/jgcms.2012040103

Yu, H., & Riedl, M. O. (2014). *Personalized interactive narratives via sequential recommendation of plot points*. IEEE Transactions on Computational Intelligence and Artificial Intelligence in Games.

Zang, P., Mehta, M., Mateas, M., & Ram, A. (2007). Towards runtime behavior adaptation for embodied characters. *Proceedings of the IJCAI*, 1557-1562.

Zhao, R., & Szafron, D. (2014). Using cyclic scheduling to generate believable behavior in games. *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE'14)*.

This research was previously published in Integrating Cognitive Architectures into Virtual Character Design edited by Steve DiPaola, Michael Nixon, Jeremy Owen Turner, and Ulysses Bernardet; pages 65-93; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 29
Contour Reconstruction:
2D Object Modeling

Dariusz Jacek Jakóbczak
Technical University of Koszalin, Poland

ABSTRACT

The method of Probabilistic Features Combination (PFC) enables interpolation and modeling of high-dimensional N data using features' combinations and different coefficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function. Functions for γ calculations are chosen individually at each data modeling and it is treated as N-dimensional probability distribution function: γ depends on initial requirements and features' specifications. PFC method leads to data interpolation as handwriting or signature identification and image retrieval via discrete set of feature vectors in N-dimensional feature space. So PFC method makes possible the combination of two important problems: interpolation and modeling in a matter of image retrieval or writer identification. Main features of PFC method are: PFC interpolation develops a linear interpolation in multidimensional feature spaces into other functions as N-dimensional probability distribution functions.

INTRODUCTION

Probabilistic modeling is still developing branch of the computer science: operational research (for example probabilistic model-based prognosis) (Lorton, Fouladirad & Grall, 2013), decision making techniques and probabilistic modeling (Pergler & Freeman, 2008), artificial intelligence and machine learning. There are used different aspects of probabilistic methods: stochastic processes and stochastic model-based techniques, Markov processes (Cocozza-Thivent, Eymard, Mercier & Roussignol, 2006), Poisson processes, Gamma processes, a Monte Carlo method, Bayes rule, conditional probability and many probability distributions. In this chapter the goal of probability distribution function is to describe the position of unknown points between given interpolation nodes. Two-dimensional curve (opened or closed) is used to represent the data points.

DOI: 10.4018/978-1-7998-0951-7.ch029

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

So problem statement of this chapter is: how to reconstruct (interpolate) missing points of 2D curve having the set of interpolation nodes (key points) and using the information about probabilistic distribution of unknown points. For example the simplest basic distribution leads to the easiest interpolation – linear interpolation. Apart from probability distribution, additionally there is the second factor of proposed interpolation method: nodes combination. The simplest nodes combination is zero. Thus proposed curve modeling is based on two agents: probability distribution and nodes combination.

Curve interpolation (Collins, 2003) represents one of the most important problems in mathematics and computer science: how to model the curve (Chapra, 2012) via discrete set of two-dimensional points (Ralston & Rabinowitz, 2001)? Also the matter of shape representation (as closed curve-contour) and curve parameterization is still opened (Zhang & Lu, 2004). For example pattern recognition, signature verification or handwriting identification problems are based on curve modeling via the choice of key points. So interpolation is not only a pure mathematical problem but important task in computer vision and artificial intelligence. The chapter wants to approach a problem of curve modeling by characteristic points. Proposed method relies on nodes combination and functional modeling of curve points situated between the basic set of key points. The functions that are used in calculations represent whole family of elementary functions with inverse functions: polynomials, trigonometric, cyclometric, logarithmic, exponential and power function. These functions are treated as probability distribution functions in the range [0;1]. Curve interpolation represents one of the most important problems in mathematics and computer science: how to model the curve via discrete set of two-dimensional points? Also the matter of shape representation (as closed curve-contour) and curve parameterization is still opened. For example pattern recognition, signature verification or handwriting identification problems are based on curve modeling via the choice of key points. So interpolation is not only a pure mathematical problem but important task in computer vision and artificial intelligence. The monograph wants to approach a problem of curve modeling by characteristic points. Proposed method relies on nodes combination and functional modeling of curve points situated between the basic set of key points. The functions that are used in calculations represent whole family of elementary functions with inverse functions: polynomials, trigonometric, cyclometric, logarithmic, exponential and power function. These functions are treated as probability distribution functions in the range [0;1]. Significant problem in machine vision and computer vision is that of appropriate 2D shape representation and reconstruction. Classical discussion about shape representation is based on the problem: contour versus skeleton. This monograph is voting for contour which forms boundary of the object. Contour of the object, represented by successive contour points, consists of information which allows us to describe many important features of the object as shape coefficients. 2D curve modeling and generation is a basic subject in many branches of industry and computer science, for example in the cad/cam software. The representation of shape can have a great impact on the accuracy and effectiveness of object recognition. In the literature, shape has been represented by many options including curves, graph-based algorithms and medial axis to enable shape-based object recognition. Digital 2D curve (open or closed) can be represented by chain code (Freeman's code). Chain code depends on selection of the started point and transformations of the object. So Freeman's code is one of the method how to describe and to find contour of the object. Analog (continuous) version of Freeman's code is the curve α-s. Another contour representation and reconstruction is based on Fourier coefficients calculated in discrete Fourier transformation (DFT). These coefficients are used to fix similarity of the contours with different sizes or directions. If we assume that contour is built from segments of a

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

line and fragments of circles or ellipses, hough transformation is applied to detect contour lines. Also geometrical moments of the object are used during the process of object shape representation. Contour is also applied in shape decomposition. Many branches of medicine, industry and manufacturing are looking for methods connected with geometry of the contour. Why and when should we use MHR and PNC methods? Interpolation methods and curve fitting represent so huge problem that each individual interpolation is exceptional and requires specific solutions. PNC method is such a novel tool with its all pros and cons. The user has to decide which interpolation method is the best in a single situation. The choice is yours if you have any choice. Presented method is such a new possibility for curve fitting and interpolation when specific data (for example handwritten symbol or character) starts up with no rules for polynomial interpolation. This book consists of two generalizations: generalization of previous MHR method with various nodes combinations and generalization of linear interpolation with different (no basic) probability distribution functions and nodes combinations. The method of Probabilistic Nodes Combination (PNC) enables interpolation and modeling of two-dimensional curves using nodes combinations and different coefficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function, also inverse functions. This probabilistic view is novel approach a problem of modeling and interpolation. Computer vision and pattern recognition are interested in appropriate methods of shape representation and curve modeling. PNC method represents the possibilities of shape reconstruction and curve interpolation via the choice of nodes combination and probability distribution function for interpolated points. It seems to be quite new look at the problem of contour representation and curve modeling in artificial intelligence and computer vision. Function for γ calculations is chosen individually at each curve modeling and it is treated as probability distribution function: γ depends on initial requirements and curve specifications. PNC method leads to curve interpolation as handwriting modeling via discrete set of fixed knots. So PNC makes possible the combination of two important problems: interpolation and modeling. Nodes are treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PFC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. Two-dimensional object is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function. Functions for γ calculations are chosen individually at each data modeling and it is treated as 2-dimensional probability distribution function: γ depends on initial requirements and features' specifications. PFC method leads to data interpolation as handwriting or signature identification and image retrieval via discrete set of feature vectors in 2-dimensional feature space. So PFC method makes possible the combination of two important problems: interpolation and modeling in a matter of image retrieval or writer identification. PFC interpolation develops a linear interpolation in multidimensional feature spaces into other functions as two-dimensional probability distribution functions. Future works are going to applications of PFC method in biometric recognition, computer vision and artificial intelligence.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

BACKGROUND

Shape Representation and Curve Reconstruction

An important problem in machine vision and computer vision (Ballard, 1982) is that of appropriate shape representation and reconstruction. Classical discussion about shape representation is based on the problem: contour versus skeleton. This chapter is voting for contour which forms boundary of the object. Contour of the object, represented by contour points, consists of information which allows us to describe many important features of the object as shape coefficients (Tadeusiewicz & Flasiński, 1991). In the chapter contour is dealing with a set of curves. Curve modeling and generation is a basic subject in many branches of industry and computer science, for example in the CAD/CAM software.

The representation of shape has a great impact on the accuracy and effectiveness of object recognition (Saber, Yaowu & Murat Tekalp, 2005). In the literature, shape has been represented by many options including curves (Sebastian & Klein, 2003), graph-based algorithms and medial axis (Liu & Geiger, 1999) to enable shape-based object recognition. Digital curve (open or closed) can be represented by chain code (Freeman's code). Chain code depends on selection of the started point and transformations of the object. So Freeman's code is one of the method how to describe and to find contour of the object. An analog (continuous) version of Freeman's code is the curve α-*s*. Another contour representation and reconstruction is based on Fourier coefficients calculated in Discrete Fourier Transformation (DFT). These coefficients are used to fix similarity of the contours with different sizes or directions. If we assume that contour is built from segments of a line and fragments of circles or ellipses, Hough transformation is applied to detect contour lines. Also geometrical moments of the object are used during the process of object shape representation (Choraś, 2005).

Nowadays methods apply mainly polynomial functions, for example Bernstein polynomials in Bezier curves, splines (Schumaker, 2007) and NURBS (Rogers, 2001). But Bezier curves don't represent the interpolation method and cannot be used for example in handwriting modeling with key points (interpolation nodes). In comparison PNC method with Bézier curves, Hermite curves and B-curves (*B-splines*) or NURBS one unpleasant feature of these curves has to be mentioned: small change of one characteristic point can result in unwanted change of whole reconstructed curve. Such a feature does not appear in proposed PNC method. Numerical methods for data interpolation are based on polynomial or trigonometric functions, for example Lagrange, Newton, Aitken and Hermite methods. These methods have many weak sides (Dahlquist & Bjoerck, 1974) and are not sufficient for curve interpolation in the situations when the curve cannot be build by polynomials or trigonometric functions. Also there exists several well established methods of curve modeling, for example shape-preserving techniques (Dejdumrong, 2007), subdivision algorithms (Dyn, Levin & Gregory, 1987) and others (Kozera, 2004) to overcome difficulties of polynomial interpolation, but probabilistic interpolation with nodes combination seems to be quite novel in the area of shape modeling. Proposed 2D curve interpolation is the functional modeling via any elementary functions and it helps us to fit the curve during the computations.

This chapter presents novel Probabilistic Nodes Combination (PNC) method of curve interpolation. This chapter takes up new PNC method of two-dimensional curve modeling via the examples using the family of Hurwitz-Radon matrices (MHR method) (Jakóbczak, 2007), but not only (other nodes combinations). The method of PNC requires minimal assumptions: the only information about a curve is the set of at least two nodes. Proposed PNC method is applied in curve modeling via different coefficients: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

tan, arc cot or power. Function for PNC calculations is chosen individually at each interpolation and it represents probability distribution function of parameter $\alpha \in [0;1]$ for every point situated between two interpolation knots. PNC method uses two-dimensional vectors (x,y) for curve modeling-knots $p_i = (x_i,y_i) \in \boldsymbol{R}^2$ in PNC method, $i = 1,2,\ldots n$:

1. PNC needs 2 knots or more ($n \geq 2$);
2. If first node and last node are the same ($p_1 = p_n$), then curve is closed (contour);
3. For more precise modeling knots ought to be settled at key points of the curve, for example local minimum or maximum and at least one node between two successive local extrema.

Condition 3 means for example the highest point of the curve in a particular orientation, convexity changing or curvature extrema. So this chapter wants to answer the question: how to interpolate the curve by a set of knots (Jakóbczak, 2010)?

Algorithm of PNC interpolation and modeling (1) looks as follows:

Step 1: Choice of knots p_i at key points.
Step 2: Choice of nodes combination $h(p_1,p_2,\ldots,p_m)$.
Step 3: Choice of distribution $\gamma = F(\alpha)$ (continuous or discrete).
Step 4: Determining values of α: $\alpha = 0.1, 0.2\ldots0.9$ (nine points) or $0.01, 0.02\ldots0.99$ (99 points) or others.
Step 5: The computations.

These five steps can be treated as the algorithm of PNC method of curve modeling and interpolation. Without knowledge about the formula of curve or function, PNC interpolation has to implement the coefficients γ, but PNC is not limited only to these coefficients. Each strictly monotonic function F between points (0;0) and (1;1) can be used in PNC modeling.

HANDWRITTEN CHARACTER AND SHAPE MODELING

Issues

Curve knots $p_1 = (0.1;10), p_2 = (0.2;5), p_3 = (0.4;2.5), p_4 = (1;1)$ and $p_5 = (2;5)$ are used in some examples of PNC method in handwritten character modeling. Figures 1-8 represent PNC as MHR interpolation (Jakóbczak, 2011) with different γ. Points of the curve are calculated with no matrices ($N = 1$) and $\gamma = \alpha$ in example 1 and with matrices of dimension $N = 2$ in examples 2-8 for $\alpha = 0.1, 0.2,\ldots,0.9$.

Example 1: (refer to Figure 1)

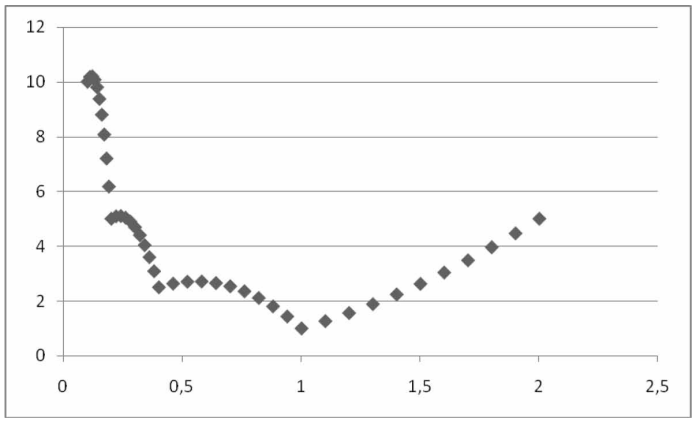
PNC curve interpolation for $\gamma = \alpha$ and

$$h(p_1, p_2) = \frac{y_1}{x_1} x_2 + \frac{y_2}{x_2} x_1 .$$

For $N = 2$ (examples 2 – 8) MHR version (Jakóbczak, 2011) as PNC method gives us:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. PNC character modeling for nine reconstructed points between nodes



$$h(p_1,p_2,p_3,p_4)=\frac{1}{x_1^{\,2}+x_3^{\,2}}(x_1x_2y_1+x_2x_3y_3+x_3x_4y_1-x_1x_4y_3)$$
$$+\frac{1}{x_2^{\,2}+x_4^{\,2}}(x_1x_2y_2+x_1x_4y_4+x_3x_4y_2-x_2x_3y_4)$$

Example 2: (refer to Figure 2)
PNC sinusoidal interpolation with $\gamma = sin(\alpha \cdot \pi/2)$.
Example 3: (refer to Figure 3)
PNC tangent interpolation for $\gamma = tan(\alpha \cdot \pi/4)$.
Example 4: (refer to Figure 4)
PNC tangent interpolation with $\gamma = tan(\alpha^s \cdot \pi/4)$ and $s = 1.5$.

Proposed method, called Probabilistic Nodes Combination (PNC), is the method of 2D curve interpolation and extrapolation using the set of key points (knots or nodes). Nodes can be treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PNC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. Two-dimensional curve is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function. Novelty of this book consists of two generalizations: generalization of previous MHR method with various nodes combinations and generalization of linear interpolation with different (no basic) probability distribution functions and nodes combinations. Computer vision needs suitable methods of shape representation and contour reconstruction. One of them, invented by the author and called method of Hurwitz-Radon Matrices (MHR), can be used in representation and reconstruction of shapes of the objects in the plane. Proposed method is based on a family of Hurwitz-Radon (HR) matrices. The matrices are skew-symmetric and possess columns composed of orthogonal vectors. 2D shape is represented by the set of successive nodes. It is shown how to create the orthogonal and discrete OHR operator and how to use it in a process of shape

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Sinusoidal modeling with nine reconstructed curve points between nodes

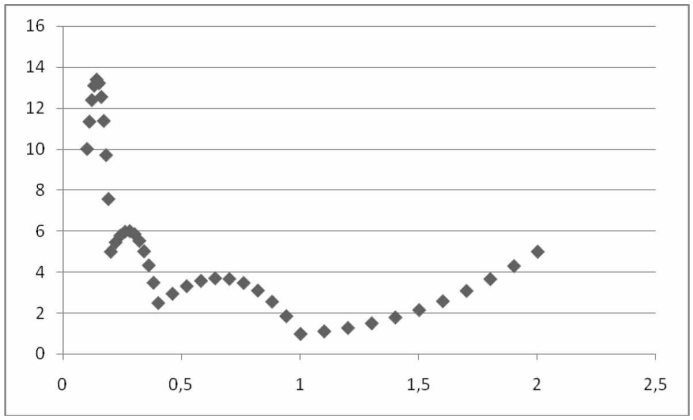


Figure 3. Tangent character modeling with nine interpolated points between nodes

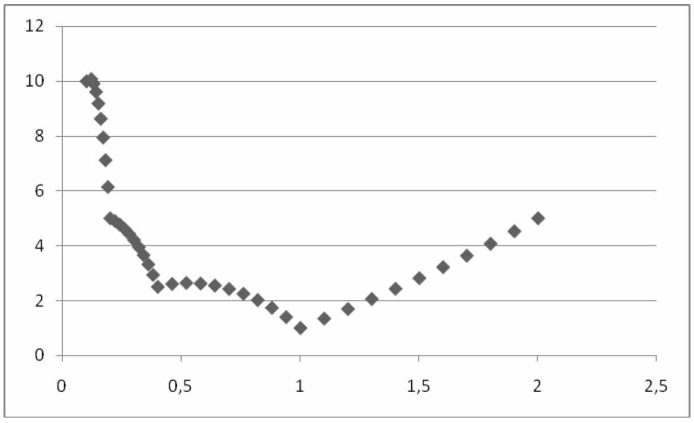
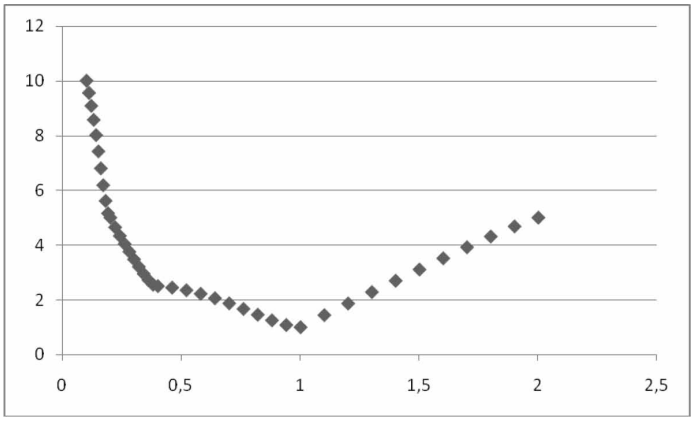


Figure 4. Tangent curve modeling with nine recovered points between nodes



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

representation and reconstruction. Then MHR method is generalized to Probabilistic Nodes Combination (PNC) method.

Example 5: (refer to Figure 5)
PNC tangent curve interpolation for $\gamma = tan(\alpha^s \cdot \pi/4)$ and $s = 1.797$.
Example 6: (refer to Figure 6)
PNC sinusoidal interpolation with $\gamma = sin(\alpha^s \cdot \pi/2)$ and $s = 2.759$.
Example 7: (refer to Figure 7)
PNC power function modeling for $\gamma = \alpha^s$ and $s = 2.1205$.
Example 8: (refer to Figure 8)
PNC logarithmic curve modeling with $\gamma = log_2(\alpha^s + 1)$ and $s = 2.533$.

These eight examples demonstrate possibilities of PNC curve interpolation and handwritten character modeling for key nodes in MHR version. And here are other examples of PNC modeling (but not MHR):

Figure 5. Tangent symbol modeling with nine reconstructed points between nodes

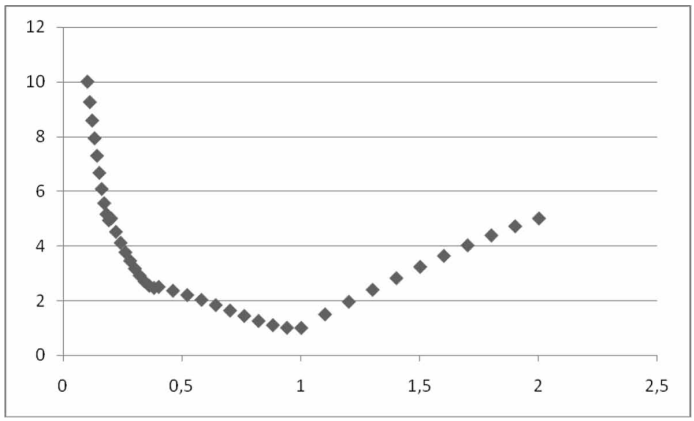
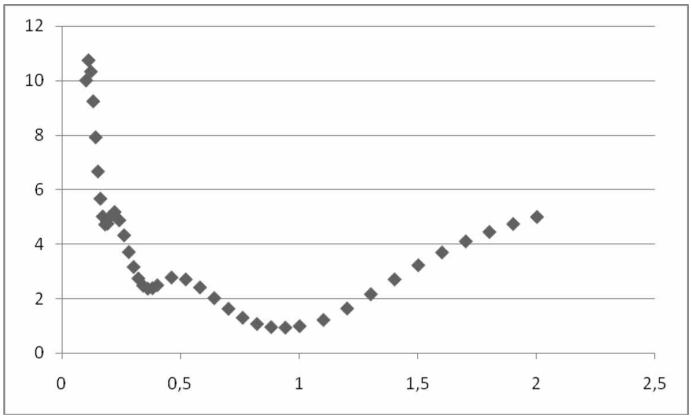


Figure 6. Sinusoidal modeling with nine interpolated curve points between nodes



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Power function curve modeling with nine recovered points between nodes

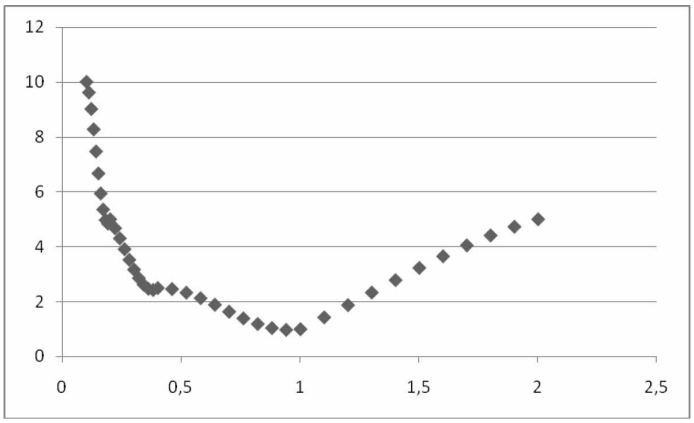
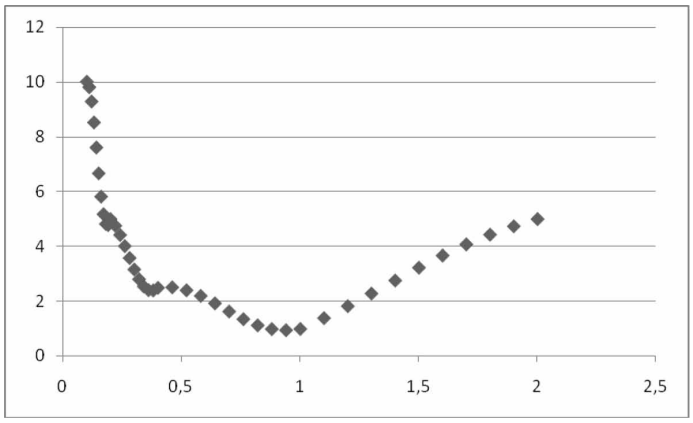


Figure 8. Logarithmic character modeling with nine reconstructed points between nodes



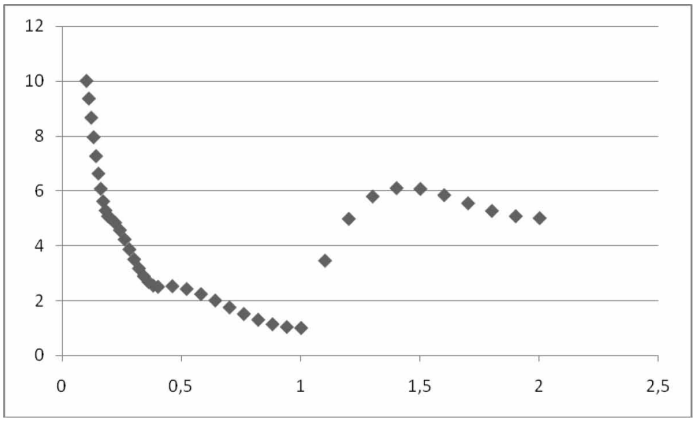
Example 9: (refer to Figure 9)

PNC for $\gamma = \alpha^2$ and $h(p_1,p_2) = x_1,y_1 + x_2y_2$.

Curve interpolation represents one of the most important problems in mathematics and computer science: how to model the curve via discrete set of two-dimensional points? Also the matter of shape representation (as closed curve-contour) and curve parameterization is still opened. For example pattern recognition, signature verification or handwriting identification problems are based on curve modeling via the choice of key points. So interpolation is not only a pure mathematical problem but important task in computer vision and artificial intelligence. The monograph wants to approach a problem of curve modeling by characteristic points. Proposed method relies on nodes combination and functional modeling of curve points situated between the basic set of key points. The functions that are used in calculations represent whole family of elementary functions with inverse functions: polynomials, trigonometric, cyclometric, logarithmic, exponential and power function. These functions are treated as probability distribution functions in the range [0;1].

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Quadratic symbol modeling with nine reconstructed points between nodes



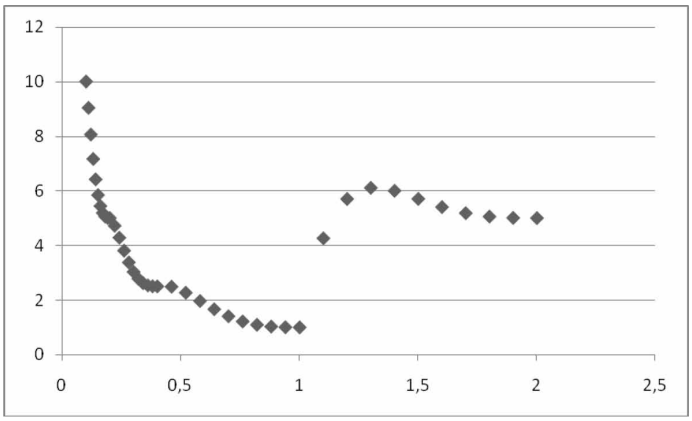
Example 10: (refer to Figure 10)
PNC for $\gamma = \alpha^3$ and $h(p_1,p_2) = x_1,y_1 + x_2y_2$.

If we consider Figure 1 as closed curve (contour) with the node $p_6 = p_1 = (0.1;10)$ then examples 8 and 9 give the shapes:

Example 11: (refer to Figure 11)
PNC for $\gamma = \alpha^2$ and $h(p_1,p_2) = x_1,y_1 + x_2y_2$.
Example 12: (refer to Figure 12)
PNC for $\gamma = \alpha^3$ and $h(p_1,p_2) = x_1,y_1 + x_2y_2$.

Every man has individual style of handwriting. Recognition of handwritten letter or symbol need modeling and the model of each individual symbol or character can be built by choice of γ and h. PNC modeling via nodes combinations h and parameter γ as probability distribution function enables curve interpolation for each specific letter or symbol.

Figure 10. Cubic character modeling with nine reconstructed points between nodes



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Quadratic contour modeling with nine reconstructed points between nodes

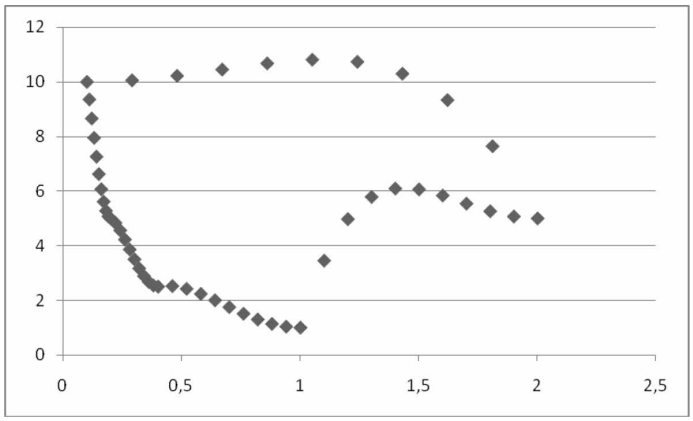
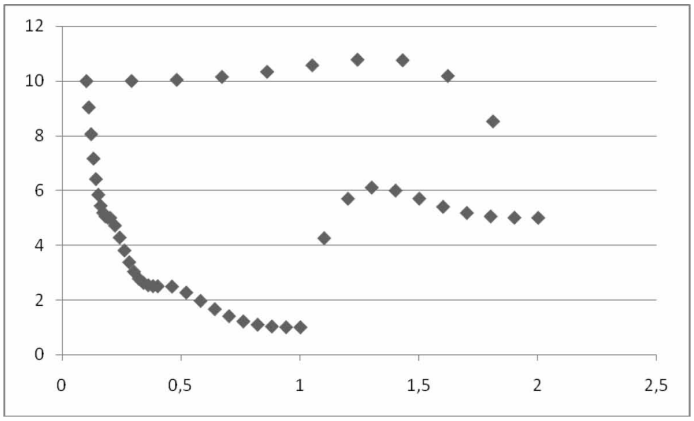


Figure 12. Cubic shape modeling with nine reconstructed points between nodes



Number of reconstructed points depends on a user by value α. If for example $\alpha = 0.01, 0.02,\ldots,0.99$ then 99 points are interpolated for each pair of nodes. Reconstructed values and interpolated points, calculated by PNC method, are applied in the process of curve modeling. Every curve can be interpolated by some distribution function as parameter γ and nodes combination h. Parameter γ is treated as probability distribution function for each curve.

Beta Distribution

Basic distribution ($\gamma = \alpha$) with nodes combination $h = 0$ turns PNC interpolation (1) to linear interpolation. What about PNC in the case of yet another distribution on the range [0;1]: beta distribution? Power functions as γ used in examples 1, 7 and 9-12 are also connected with beta distribution. Here are the examples of PNC modeling for beta distribution with nodes combination $h = 0$.

Example 13: (refer to Figure 13)

PNC for $\gamma = 3\alpha^2-2\alpha^3$ and $h(p_1,p_2)=0$.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Example 14: (refer to Figure 14)
PNC for $\gamma = 4\alpha^3\text{-}3\alpha^4$ and $h(p_1,p_2)=0$.
Example 15: (refer to Figure 15)
PNC for $\gamma = 2\alpha\text{-}\alpha^2$ and $h(p_1,p_2)=0$.

Examples 9-12 represent beta distribution with $h(p_1,p_2) = x_1,y_1 + x_2y_2$.

Exponential Distribution

Exponential distribution is dealing with random variable ≥ 0, but in PNC interpolation random variable $\alpha\in[0;1]$. Then exponential distribution is represented by distribution function:

$\gamma = F(\alpha) = 1 - (1\text{-}\alpha)e^{\alpha}$.

Example 16: (refer to Figure 16)
PNC for $\gamma = 1\text{-}(1\text{-}\alpha)e^{\alpha}$ and $h(p_1,p_2)=0$.
Example 17: (refer to Figure 17)
PNC for $\gamma = 1\text{-}(1\text{-}\alpha)e^{\alpha}$ and $h(p_1, p_2) = \frac{y_2}{y_1} + \frac{x_2}{x_1}$:

Figure 13. Beta distribution in handwritten character modeling

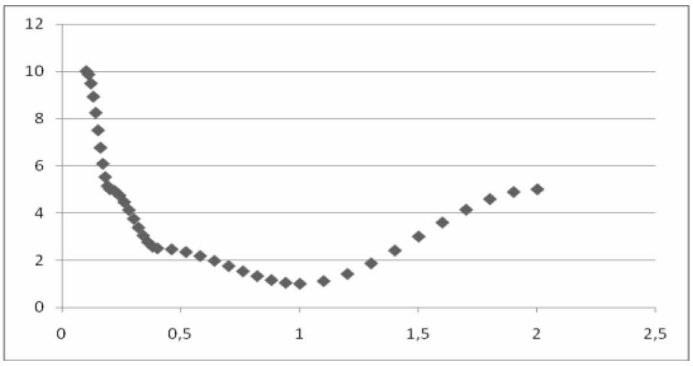
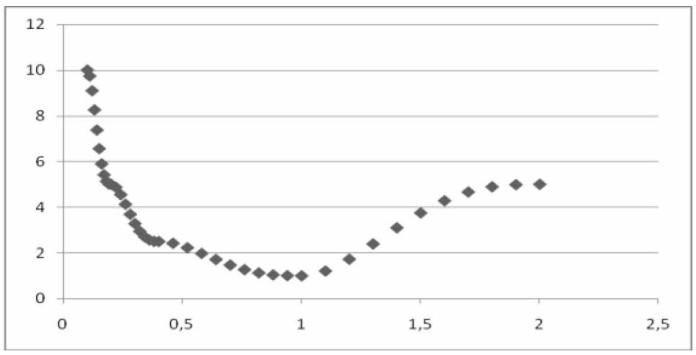


Figure 14. Beta distribution in handwritten symbol modeling



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 15. Beta distribution in handwritten letter modeling

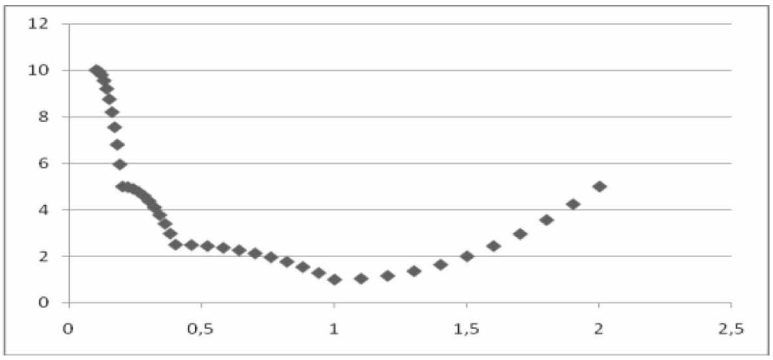


Figure 16. Exponential distribution in handwritten character modeling

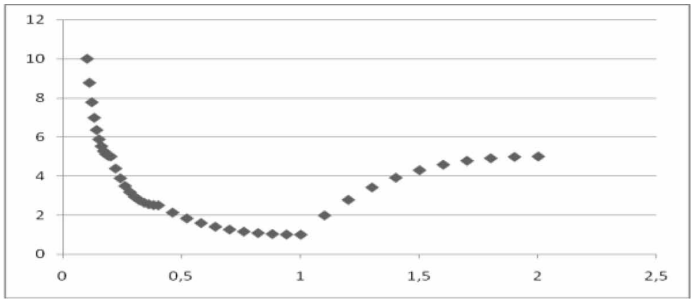
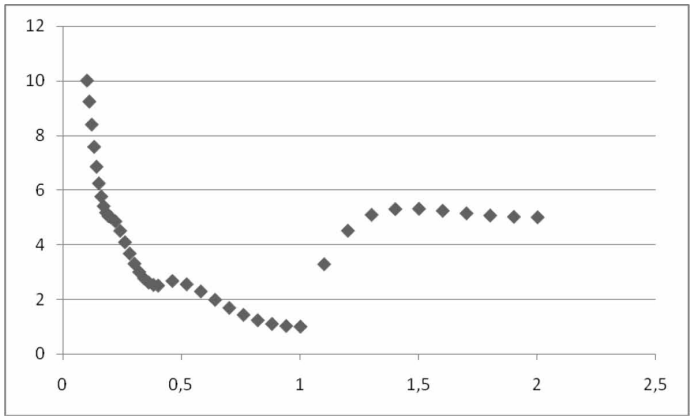


Figure 17. Exponential distribution in handwritten symbol modeling



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

These examples show the variety of possibilities in curve modeling via the choice of nodes combination and probability distribution function for interpolated points.

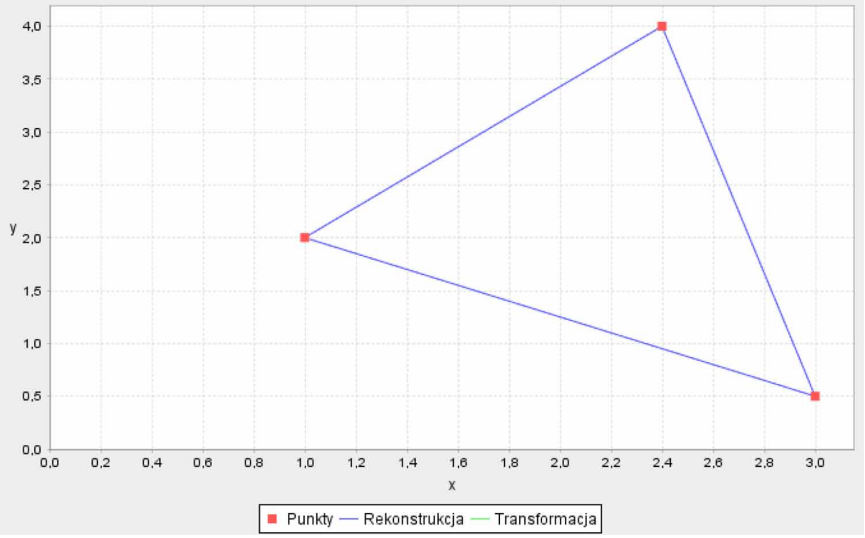
2D OBJECT MODELING

What about contours – closed curves? The only assumption is that first node and last node are the same. The simplest way is the linear interpolation for h=0 and $\gamma=\alpha$. (Figure 18).

Proposed method, called Probabilistic Nodes Combination (PNC), is the method of 2D curve interpolation and extrapolation using the set of key points (knots or nodes). Nodes can be treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PNC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. Two-dimensional curve is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function. Novelty of this book consists of two generalizations: generalization of previous MHR method with various nodes combinations and generalization of linear interpolation with different (no basic) probability distribution functions and nodes combinations. Here are examples of shape modeling for five (or rather six) nodes:

Significant problem in machine vision and computer vision is that of appropriate 2D shape representation and reconstruction. Classical discussion about shape representation is based on the problem: contour versus skeleton. This monograph is voting for contour which forms boundary of the object. Contour of the object, represented by successive contour points, consists of information which allows us to describe many important features of the object as shape coefficients. 2D curve modeling and generation is a basic subject in many branches of industry and computer science, for example in the cad/cam software. The representation of shape can have a great impact on the accuracy and effectiveness of object recognition.

Figure 18. Contour reconstructed 1for h=0 and γ=α



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Shape modeling for nodes

Lp.	*x*	*y*
1	1	2.5
2	1.7	4
3	2.6	3
4	2.3	0.5
5	1.5	1

Figure 19. PFC 2D modeling for s=1, h=0, $\gamma=\alpha^s$ (linear modeling)

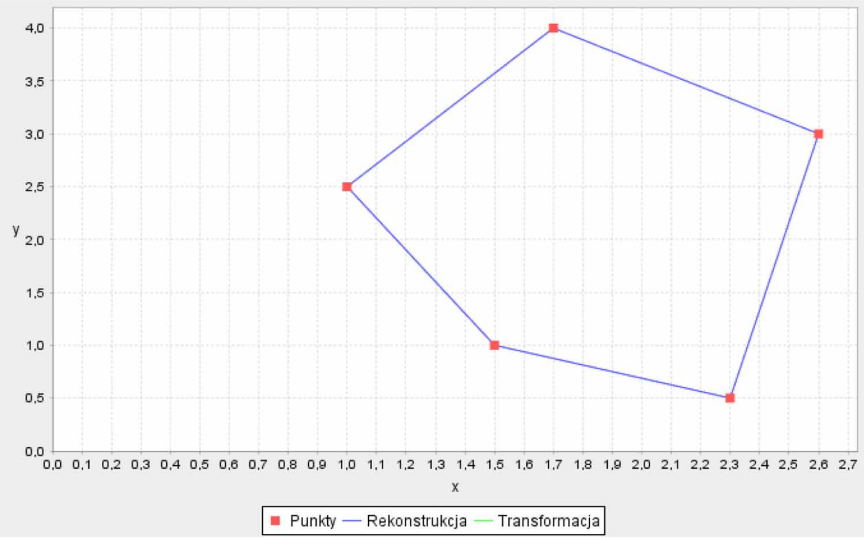
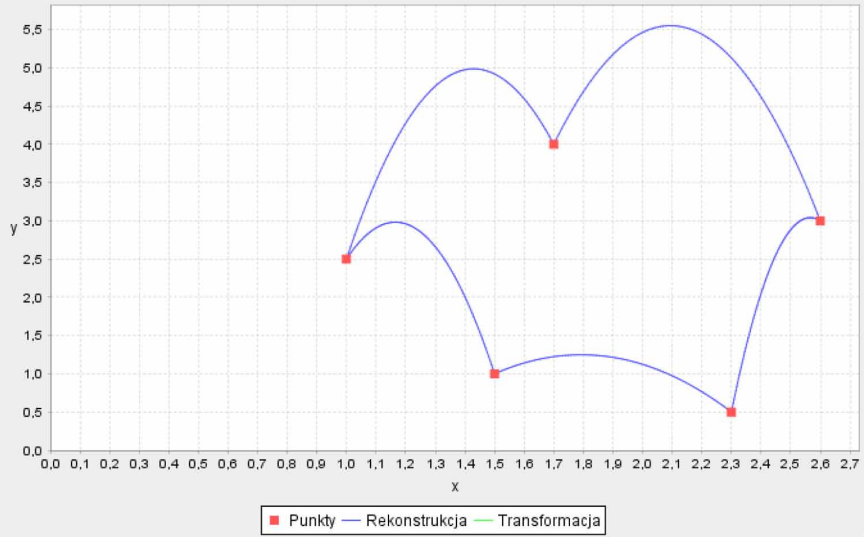


Figure 20. PFC 2D modeling for s=1, h= $x_1x_2 + y_1y_2$, $\gamma=\alpha^s$



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 21. PFC 2D modeling for s=0.7, h=0, γ=α^s

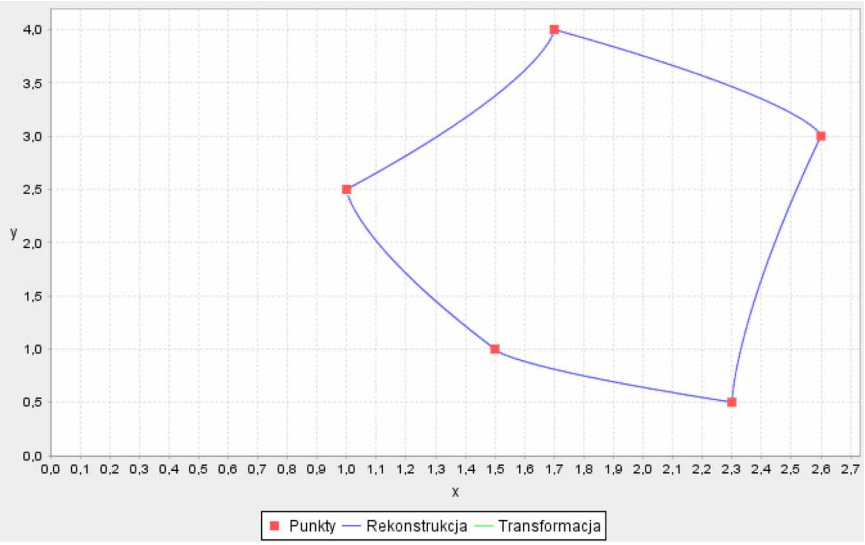
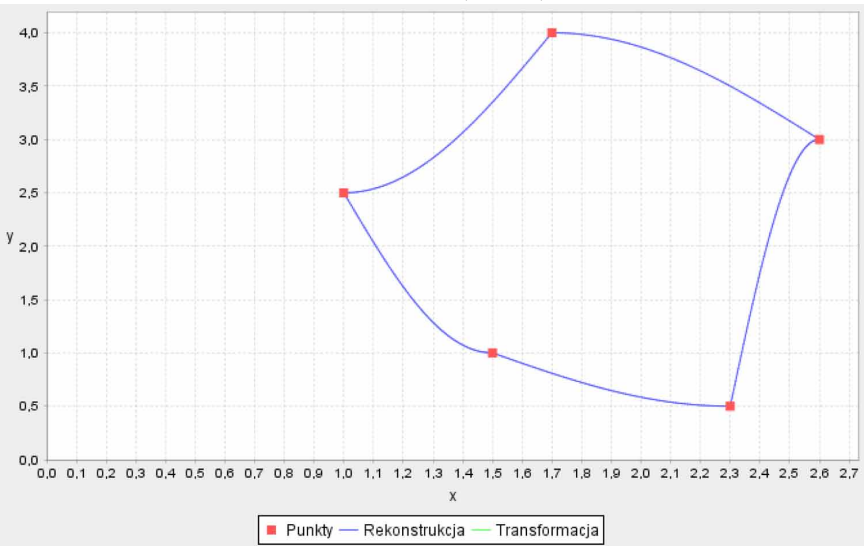


Figure 22. PFC 2D modeling for s=1, h=0, $\gamma = \sin\left(\alpha^s \cdot \frac{\pi}{2}\right)$



The method of Probabilistic Nodes Combination (PNC) enables interpolation and modeling of two-dimensional curves using nodes combinations and different coefficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function, also inverse functions. This probabilistic view is novel approach a problem of modeling and interpolation. Computer vision and pattern recognition are interested in appropriate methods of shape representation and curve modeling. PNC method represents the possibilities of shape reconstruction and curve interpolation via the choice of nodes combination and probability distribution function for

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 23. PFC 2D modeling for s=1, h= $x_1x_2 + y_1y_2$, $\gamma = 1 - \cos\left(\alpha^s \cdot \frac{\pi}{2}\right)$

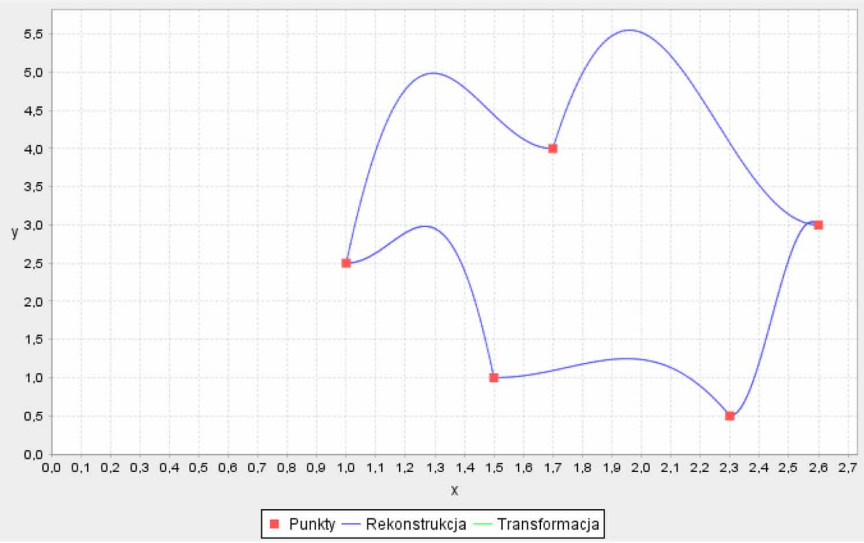
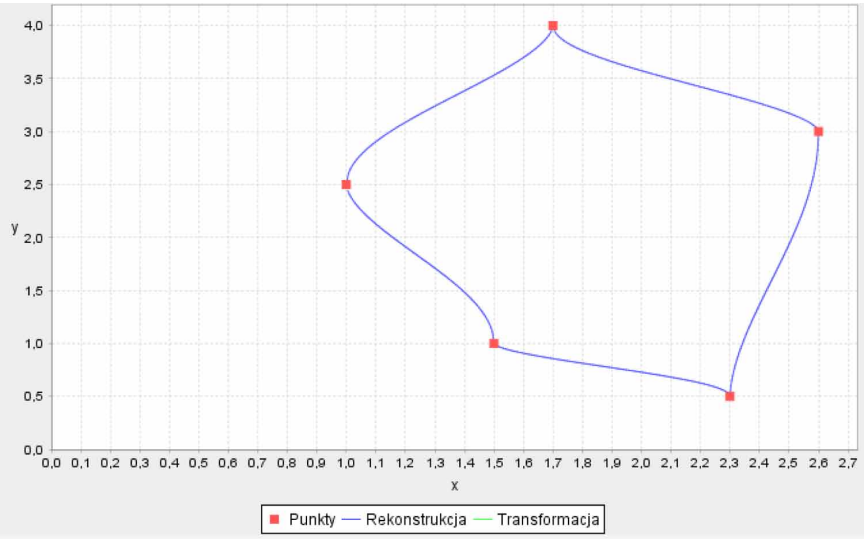


Figure 24. PFC 2D modeling for s=0.6, h=0, $\gamma = \frac{2}{\pi} \cdot \arcsin\left(\alpha^s\right)$



interpolated points. It seems to be quite new look at the problem of contour representation and curve modeling in artificial intelligence and computer vision.

Function for γ calculations is chosen individually at each curve modeling and it is treated as probability distribution function: γ depends on initial requirements and curve specifications. PNC method leads to curve interpolation as handwriting modeling via discrete set of fixed knots. So PNC makes possible the combination of two important problems: interpolation and modeling.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 25. Contour interpolation for: 2s=0.6, h= $x_1x_2 + y_1y_2$, $\gamma = \frac{2}{\pi} \cdot \arcsin\left(\alpha^s\right)$

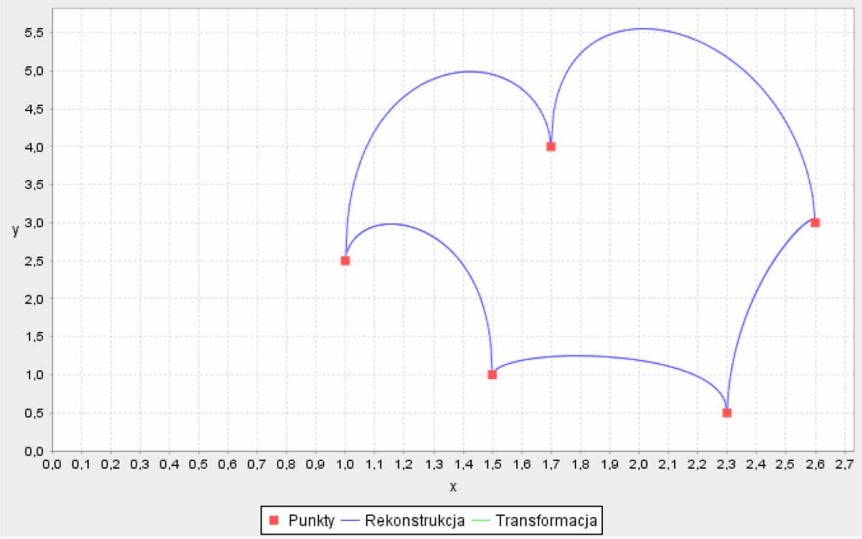
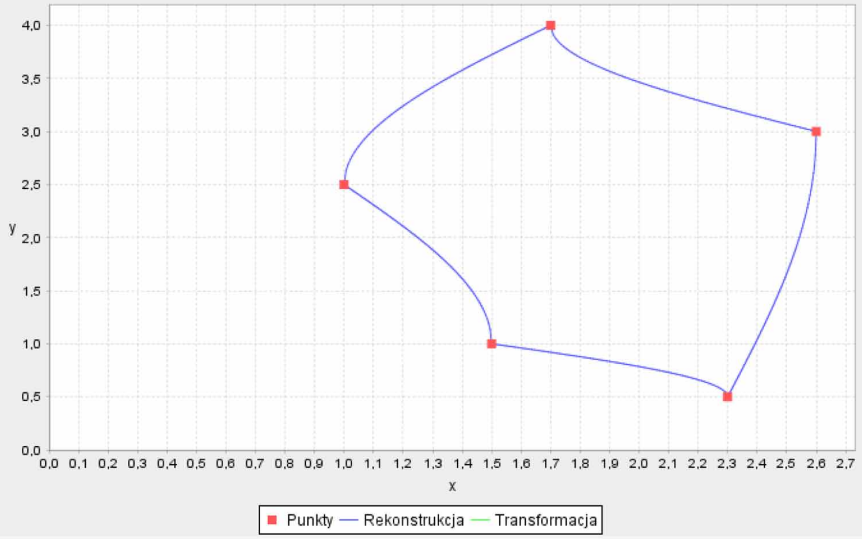


Figure 26. Shape reconstruction for: s=1, h=0, $\gamma = 1 - \frac{2}{\pi} \cdot \arccos\left(\alpha^s\right)$



For the same set of nodes above contours are reconstructed via modeling functions γ and parameters *s*, *h*. Minimal set of characteristic points in contour reconstruction consists of three nodes. Here are PNC shape modeling for nodes: (1;1), (10;10), (20;5).

Significant problem in machine vision and computer vision is that of appropriate 2D shape representation and reconstruction. Classical discussion about shape representation is based on the problem: contour versus skeleton. This monograph is voting for contour which forms boundary of the object. Contour of the object, represented by successive contour points, consists of information which allows us to describe

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 27. Contour modeling for: s=1, h= $x_1x_2 + y_1y_2$, $\gamma = 1 - \frac{2}{\pi} \cdot \arccos\left(\alpha^s\right)$

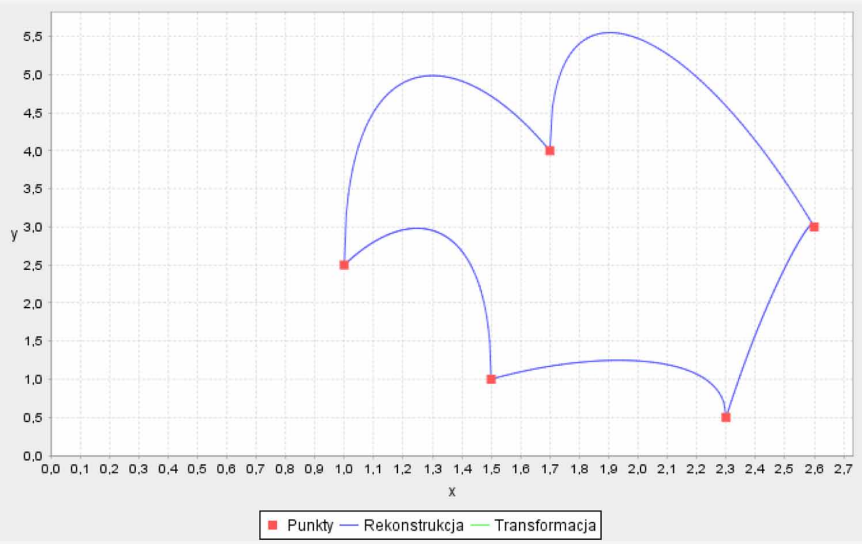
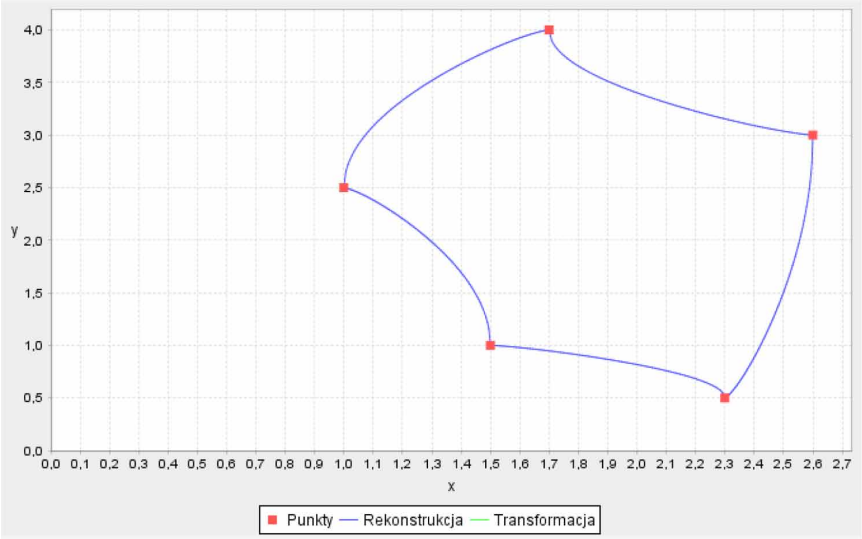


Figure 28. Object reconstruction for: s=1.3, h=0, $\gamma = 1 - \frac{2}{\pi} \cdot \arccos\left(\alpha^s\right)$



many important features of the object as shape coefficients. 2D curve modeling and generation is a basic subject in many branches of industry and computer science, for example in the cad/cam software. The representation of shape can have a great impact on the accuracy and effectiveness of object recognition.

Nodes are treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PFC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. Two-dimensional object is extrapolated and interpolated via nodes combination and

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 29. Shape interpolation for: s=1.3, h= $x_1x_2 + y_1y_2$, $\gamma = 1 - \frac{2}{\pi} \cdot \arccos\left(\alpha^s\right)$

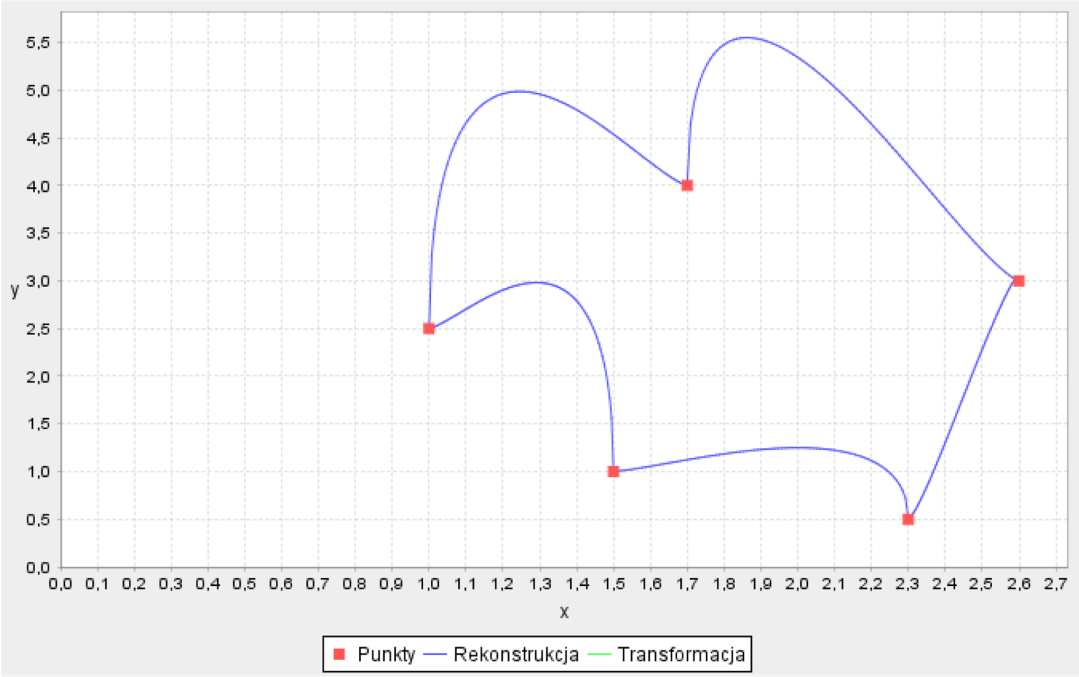
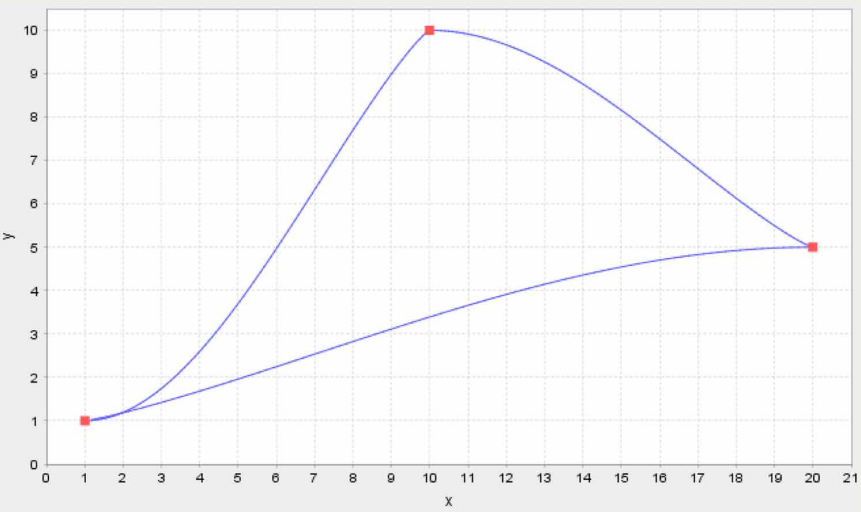


Figure 30. Contour reconstruction for s = 1.2, h = 0, $\gamma = \sin\left(\alpha^s * \frac{\pi}{2}\right)$



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 31. Shape interpolation for s = 1.2, h = 0, $\gamma = \log_2\left(\alpha^s + 1\right)$

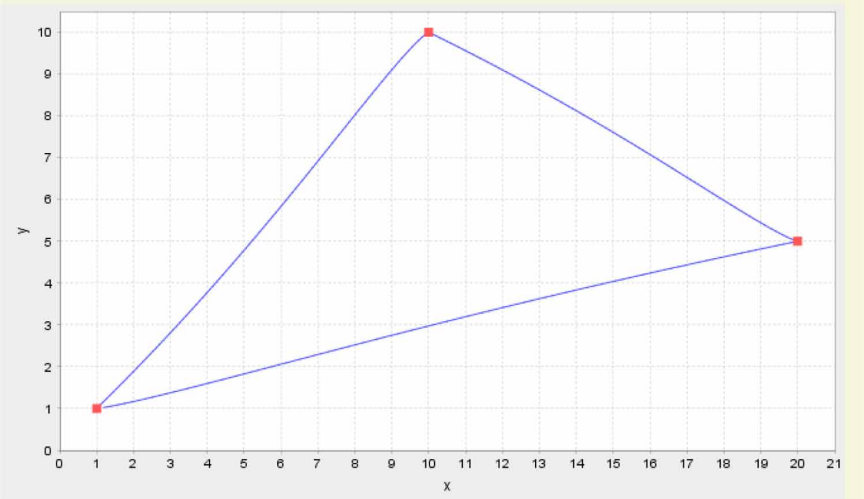
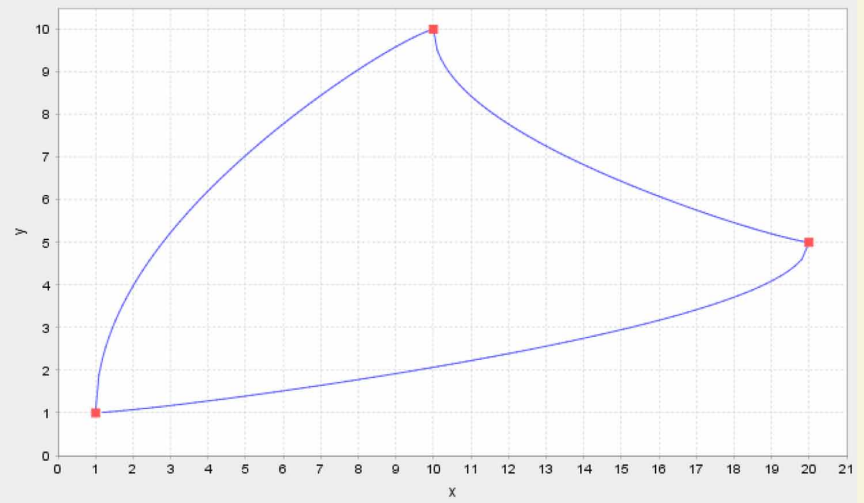


Figure 32. Object modeling for s = 1.2, h = 0, $\gamma = \frac{2}{\pi} * \arcsin\left(\alpha^s\right)$



different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function. Functions for γ calculations are chosen individually at each data modeling and it is treated as 2-dimensional probability distribution function: γ depends on initial requirements and features' specifications. PFC method leads to data interpolation as handwriting or signature identification and image retrieval via discrete set of feature vectors in 2-dimensional feature space. So PFC method makes possible the combination of two important problems: interpolation and modeling in a matter of image retrieval or writer identification. PFC interpolation develops a linear interpolation in multidimensional feature spaces into other functions as two-dimensional probability distribution functions.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 33. Shape reconstruction for s = 1.4, h = $y_1\left(\frac{x_2}{x_1}\right)+y_2\left(\frac{x_1}{x_2}\right),\ \gamma=1-\cos\left(\alpha^s*\frac{\pi}{2}\right)$

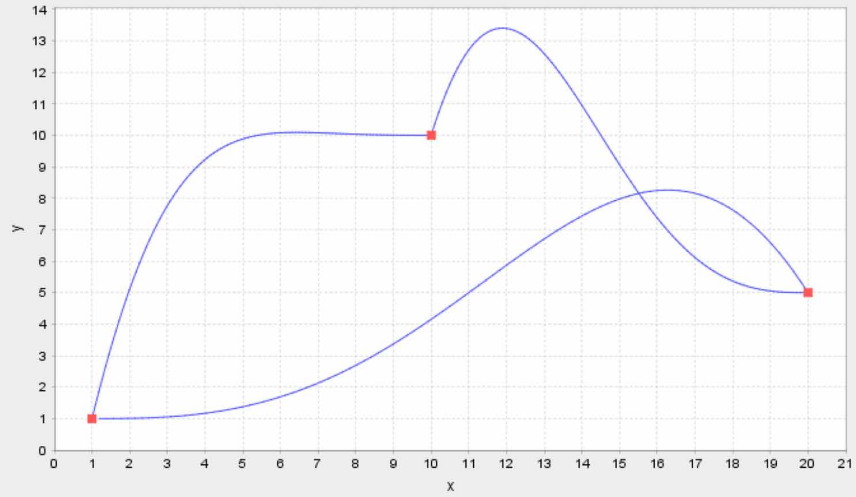
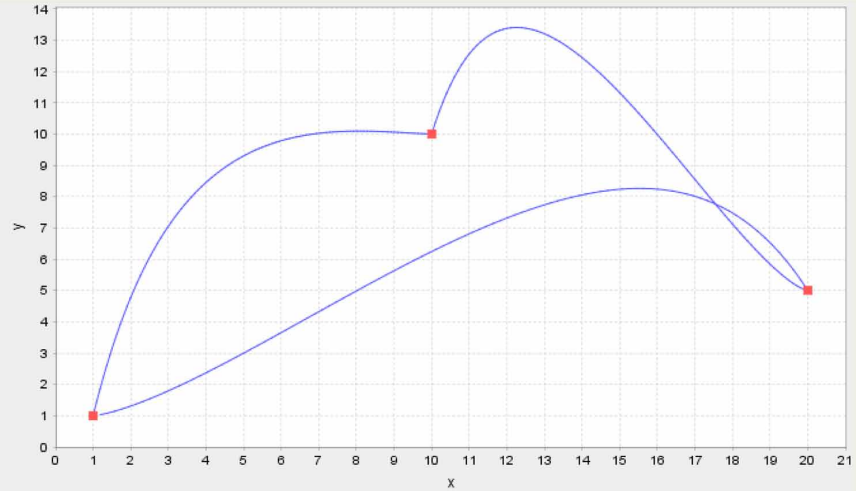


Figure 34. Contour interpolation for s = 1.4, h = $y_1\left(\frac{x_2}{x_1}\right)+y_2\left(\frac{x_1}{x_2}\right),\ \gamma=ctg\left(\frac{\pi}{2}-\alpha^s*\frac{\pi}{4}\right)$



These examples of shape modeling and contour reconstruction show the variety of PNC interpolations via chosen parameters and set of nodes.

RESULT ANALYSIS

PFC method is interpolating a curve between each pair of nodes using modeling function $\gamma = F(\alpha)$ and nodes combination *h*. The simplest way of comparing PFC with another method is to see the example. Here is the application of PFC method for function $f(x)=1/x$ and nine nodes: $y = 0.2, 0.4, 0.6, 0.8, 1,$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 35. Object modeling with s = 1.4, $h = y_1\left(\frac{x_2}{x_1}\right) + y_2\left(\frac{x_1}{x_2}\right),\ \gamma = 1 - \frac{2}{\pi} * \arccos\left(\alpha^s\right)$

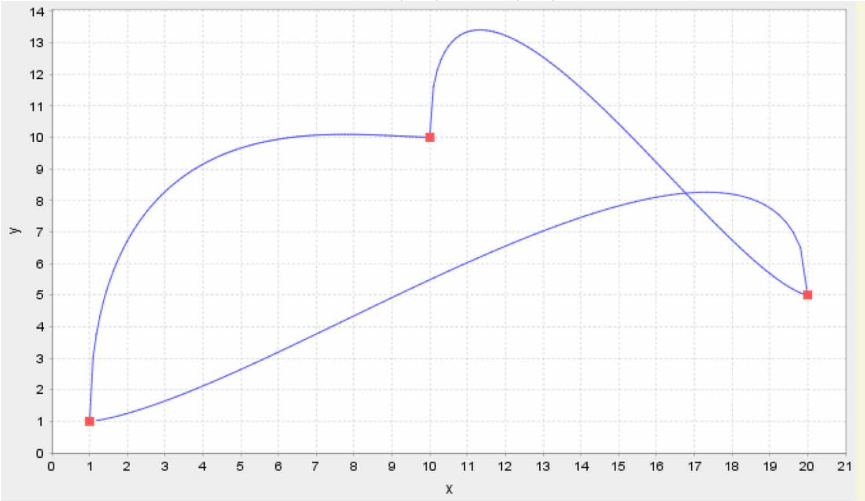
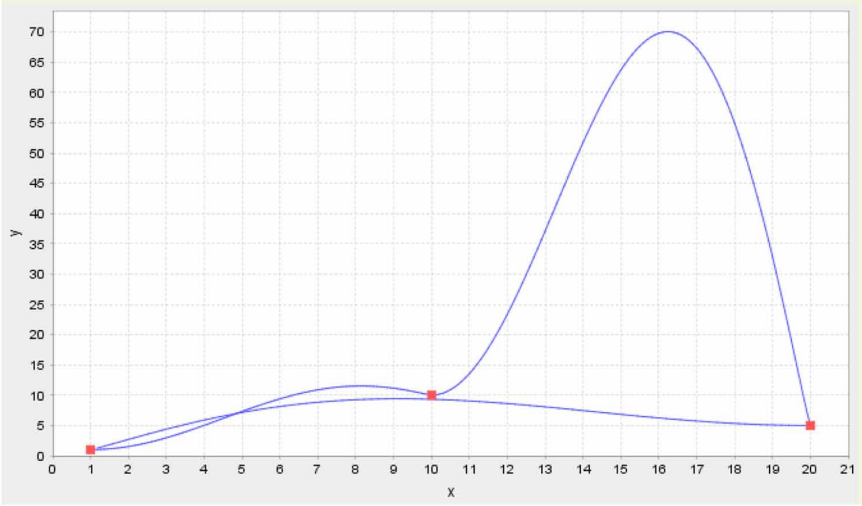


Figure 36. Shape interpolation with s = 1.1, h= $x_1x_2 + y_1y_2,\ \gamma = \sin\left(\alpha^s * \frac{\pi}{2}\right)$



1.2, 1.4, 1.6, 1.8. PFC represents (Figure 42) much more precise interpolation than Lagrange or Newton polynomial interpolation (Figure 43).

PFC interpolation is more accurate than polynomial interpolation by Newton or Lagrange. Very important matter is dealing with closed curves. PFC reconstruction of the contour or shape is done with the same formulas. Another important problem is connected with extrapolation. PFC method gives the tool of data anticipation or prediction.

PFC is a novel approach to the matter of data modeling, reconstruction and extrapolation.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 37. Object reconstruction with s = 1.1, h= $x_1x_2 + y_1y_2$, $\gamma = ctg\left(\frac{\pi}{2} - \alpha^s * \frac{\pi}{4}\right)$

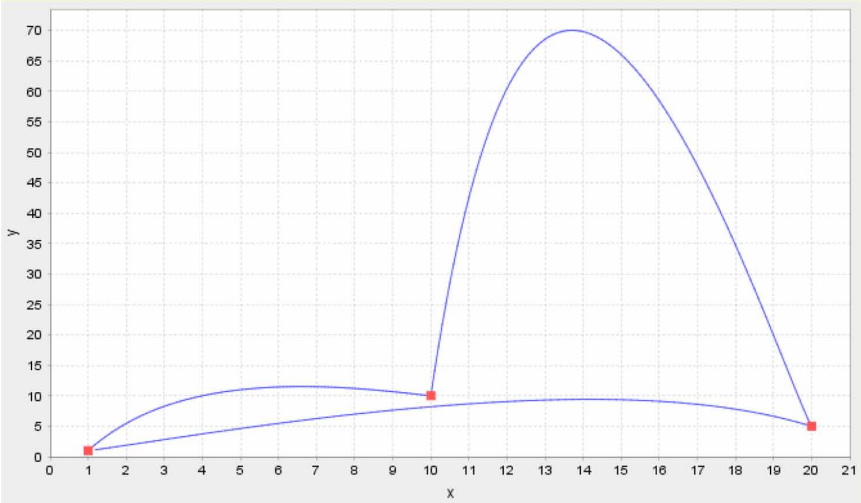
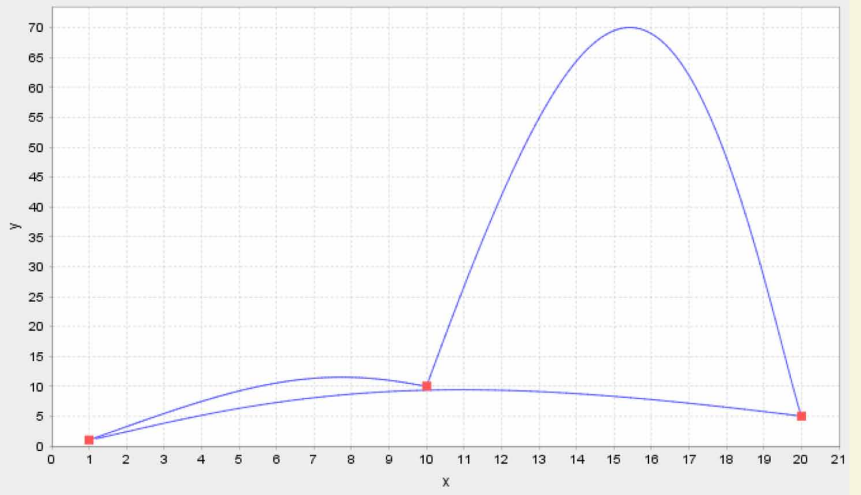


Figure 38. Contour modeling with s = 1.1, h= $x_1x_2 + y_1y_2$, $\gamma = \frac{4}{\pi} * \arctan\left(\alpha^s\right)$



SOLUTIONS AND RECOMMENDATIONS

Proposed method, called Probabilistic Features Combination (PFC), is the method of 2D curve interpolation and extrapolation using the set of key points (knots or nodes). Nodes can be treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PFC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. Two-dimensional curve is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 39. Shape modeling with s = 1.3, h= $x_1x_2 + y_1y_2$, $\gamma = 1 - \cos\left(\alpha^s * \frac{\pi}{2}\right)$

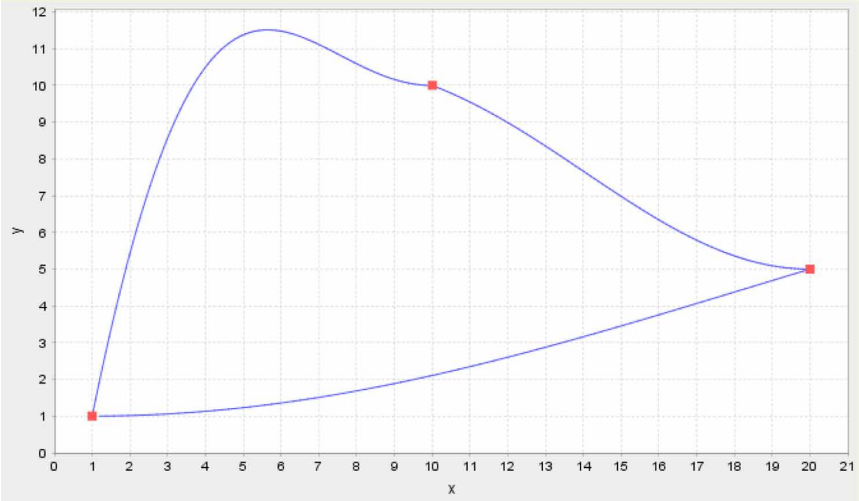
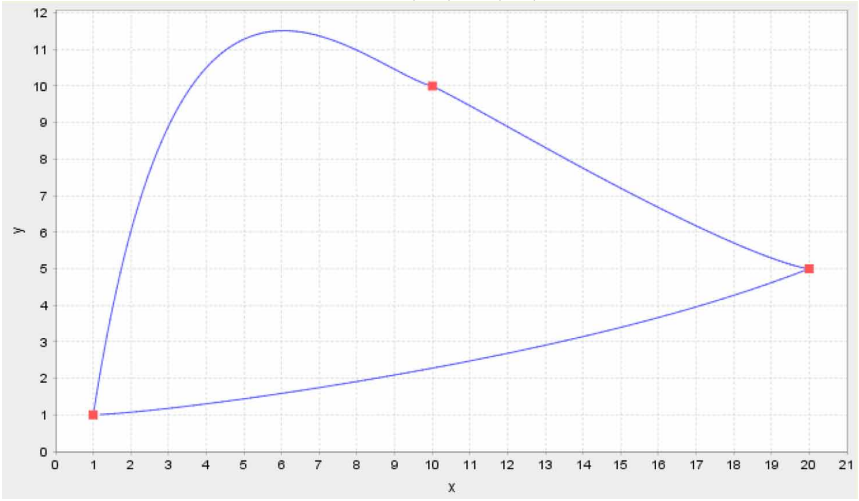


Figure 40. Object interpolation with s = 1.3, h = $\left(\frac{x_2}{x_1}\right) + \left(\frac{y_2}{y_1}\right)$, $\gamma = ctg\left(\frac{\pi}{2} - \alpha^s * \frac{\pi}{4}\right)$



FUTURE RESEARCH DIRECTIONS

Future trends will go to various directions: how to fix the best probability distribution function for the nodes, how to calculate the most appropriate nodes combination and what extrapolation is the most valuable in decision making and risk analysis.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 41. Contour reconstruction with s = 1.3, h = $\left(\frac{x_2}{x_1}\right) + \left(\frac{y_2}{y_1}\right)$, $\gamma = \frac{2}{\pi} * \arcsin\left(\alpha^s\right)$

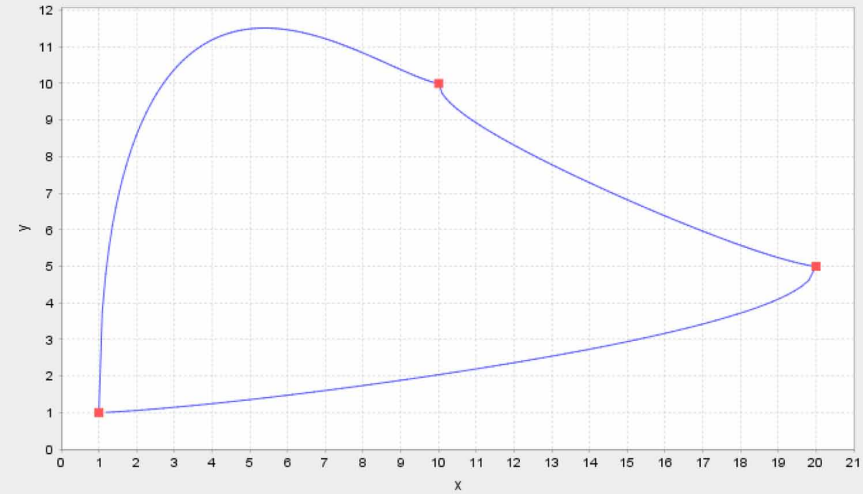
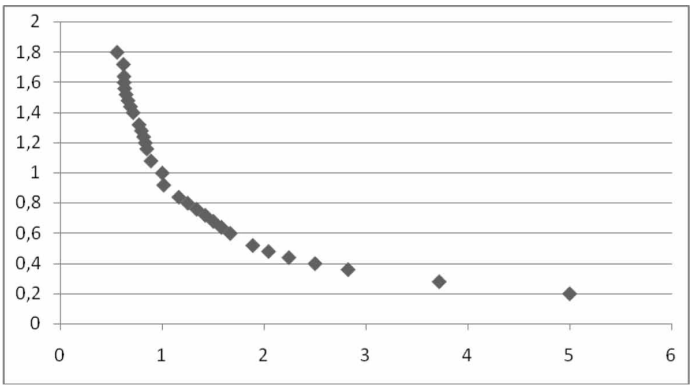


Figure 42. Points of function f(x) = 1/x using PFC method with 9 nodes – better than polynomial interpolation



CONCLUSION

The method of Probabilistic Nodes Combination (PNC) enables interpolation and modeling of two-dimensional curves using nodes combinations and different coefficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function, also inverse functions. This probabilistic view is novel approach a problem of modeling and interpolation. Computer vision and pattern recognition are interested in appropriate methods of shape representation and curve modeling. PNC method represents the possibilities of shape reconstruction and curve interpolation via the choice of nodes combination and probability distribution function for interpolated points. It seems to be quite new look at the problem of contour representation and curve modeling in artificial intelligence and computer vision. Function for γ calculations is chosen individually at each curve modeling and it is treated as probability distribution function: γ depends on initial requirements

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 43. Interpolation polynomial of function f(x) = 1/x is completely wrong

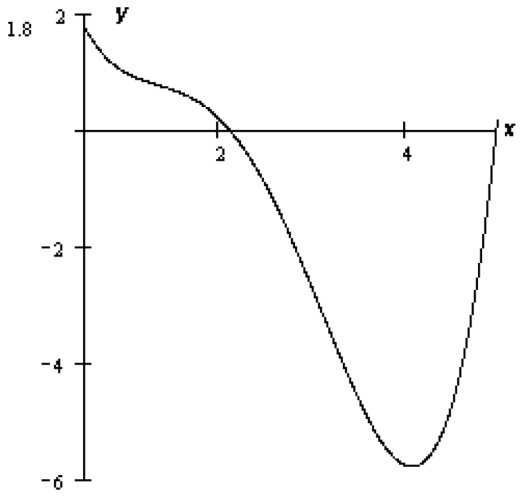
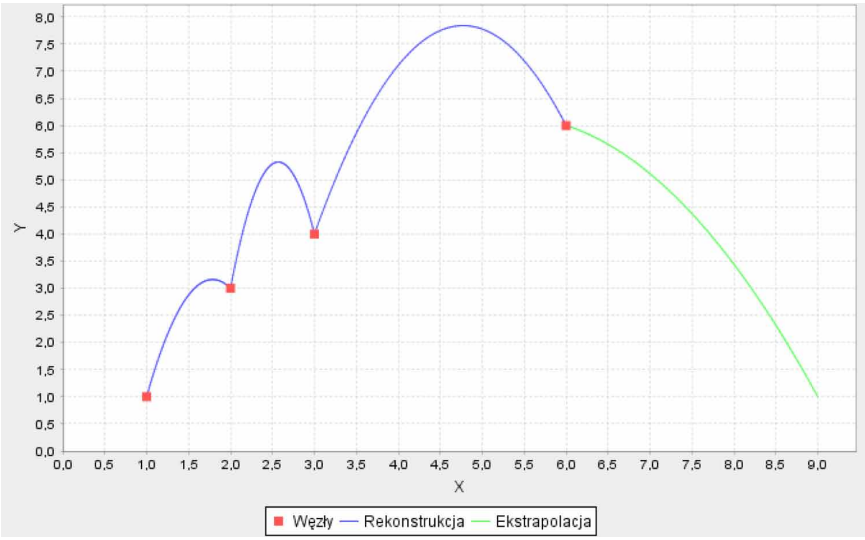


Figure 44. Extrapolation of data right of the last node

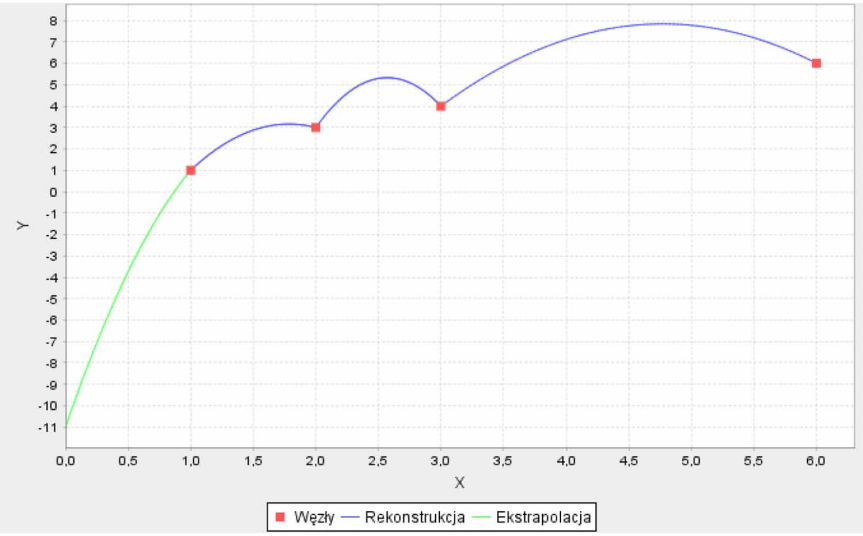


and curve specifications. PNC method leads to curve interpolation as handwriting modeling via discrete set of fixed knots. So PNC makes possible the combination of two important problems: interpolation and modeling. Main features of PNC method are:

1. The smaller distance between knots the better;
2. Calculations for coordinates close to zero and near by extremum require more attention because of importance of these points;
3. PNC interpolation develops a linear interpolation into other functions as probability distribution functions;
4. PNC is a generalization of MHR method via different nodes combinations;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 45. Prediction of values left of first node



5. Interpolation of L points is connected with the computational cost of rank $O(L)$ as in MHR method;
6. Nodes combination and coefficient γ are crucial in the process of curve probabilistic parameterization and interpolation: they are computed individually for a single curve.

What is the most important feature of MHR and PNC methods? Interpolation methods and curve fitting represent so huge problem that each individual interpolation is exceptional and requires specific solutions. PNC method is such a novel tool with its all pros and cons. The user has to decide which interpolation method is the best in a single situation. The choice is yours if you have any choice. Presented method is such a new possibility for curve fitting and interpolation when specific data (for example handwritten symbol or character) starts up with no rules for polynomial interpolation. This book consists of two generalizations: generalization of previous MHR method with various nodes combinations and generalization of linear interpolation with different (no basic) probability distribution functions and nodes combinations. The method of Probabilistic Nodes Combination (PNC) enables interpolation and modeling of two-dimensional curves using nodes combinations and different coefficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function, also inverse functions. This probabilistic view is novel approach a problem of modeling and interpolation. Computer vision and pattern recognition are interested in appropriate methods of shape representation and curve modeling. PNC method represents the possibilities of shape reconstruction and curve interpolation via the choice of nodes combination and probability distribution function for interpolated points. It seems to be quite new look at the problem of contour representation and curve modeling in artificial intelligence and computer vision. Function for γ calculations is chosen individually at each curve modeling and it is treated as probability distribution function: γ depends on initial requirements and curve specifications. PNC method leads to curve interpolation as handwriting modeling via discrete set of fixed knots. So PNC makes possible the combination of two important problems: interpolation and modeling. The method of Probabilistic Features Combination (PFC) enables interpolation and modeling of high-dimensional N data using features' combinations and different coef-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function. Functions for γ calculations are chosen individually at each data modeling and it is treated as *N*-dimensional probability distribution function: γ depends on initial requirements and features' specifications. PFC method leads to data interpolation as handwriting or signature identification and image retrieval via discrete set of feature vectors in N-dimensional feature space. So PFC method makes possible the combination of two important problems: interpolation and modeling in a matter of image retrieval or writer identification. Main features of PFC method are: PFC interpolation develops a linear interpolation in multidimensional feature spaces into other functions as *N*-dimensional probability distribution functions; PFC is a generalization of MHR method and PNC method via different nodes combinations; interpolation of *L* points is connected with the computational cost of rank $O(L)$ as in MHR and PNC method; nodes combination and coefficients γ are crucial in the process of data probabilistic parameterization and interpolation: they are computed individually for a single feature. Future works are going to applications of PFC method in signature and handwriting biometric recognition: choice and features of nodes combinations *h* and coefficients γ.

Proposed method, called Probabilistic Features Combination (PFC), is the method of *N*-dimensional data interpolation and extrapolation using the set of key points (knots or nodes). Nodes can be treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PFC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. *N*-dimensional object is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function. The method of Probabilistic Features Combination (PFC) enables interpolation and modeling of high-dimensional data using features' combinations and different coefficients γ as modeling function. Functions for γ calculations are chosen individually at each data modeling and it is treated as *N*-dimensional probability distribution function: γ depends on initial requirements and features' specifications. PFC method leads to data interpolation as handwriting or signature identification and image retrieval via discrete set of feature vectors in N-dimensional feature space. So PFC method makes possible the combination of two important problems: interpolation and modeling in a matter of image retrieval or writer identification. PFC interpolation develops a linear interpolation in multidimensional feature spaces into other functions as *N*-dimensional probability distribution functions. Future works are going to applications of PFC method in biometric recognition, computer vision and artificial intelligence. Nodes are treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PFC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. Two-dimensional object is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function. Functions for γ calculations are chosen individually at each data modeling and it is treated as 2-dimensional probability distribution function: γ depends on initial requirements and features' specifications. PFC method leads to data interpolation as handwriting or signature identification and image retrieval via discrete set of feature vectors in 2-dimensional feature space. So PFC method makes possible the combination of two important problems: interpolation and modeling in a matter of image retrieval or writer identification.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

PFC interpolation develops a linear interpolation in multidimensional feature spaces into other functions as two-dimensional probability distribution functions.

Future works are going to: application of PNC method in signature and handwriting recognition, choice and features of nodes combinations and coefficient γ, implementation of PNC in computer vision and artificial intelligence: shape geometry, contour modelling, object recognition and curve parameterization.

REFERENCES

Ballard, D. H. (1982). *Computer Vision*. New York: Prentice Hall.

Chapra, S. C. (2012). *Applied Numerical Methods*. McGraw-Hill.

Chorań, R. S. (2005). *Computer Vision*. Warsaw, Poland: Exit.

Cocozza-Thivent, C., Eymard, R., Mercier, S., & Roussignol, M. (2006). Characterization of the Marginal Distributions of Markov Processes Used in Dynamic Reliability. *Journal of Applied Mathematics and Stochastic Analysis*, 1–18.

Collins, G. W. II. (2003). *Fundamental Numerical Methods and Data Analysis*. Case Western Reserve University.

Dahlquist, G., & Bjoerck, A. (1974). *Numerical Methods*. New York: Prentice Hall.

Dejdumrong, N. (2007). A Shape Preserving Verification Techniques for Parametric Curves. *Computer Graphics, Imaging and Visualization*, 163–168.

Dyn, N., Levin, D., & Gregory, J. A. (1987). A 4-Point Interpolatory Subdivision Scheme for Curve Design. *Computer Aided Geometric Design*, *4*(4), 257–268. doi:10.1016/0167-8396(87)90001-X

Jakóbczak, D. (2007). 2D and 3D Image Modeling Using Hurwitz-Radon Matrices. *Polish Journal of Environmental Studies*, *4A*(16), 104–107.

Jakóbczak, D. (2009). Curve Interpolation Using Hurwitz-Radon Matrices. *Polish Journal of Environmental Studies*, *3B*(18), 126–130.

Jakóbczak, D. (2010a). Shape Representation and Shape Coefficients via Method of Hurwitz-Radon Matrices. *Lecture Notes in Computer Science*, *6374*, 411–419. doi:10.1007/978-3-642-15910-7_47

Jakóbczak, D. (2010b). Object Modeling Using Method of Hurwitz-Radon Matrices of Rank k. In W. Wolski & M. Borawski (Eds.), *Computer Graphics: Selected Issues* (pp. 79–90). Szczecin, Poland: University of Szczecin Press.

Jakóbczak, D. (2011a). Curve Parameterization and Curvature via Method of Hurwitz-Radon Matrices. *Image Processing & Communications- International Journal (Toronto, Ont.)*, *1-2*(16), 49–56.

Jakóbczak, D. (2011b). Data Extrapolation and Decision Making via Method of Hurwitz-Radon Matrices. *Lecture Notes in Computer Science*, *6922*, 173–182. doi:10.1007/978-3-642-23935-9_17

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Jakóbczak, D. (2011c). Curve Extrapolation and Data Analysis using the Method of Hurwitz-Radon Matrices. *Folia Oeconomica Stetinensia, 9*(17), 121-138.

Jakóbczak, D. (2013). Probabilistic Modeling of Signature using the Method of Hurwitz-Radon Matrices. *Global Perspectives on Artificial Intelligence*, *1*(1), 1–7.

Kozera, R. (2004). *Curve Modeling via Interpolation Based on Multidimensional Reduced Data*. Gliwice, Poland: Silesian University of Technology Press.

Liu, T., & Geiger, D. (1999). Approximate tree matching and shape similarity. *Int. Conf. Computer Vision,* Corfu, Greece.

Lorton, A., Fouladirad, M., & Grall, A. (2013). A Methodology for Probabilistic Model-based Prognosis. *European Journal of Operational Research*, *225*(3), 443–454. doi:10.1016/j.ejor.2012.10.025

Pergler, M., & Freeman, A. (2008). Probabilistic Modeling as an Exploratory Decision-Making Tool. *McKinsey Working Papers on Risk, 6*, 1-18.

Ralston, A., & Rabinowitz, P. (2001). *A First Course in Numerical Analysis* (2nd ed.). New York: Dover Publications.

Rogers, D. F. (2001). *An Introduction to NURBS with Historical Perspective*. Morgan Kaufmann Publishers.

Saber, E., Xu, Y., & Murat Tekalp, A. (2005). Partial shape recognition by sub-matrix matching for partial matching guided image labeling. *Pattern Recognition*, *38*(10), 1560–1573. doi:10.1016/j.patcog.2005.03.027

Schumaker, L. L. (2007). *Spline Functions: Basic Theory*. Cambridge Mathematical Library. doi:10.1017/CBO9780511618994

Sebastian, T. B., Klein, P. N., & Kimia, B. B. (2003). On aligning curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(1), 116–124. doi:10.1109/TPAMI.2003.1159951

Tadeusiewicz, R., & Flasiński, M. (1991). *Image Recognition*. Warsaw, Poland: PWN.

Zhang, D., & Lu, G. (2004). Review of Shape Representation and Description Techniques. *Pattern Recognition*, *1*(37), 1–19. doi:10.1016/j.patcog.2003.07.008

ADDITIONAL READING

Basu, S., & Bresler, Y. (2000). $O(N^2\log_2 N)$ filtered backprojection reconstruction algorithm for tomography. *IEEE Transactions on Image Processing*, *9*(10), 1760–1773. doi:10.1109/83.869187 PMID:18262914

Brankov, J. G., Yang, Y., & Wernick, M. N. (2004). Tomographic image reconstruction based on a Content – Adaptive Mesh Model. *IEEE Transactions on Medical Imaging*, *2*(23), 202–212. doi:10.1109/TMI.2003.822822 PMID:14964565

Brasse, D., & Defrise, M. (2004). Fast fully 3-D image reconstruction in PET using planograms. *IEEE Transactions on Medical Imaging*, *4*(23), 413–425. doi:10.1109/TMI.2004.824231 PMID:15084067

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bulacu, M., & Schomaker, L. (2007). Text-independent writer identification and verification using textural and allographic features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(4), 701–717. doi:10.1109/TPAMI.2007.1009 PMID:17299226

Bulacu, M., Schomaker, L., & Brink, A. (2007). *Text-independent writer identification and verification on off-line Arabic handwriting*. In: *International Conference on Document Analysis and Recognition*. 769–773.

Cetin, M., Karl, W. C., & Willsky, A. S. (2002, September). *Edge – preserving image reconstruction for coherent imaging application.* Paper presented at the IEEE International Conference on Image Processsing, Rochester, NY, USA. 10.1109/ICIP.2002.1039992

Chen, J., Cheng, W., & Lopresti, D. (2011). Using perturbed handwriting to support writer identification in the presence of severe data constraints. In: Document Recognition and Retrieval. 1–10. doi:10.1117/12.876497

Chen, J., Lopresti, D., & Kavallieratou, E. (2010). *The impact of ruling lines on writer identification*. In: *International Conference on Frontiers in Handwriting Recognition*. 439–444. 10.1109/ICFHR.2010.75

Chlebus, E., & Cholewa, M. (1999). Rapid prototyping – rapid tooling. *CADCAM Forum, 11*, 23-28.

Cormen, T. H., Leiserson, C. E., & Rivest, R. L. (1996). *Introduction to algorithms*. Massachusetts, USA: the Massachusetts Institute of Technology Press and McGraw-Hill.

Defrise, M. (2001). A short readers guide to 3D tomographic reconstruction. *Computerized Medical Imaging and Graphics*, *25*(2), 113–116. doi:10.1016/S0895-6111(00)00061-6 PMID:11137787

Djeddi, C., & Souici-Meslati, L. (2010). *A texture based approach for Arabic writer identification and verification*. In: *International Conference on Machine and Web Intelligence*. 115–120.

Djeddi, C., & Souici-Meslati, L. (2011). *Artificial immune recognition system for Arabic writer identification*. In: *International Symposium on Innovation in Information and Communication Technology*. 159–165. 10.1109/ISIICT.2011.6149612

Dryja, M., Jankowska, J., & Jankowski, M. (1982). *Survey of numerical methods and algorithms. Part II*. Warsaw, Poland: WNT.

Eldar, Y. C. (2001). *Quantum Signal Processing.* (Unpublished doctoral dissertation). Massachusetts Institute of Technology, USA.

Eldar, Y. C., & Oppenheim, A. V. (2002). Quantum Signal Processing. *IEEE Signal Processing Magazine*, *6*(19), 12–32. doi:10.1109/MSP.2002.1043298

Fortuna, Z., Macukow, B., & Wąsowski, J. (1982). *Numerical methods*. Warsaw, Poland: WNT.

Galloway, M. M. (1975). Texture analysis using gray level run lengths. *Computer Graphics and Image Processing*, *4*(2), 172–179. doi:10.1016/S0146-664X(75)80008-6

Garain, U., & Paquet, T. (2009). *Off-line multi-script writer identification using AR coefficients*. In: *International Conference on Document Analysis and Recognition*. 991–995. 10.1109/ICDAR.2009.222

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ghiasi, G., & Safabakhsh, R. (2013). Offline text-independent writer identification using codebook and efficient code extraction methods. *Image and Vision Computing*, *31*(5), 379–391. doi:10.1016/j.imavis.2013.03.002

Jakóbczak, D. (2005). Hurwitz-Radon matrices and their children. *Computer Science*, *5*(8), 29–38.

Jakóbczak, D. J. (2014). *2D Curve Modeling via the Method of Probabilistic Nodes Combination-Shape Representation, Object Modeling and Curve Interpolation-Extrapolation with the Applications*. Saarbrucken: LAP Lambert Academic Publishing.

Jankowska, J., & Jankowski, M. (1981). *Survey of numerical methods and algorithms. Part I*. Warsaw, Poland: WNT.

Kontaxakis, G., & Strauss, L. G. (1998). Maximum likelihood algorithms for image reconstruction in Positron Emission Tomography. *Radionuclides for Oncology – Current Status and Future Aspects, 1998,* 73-106.

Kowalczuk, Z., & Wiszniewski, B. (Eds.). (2007). *Intelligent data mining in diagnostic purposes: Automatics and informatics*. Gdansk, Poland: PWNT.

Kundur, D., & Hatzinakos, D. (1998). A novel blind deconvolution scheme for image restoration using recursive filtering. *IEEE Transactions on Signal Processing*, *2*(46), 375–390. doi:10.1109/78.655423

Laine, A., & Zong, X. (1996). *Border identification of echocardiograms via multiscale edge detection and shape modeling*. Paper presented at the IEEE International Conference on Image Processsing, Lausanne, Switzerland. 10.1109/ICIP.1996.560486

Lang, S. (1970). *Algebra*. Reading, Massachusetts, USA: Addison-Wesley Publishing Company.

Le Buhan Jordan, C., Bossen, F., & Ebrahimi, T. (1997). *Scalable shape representation for content based visual data compression*. Paper presented at the International Conference on Image Processing, Santa Barbara, CA, USA. 10.1109/ICIP.1997.647962

Marker, J., Braude, I., Museth, K., & Breen, D. (2006). Contour-based surface reconstruction using implicit curve fitting, and distance field filtering and interpolation. *Volume Graphics*, *2006*, 1–9.

Marti, U.-V., & Bunke, H. (2002). The IAM-database: An English sentence database for offline handwriting recognition. *Int. J. Doc. Anal. Recognit.*, *5*(1), 39–46. doi:10.1007100320200071

Meyer, Y. (1993). *Wavelets: algorithms & applications*. Philadelphia, USA: Society for Industrial and Applied Mathematics.

Nosary, A., Heutte, L., & Paquet, T. (2004). Unsupervised writer adaption applied to handwritten text recognition. *Pattern Recognition Letters*, *37*(2), 385–388. doi:10.1016/S0031-3203(03)00185-7

Ozaki, M., Adachi, Y., & Ishii, N. (2006). *Examination of effects of character size on accuracy of writer recognition by new local arc method*. In: International Conference on Knowledge- Based Intelligent Information and Engineering Systems. 1170–1175. 10.1007/11893004_148

Poggio, T., & Smale, S. (2003). The mathematics of learning: Dealing with data. *Notices of the American Mathematical Society*, *5*(50), 537–544.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Przelaskowski, A. (2005). *Data compression*. Warsaw, Poland: BTC.

Rutkowski, L., Siekmann, J., Tadeusiewicz, R., & Zadeh, A. (Eds.). (2004). *Lecture notes on artificial intelligence: Artificial intelligence and soft computing*. Berlin-Heidelberg, Germany: Springer-Verlag.

Schlapbach, A., & Bunke, H. (2004). *Using HMM based recognizers for writer identification and verification*. *9th Int. Workshop on Frontiers in Handwriting Recognition*. 167–172. 10.1109/IWFHR.2004.107

Schlapbach, A., & Bunke, H. (2006). *Off-line writer identification using Gaussian mixture models*. In: *International Conference on Pattern Recognition*. 992–995.

Schlapbach, A., & Bunke, H. (2007). A writer identification and verification system using HMM based recognizers. *Pattern Analysis & Applications*, *10*(1), 33–43. doi:10.100710044-006-0047-5

Schomaker, L., Franke, K., & Bulacu, M. (2007). Using codebooks of fragmented connected- component contours in forensic and historic writer identification. *Pattern Recognition Letters*, *28*(6), 719–727. doi:10.1016/j.patrec.2006.08.005

Shahabinejad, F., & Rahmati, M. (2007). A new method for writer identification and verification based on Farsi/Arabic handwritten texts, *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*. 829–833.

Siddiqi, I., Cloppet, F., & Vincent, N. (2009). Contour based features for the classification of ancient manuscripts. In: Conference of the International Graphonomics Society. 226–229.

Siddiqi, I., & Vincent, N. (2010). Text independent writer recognition using redundant writing patterns with contour-based orientation and curvature features. *Pattern Recognition Letters*, *43*(11), 3853–3865. doi:10.1016/j.patcog.2010.05.019

Vakhania, N. (1993). Orthogonal random vectors and the Hurwitz – Radon-Eckmann theorem. *Proc. of the Georgian Academy of Sciences-Mathematics, 1(1)*, 109-125.

Van, E. M., Vuurpijl, L., Franke, K., & Schomaker, L. (2005). The WANDA measurement tool for forensic document examination. *J. Forensic Doc. Exam.*, *16*, 103–118.

Willis, M. (2000). *Algebraic reconstruction algorithms for remote sensing image enhancement*. Unpublished doctoral dissertation, Department of Electrical and Computer Engineering, Brigham Young University.

Xu, Fang, & Mueller, K. (2005). Accelerating popular tomographic reconstruction algorithms on commodity PC graphics hardware. *IEEE Transactions on Nuclear Science*, *3*(52), 654–661.

Zaletelj, J., & Tasic, J. F. (2003). *Optimization and tracking of polygon vertices for shape coding*. Berlin-Heidelberg, Germany: Springer-Verlag. doi:10.1007/978-3-540-45179-2_52

Zhang, J. K., Davidson, T., & Wong, K. M. (2004). Efficient design of orthonormal wavelet bases for signal representation. *IEEE Transactions on Signal Processing*, *7*(52), 1983–1996. doi:10.1109/TSP.2004.828923

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

Contour: Closed curve.

Curve Interpolation: Computing new and unknown points of a curve and creating a graph of a curve using existing data points – interpolation nodes.

Data Extrapolation: Calculation of unknown values for the points situated outside the ranges of nodes.

Hurwitz – Radon Matrices: A family of skew – symmetric and orthogonal matrices with columns and rows that create, together with identical matrix, the base in vector spaces of dimensions $N = 2, 4$ or 8.

MHR Method: The method of curve interpolation and extrapolation using linear (convex) combinations of OHR operators.

OHR Operator: Matrix operator of Hurwitz – Radon built from coordinates of interpolation nodes.

Shape Representation: Data that are used in shape reconstruction.

Value Anticipation: Foreseeing next value when last value is known.

This research was previously published in Probabilistic Nodes Combination (PNC) for Object Modeling and Contour Reconstruction edited by Dariusz Jacek Jakóbczak; pages 132-177; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 30
Statistical Features for Extractive Automatic Text Summarization

Yogesh Kumar Meena
MNIT Jaipur, India

Dinesh Gopalani
MNIT Jaipur, India

ABSTRACT

Automatic Text Summarization (ATS) enables users to save their precious time to retrieve their relevant information need while searching voluminous big data. Text summaries are sensitive to scoring methods, as most of the methods requires to weight features for sentence scoring. In this chapter, various statistical features proposed by researchers for extractive automatic text summarization are explored. Features that perform well are termed as best features using ROUGE evaluation measures and used for creating feature combinations. After that, best performing feature combinations are identified. Performance evaluation of best performing feature combinations on short, medium and large size documents is also conducted using same ROUGE performance measures.

INTRODUCTION

Day by day the information content on World Wide Web is exponentially increasing (BIG Data) along with the increase in the number of web users. In current scenario volume of information available on World Wide Web exceeded the textual information available in the printed form in libraries. The rapid growth of online information services makes it difficult to retrieve the relevant information quickly. While searching a particular information content, many a times information system users realize that the extracted information using currently available popular tools, that is present in the form of text, is not relevant to their information need at all, even after reading the whole list of text documents. They perhaps only waste their valuable time in reading the irrelevant text document. This problem can be

DOI: 10.4018/978-1-7998-0951-7.ch030

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

solved if users can be provided with a summary of the given text document. However, due to a large volume of available text data, that too dynamic in nature, it is very cumbersome for human experts to summarize all the documents manually. This issue leads to the requirement of an automated system that summarize the given text document automatically. This system that condenses the text document automatically and preserving its overall information content using a computer is termed as Automatic Text Summarization (ATS) system. There are numerous applications of automatic text summarization such as a snippet in information retrieval systems, news headlines as replacement to full story of news, electronic program guide in television systems. ATS is widely used for various domains such as business, news, legal and medical domains.

The researchers have classified text summarization systems (Hovy & Lin, 1995; Jones, 1997; Mani & Maybury, 1999; Gupta & Lehal, 2010; Lloret & Palomar, 2012; Nenkova & McKeown, 2012) on the basis of three main perspectives namely input, purpose and output. The third aspect which is most popular among these (i.e. output) considers summarization systems as either extractive or abstractive, and is mostly used by researchers. In extractive automatic text summarization (EATS), a subset of sentences from the original text document is selected for the final summary. Whereas abstractive automatic text summarization (AATS), sentences are fused and regenerated using natural language resources or rules. Abstractive text summarization requires deep knowledge resources, lexical/language resources, parsers and language generators. Because of these resource requirements, it is practically infeasible to use abstractive methods for automatic text summarization. Therefore, researches mainly focused on extractive text summarization instead of abstractive text summarization. The research work carried out is also focused on extractive automatic text summarization.

A typical extractive text summarization process completes in three steps namely pre-processing, sentence scoring, and summary generation. The first step in text summarization is preprocessing the input text document. Preprocessing mainly includes the sentence segmentation, stemming, stopword removal and special symbol removal. In the step which is sentence scoring, a score is assigned to each sentence of the document based on certain specific criteria. In literature various features are defined for sentence scoring such as term frequency, numerical data inclusion, sentence location, etc. After calculating each sentence score, rank is assigned to each sentence based on these scores. The sentences with the higher score are considered more important than the sentences with lower scores and are assigned high ranks consequently. After assigning a rank to each sentence, in third step, top ranked sentences are selected to generate the summary. Total number of sentences in the summary depends on the required length of it.

This chapter is focused on statistical features used in extractive approaches. In this chapter performance of these features is analyzed individually. After that, features that perform well termed as best features using ROUGE evaluation measures are used for creating feature combinations. After that, best performing feature combinations are identified. A deep performance evaluation of best performing feature combinations on short, medium and large size documents is also conducted using same ROUGE performance measures.

The rest of the chapter is organized as follows: After discussing challenges of ATS, background methods are presented followed by features used for sentence scoring. . Next, author discusses generic summarization process used for ATS. This is subsequently followed by evaluation matrices for ATS, followed by proposed impact analysis. Features and their combinations are properly analyzed before the chapter is finally concluded.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

GENERAL CHALLENGES IN TEXT SUMMARIZATION

Major issues (Jones, 1998, Mani & Maybury, 1999) in generating automatic extractive text summaries are listed below:

1. **Redundancy:** Extractive text summaries are generated by selecting top scoring sentences from the text document. This sometimes led to the selection of redundant sentences in the final summary.
2. **Cohesion:** Sometimes, top scoring sentences becomes totally different from each other. Each neighboring sentence contains totally different terms without any of context in between. It becomes an issue, because if sentence are not cohesive then it becomes difficult to interpret the intended gist of text.
3. **Coherence:** Textual relationship beyond more than one sentence and through a discourse, if not meaningful then there is no meaning of generated summary. As selected sentences on the basis of some scoring function might not produce quality summary due to lack of coherence in between the selected ones.
4. **Anaphora:** Scoring function do not reproduce sentence as they only select the sentences as it is from the original text, therefore referring pronouns in the resulted summary creates issues in terms of interpretation and resolution.
5. **Coverage:** Extractive summaries do not guarantee that all inherent topics or themes are present in the finally produced summary as selection is purely on the basis of score of sentences obtained using some scoring methodology.

BACKGROUND AND ISSUES

Work in the area of EATS started in the early 1950s, however as of now; no such system is available that can generate summaries as efficient as experts. In early stage, pioneered features such as term frequency (Luhn, 1958), sentence location (Baxendale, 1958), cue words and title similarity (Edmundson, 1969) were proposed for text summarization process. As the time progressed, new features were added for the text summarization process like tf-isf, gain, etc. The basic methodology behind these summarization processes is to assign score to each of the sentences present in the text by using either a single feature or linear sum of features. However, applying features alone do not improve the quality of text summaries significantly, as each feature processes sentences in a different aspect. In literature many feature based statistical, graph based, lexical, machine learning and evolutionary methods for extractive summarization are proposed by researchers.

The key component of the statistical approaches(Rush et al.,1971; Pollock & Zamora, 1975; Brandow et al., 1995; Salton, 1983; Church & Gale, 1995; Hovy & Lin, 1997; Hovy & Lin, 1998; Baldwin & Morton, 1998; Radev et al., 2001; Nobata et al., 2001; Mori, 2002; Nobata et al., 2002; Pingali et al., 2005; Liu et al., 2009) of extractive text summarization is assigning weights to words or sentences based on certain statistics such as frequency of appearance of words in the text. Statistical approaches are simple to use as no knowledge source is required in these approaches. However problems with these approaches are similar to extractive text summarization approaches such as problem of ambiguous references i.e. anaphors (such as pronouns which refers to some words that appears earlier in the text) and cataphors (ambiguous words which signals to word that appears later in the text).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In graph based approaches (Michalcea & Tarau, 2004), sentence scores are generated based on the relationship among the sentences. Although graph based approaches work well as graph based approaches do not rely only on the local context of a text unit (vertex), rather take into account information recursively drawn from the entire text (graph). However graph based approaches also have the issue of ambiguous references. In addition to this issue, graph based approaches also have an issue of the requirement of multiple iterations as these approaches require multiple iterations to converge.

Later machine learning methods and evolutionary methods (Kupiec, 1995; Hirao et al., 2002; Varma et al., 2005; Fattah & Ren, 2009; Prasad & Kulakarni, 2010; Xiaoyan et al., 2010; Abuobieda et al., 2013; Abuobieda et al., 2012; Wang & Maches, 2013; Mendoza et al., 2014) were proposed with the increase number of features. These methods, used to identify a suitable set of features and their applicable weights. On the trained data the performance, most of the times, is optimal; however applying the same process to unknown data may result in the generation of poor quality summaries. This issue of generation of poor quality summaries for unknown data is known as the generalization issue. Although machine learning algorithms allow to test performance with a high number of features in an easy way, however at the same time, all these approaches require labelled corpus of sentences. This requirement of labelled corpus becomes a critical issue while summarization because it is difficult to mark summary sentences similar to abstracts, manually. Also these methods are not able to resolve the issue of coherence and cohesiveness.

Semantic methods such as feature based semantic similarity and lexical chains were proposed to alleviate the issue of coherence and cohesiveness. In lexical chain based approaches, lexical chains are constructed after the pre-processing step. Thereafter in the next step, scoring of the chains is conducted. For this purpose various matrices are applied such as chain length, chain distribution in the text, text span covered by the chain, density, graph topology (diameter of the graph of the words) and number of repetitions. The main issue with lexical chain based methods is the requirement of knowledge sources on hand such as WordNet to find semantically related words and chains, and secondly interpretation of semantic information such as semantic similarity threshold. These practical issues limits and puts constraints on the usefulness of semantic methods.

In order to improve the performance of extractive text summarization systems, the information-rich sentences need to be selected to produce the final summary. Hence sentence selection is the most important step of extractive text summarization process. For the purpose a numerical measure of usefulness may be assigned to each of the sentences in the given text. Information rich sentences are then selected according to the specified heuristics. Aforesaid weight assigning process by using a specific heuristic is termed as sentence scoring. For the final summary, sentences are selected on the basis of their scores computed using different scoring methods. Therefore, the overall performance of extractive text summarization system mainly depends on the sentence scoring methods employed to score the sentences. Several other factors like pre-processing, also influences the quality of text summaries.

From early 1960s, researchers are continuously trying to analyze the feature from different perspectives in order to generate efficient text summaries. Rafael et al., 2013 assessed a broader set of features that includes term frequency, tf-isf, upper case, proper noun, word co-occurrence, lexical similarity, cue word, numerical data, sentence length, sentence position, title similarity, aggregate similarity, bushy path and textrank. They tested each and every algorithm for three different datasets. It was concluded that term frequency, tf-isf, sentence length, lexical similarity and text rank are better features for sentence scoring as compared to the rest of the features listed. Later Rafael et al., 2014 analyzed the performance of previously proposed algorithms using different combinations. The proposed methodology was termed

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

as context based summarization and three groups of features were created namely word level, sentence level and graph level. All combinations of intra-group and inter-group word level, sentence level and graph level features are tried to test their impact on sentence scoring. This approach concluded that word frequency; title similarity and sentence location are the best features specifically for News domain. However, in all assessments listed above impact analysis of features such as sentence entropy, sentence to sentence cohesion etc. has not been carried out and a single evaluation measure ROUGE-1 for extracting best feature combinations is used. Moreover, some semantic features are also analyzed with statistical features. However, lack of abundant knowledge sources makes it difficult to use semantic features. This motivated to use exhaustive set of features for impact analysis other than semantic features (due to lack of knowledge sources) and evaluation measures (ROUGE-1, ROUGE-2, ROUGE-L and ROUGE-W).

FEATURE ENGINEERING

A typical sentence scoring module of extractive automatic text summarization requires various features to score sentences. A number of statistical and linguistics features have been proposed by researchers for the aforesaid scoring process. These features are categorized into three groups of categories according to their processing level. The categories are word level scoring, sentence level scoring and graph based scoring and these are described as following:

1. **Word Level Scoring:** The features used for word level scoring methods score each term present in the given text according to a particular criterion. Sentence score is computed as the sum of each term present in a given sentence.
2. **Graph Level Scoring:** The sentence score in case of graph level scoring method is calculated using the relationship of a sentence with other sentences. If two sentences are linked via some defined parameter, then an edge is established in between them. Their weights are used to generate sentence scores.
3. **Sentence Level Scoring:** The features used for sentence level scoring methods use sentence level statistics for computing score for a given sentence.

All features are summarized in Table 1.

GENERIC SUMMARIZATION PROCESS

A typical extractive text summarization process completes in three steps namely pre-processing, sentence scoring and summary generation. The details of process flow are discussed in the following subsection. Whole generic text summarization process flow consisting of three steps is shown in Figure 1.

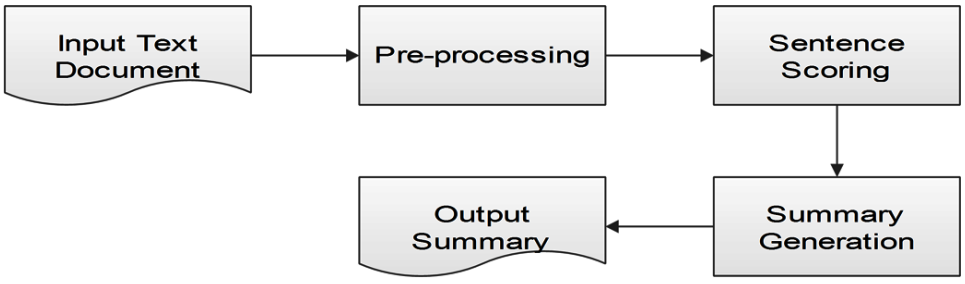
1. **Pre Processing:** First step in text summarization is preprocessing the input text document. Preprocessing mainly includes the following steps sentence segmentation, stemming (Porters, 1980), stopword removal and special symbol removal.
2. **Feature Based Scoring:** After preprocessing, the next step in generic text summarization is sentence scoring. In this step, a score is assigned to each sentence of the document based on certain specific

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Feature algorithms

S. No.	Short Name	Author, Year	Feature Algorithm Name	Level
1	FA1	Salton, 1997	Bushy Path	Graph Level
2	FA2	Neto et al., 2002	Cosine Similarity With Title	Sentence Level
3	FA3	Edmundson, 1969	Cue Words	Word Level
4	FA4	Neto et al., 2002	Depth of Sentence in the Tree	Sentence Level
5	FA5	Mori, 2002	Gain	Word Level
6	FA6	Torres et al., 2002	Hamming Distance	Sentence Level
7	FA7	Torres et al., 2002	Hamming Weight	Sentence Level
8	FA8	Torres et al., 2002	Interaction Between the Sentences	Sentence Level
9	FA9	Erkan et al., 2004	LexRank	Graph Level
10	FA10	Torres et al., 2002	Named Entity	Word Level
11	FA11	Lin, 1999	Numerical Data	Sentence Level
12	FA12	Salton, 1997	Aggregate Similarity	Graph Level
13	FA13	Torres et al., 2002	Sentence Entropy	Sentence Level
14	FA14	Baxendale, 1958	Sentence Location1	Sentence Level
15	FA15	Nobata et al., 2002	Sentence Location2	Sentence Level
16	FA16	Neto et al., 2002	Sentence to Centroid Cohesion	Sentence Level
17	FA17	Neto et al., 2002	Sentence to Sentence Cohesion	Sentence Level
18	FA18	Torres et al., 2002	Sum of Probability	Sentence Level
19	FA19	Torres et al., 2002	Sum of Hamming Weights of Words by Frequency	Sentence Level
20	FA20	Luhn, 1958	Term Frequency	Word Level
21	FA21	Mihalcea et al., 2004	TextRank	Graph Level
22	FA22	Brandow et al., 1995	TFIDF	Word Level
23	FA23	Edmundson, 1969	Title Similarity	Sentence Level
24	FA24	Torres et al., 2002	Total Terms in Sentence	Sentence Level
25	FA25	Liu et al., 2009	Word Co-occurrence	Sentence Level

Figure 1.



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

criteria. Various features are defined for sentence scoring such as term frequency, numerical data inclusion, sentence location, etc. After calculating each sentence score, rank is assigned to each sentence based on these scores of the sentences. The sentences with the higher score are considered as more important than the sentences with lower scores. Ranks are assigned to each sentence in the given text according to their scores.

3. **Summary Generation:** After assigning a rank to each sentence, the final step in generic text summarization is summary generation. In this step, top ranked sentences are selected to generate the summary. The number of sentences to be selected depends on the required length of the summary.

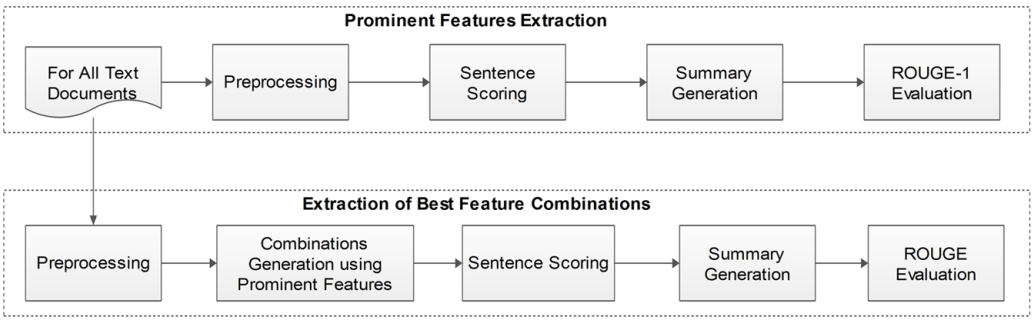
Evaluation Matrices

Automatic evaluation of text summaries is a challenging and difficult task. The gold summaries provided by human experts and summary generated by automatic text summarization systems usually do not match exactly. Lin, 2004 proposed Recall-Oriented Understudy for Gisting Evaluation (ROUGE) measures for evaluation of text summarization systems. ROUGE measures are the only available commonly used measures for the purpose of automatic evaluation. ROUGE-1 and ROUGE-2 are versions of ROUGE-N, where N is 1 and 2, respectively. ROUGE-L and ROUGE-W are computed using longest common sub-sequences.

Proposed Impact Analysis

As discussed earlier, extractive text summarization in general completes in three steps namely: pre-processing, sentence scoring and summary generation. Process flow to analyze the impact of feature combinations is shown in Figure 2. Pre-processing is applied on the text document using the process described earlier. Each sentence is then scored using all the features one by one. After sentence scoring, summary is generated using each feature. ROUGE-1 evaluation measure as described in previous section is then applied on the summary and relevant features are further used for generation of feature combinations. These relevant features are termed as prominent features. For each feature combination, a summary is generated. ROUGE evaluation measures (ROUGE-1, ROUGE-2, ROUGE-L, ROUGE-W) are further applied to get the better performing feature combinations which are termed as best feature combinations.

Figure 2.



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

RESULTS AND DISCUSSIONS ON DUC 2002 DATASET

Document Understating Conference (DUC) 2002 dataset is used for experimental evaluation of proposed impact analysis. This dataset contains News data along with their gold summaries (manual summaries provided by humans). Sixty reference documents sets, each of them consisting of approximately ten documents for the evaluation task are provided in DUC 2002. For each document, DUC 2002 provides two abstractive summaries as gold summaries with about hundred words each.

After summary generation using a specific sentence scoring ROUGE-1, ROUGE-2, ROUGE-L and ROUGE-W scores are computed for each document. Initially, for each feature algorithm precision, recall and f-Measure value is calculated for each document. Thereafter, features that perform well are identified and termed as prominent features. These feature algorithms are then used for generating feature combinations. The feature algorithm combinations which perform well are termed as best feature combinations.

Using prominent feature algorithms, different feature algorithm combinations are generated for sentence scoring. Initially, Best forty combinations for each evaluation measure ROUGE-1, ROUGE-2, ROUGE-L and ROUGE-W are selected. Combinations which are present in the lists of best performing combinations, for each of the four evaluation measures are selected as final best feature algorithm combinations. These combinations are listed in Table 2. The list comprised of features algorithms busy path, cosine similarity with title, depth of the sentence in the tree, lexrank, aggregate similarity, sentence location1, sentence to sentence cohesion and title similarity. Despite having better performance individually, sentence entropy did not produce higher ROUGE scores in combinations with other feature algorithms. All best combinations are renamed as COMB with number in the same order of their respective rank for ROUGE-2 F-measure, for the purpose of comparison in graphical form. For example: first best performing combinations is renamed as COMB1.

Pre-processing is an important step of any text processing system. Final results are very much sensitive to this pre-processing. Comparison of results obtained using ROUGE-1 evaluation measure on all documents for single feature algorithms, with stemming and without stemming is shown in Figure 3.

It is observed that results using stemming are better as compared to results without stemming. The fact that stemming use the root forms of the words, enhances the performance. Comparing terms using their stemmed version gives more number of matches, which ultimately improves the results. Comparison of results obtained using ROUGE-2 evaluation measure on all documents for single feature algorithms, with stemming and without stemming is shown in Figure 4. Here also, the performance of stemmed version is better as compared without non-stemmed version.

Precision, Recall and F-Measure comparisons of best feature algorithm combinations for all documents set using ROUGE-1 and ROUGE-2 are presented in Figure 5 and Figure 6 respectively. Results are sorted on the basis of f-measure. It is observed that their are notable variations in ROUGE scores of precision and recall for some of the combinations. This is primarily due to the size of summary to be produced i.e. compression rate.

In proposed experiments, the compression rate is hundred words. At the time, total words in the summary exceed the limit beyond hundred words, the variation in between precision and recall is visible. More the limit exceeds, more the difference is reported in precision and recall. The smooth curve shows that COMB1 performs better as compared to all other feature algorithm combinations. ROUGE-L and ROUGE-W comparisons for the similar group are shown in Figure 7 and Figure 8, respectively.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Best feature algorithm combinations details

S. No.	Combination ID	Features in Combination
1	COMB1	Aggregate Similarity + LexRank + Sentence Location1
2	COMB2	LexRank + Sentence Location1
3	COMB3	Aggregate Similarity + LexRank + Sentence Location1 + Sentence to Sentence Cohesion
4	COMB4	Bushy Path + LexRank + Sentence Location1 + Sentence to Sentence Cohesion
5	COMB5	LexRank + Sentence Location1 + Sentence to Sentence Cohesion + Sentence to Centroid Cohesion
6	COMB6	LexRank + Sentence Location1 + Sentence to Sentence Cohesion
7	COMB7	Depth of the Sentence in the Tree + LexRank + Sentence Location1
8	COMB8	Aggregate Similarity + LexRank + Sentence Location1 + Sentence to Sentence Cohesion + Sentence to Centroid Cohesion
9	COMB9	Depth of the Sentence in the Tree + LexRank + Sentence Location1 + Sentence to Centroid Cohesion
10	COMB10	LexRank + Sentence Location1 + Sentence to Sentence Cohesion + Term Frequency
11	COMB11	Aggregate Similarity + Depth of the Sentence in the Tree + LexRank + Sentence Location1 + Sentence to Sentence Cohesion
12	COMB12	Depth of the Sentence in the Tree + LexRank + Sentence Location1 + Term Frequency
13	COMB13	Aggregate Similarity + Depth of the Sentence in the Tree + LexRank + Sentence Location1
14	COMB14	Aggregate Similarity + LexRank + Sentence Location1 + Sentence to Sentence Cohesion + Term Frequency
15	COMB15	Sentence Location1 + Sentence to Sentence Cohesion
16	COMB16	Aggregate Similarity + Depth of the Sentence in the Tree + LexRank + Sentence Location1 + Sentence to Centroid Cohesion
17	COMB17	Aggregate Similarity + Frequential Sum of Probability + LexRank + Sentence Location1 + Sentence to Sentence Cohesion
18	COMB18	Aggregate Similarity + Sentence Location1 + Sentence to Sentence Cohesion
19	COMB19	Aggregate Similarity + Depth of the Sentence in the Tree + LexRank + Sentence Location1 + Term Frequency
20	COMB20	Depth of the Sentence in the Tree + Sentence Location1 +Term Frequency
21	COMB21	Aggregate Similarity + Sentence Location1
22	COMB22	Similarity + Depth of the Sentence in the Tree + Hamming distance + LexRank + Sentence Location1

Similar to ROUGE-1 and ROUGE-2, it is observed that COMB1 performs better in case of ROUGE-L and ROUGE-W evaluation as well. COMB1 comprised of feature algorithms aggregate similarity (Salton, 1997), lexrank (Erkan, 2004) and sentence location1. This is primarily due to using graph based feature algorithms such as aggregate similarity (Salton, 1997) and lexrank (Erkan, 2004) which focuses on the sentences which have more number of interlinks. In rest of the other combinations, feature algorithms such as depth of the sentence in the Tree (Neto et al., 2002), frequential sum of probability (Torres et al., 2002), hamming distance (Torres et al., 2002) sentence to sentence cohesion (Neto et al., 2002)}, term frequency (Luhn, 1958) and title similarity (Edmundson, 1969) are present other then the three features present in COMB1.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3.

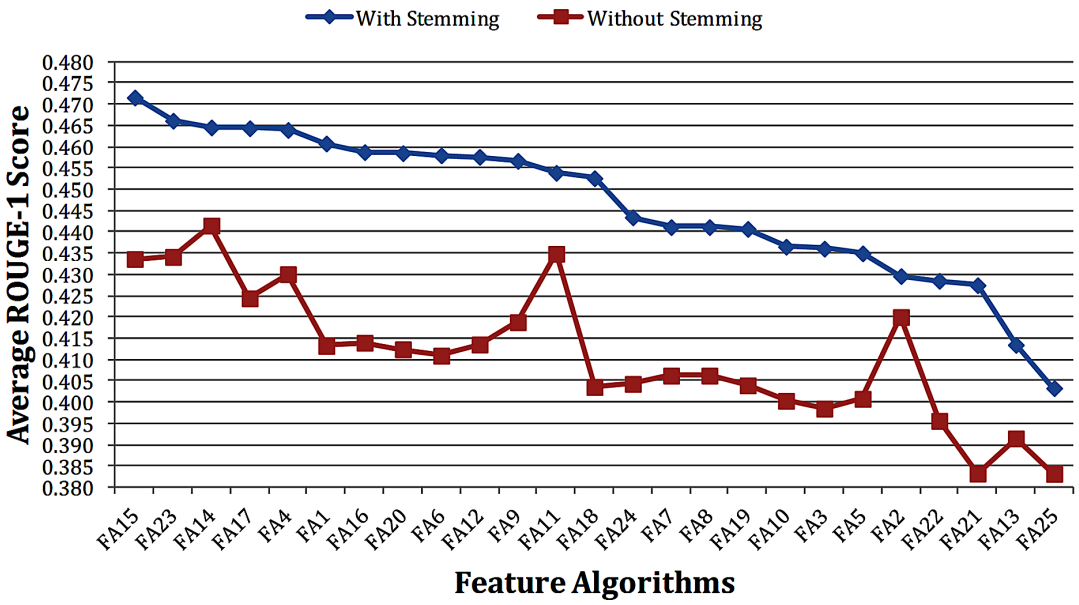
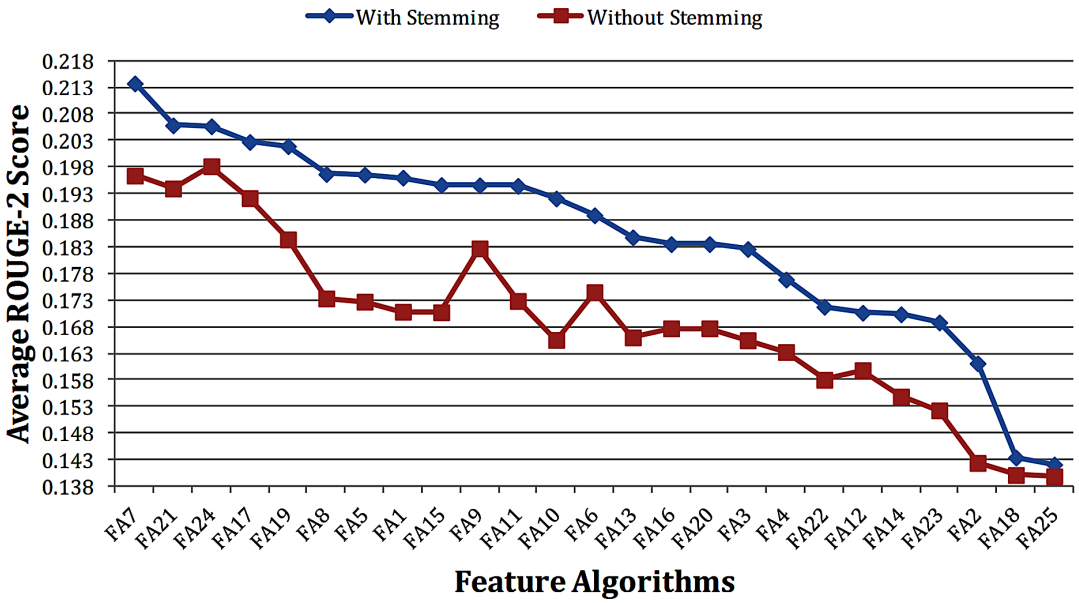


Figure 4.



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5.

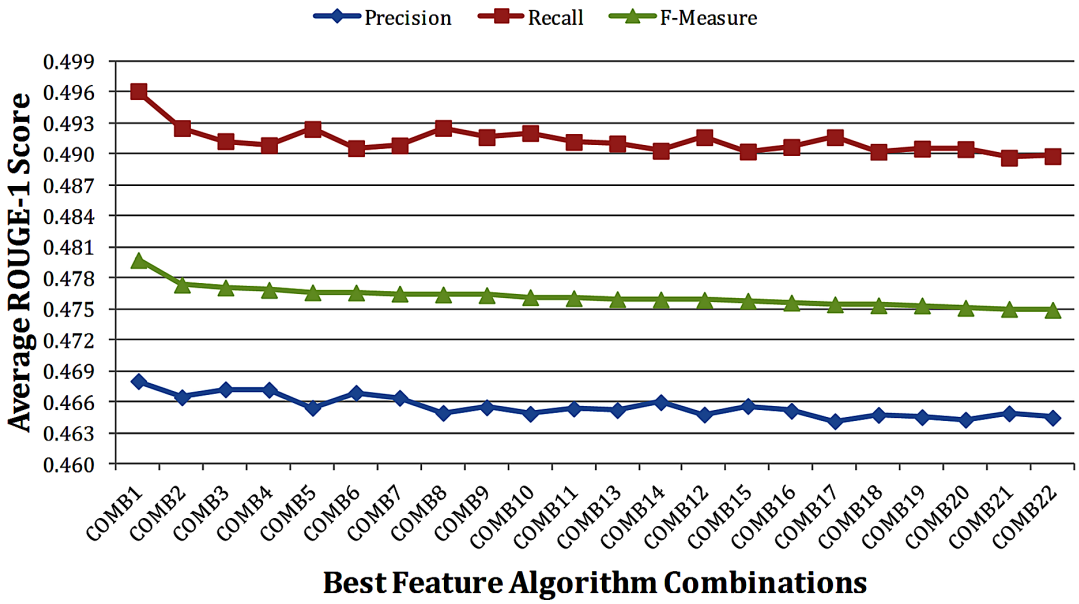
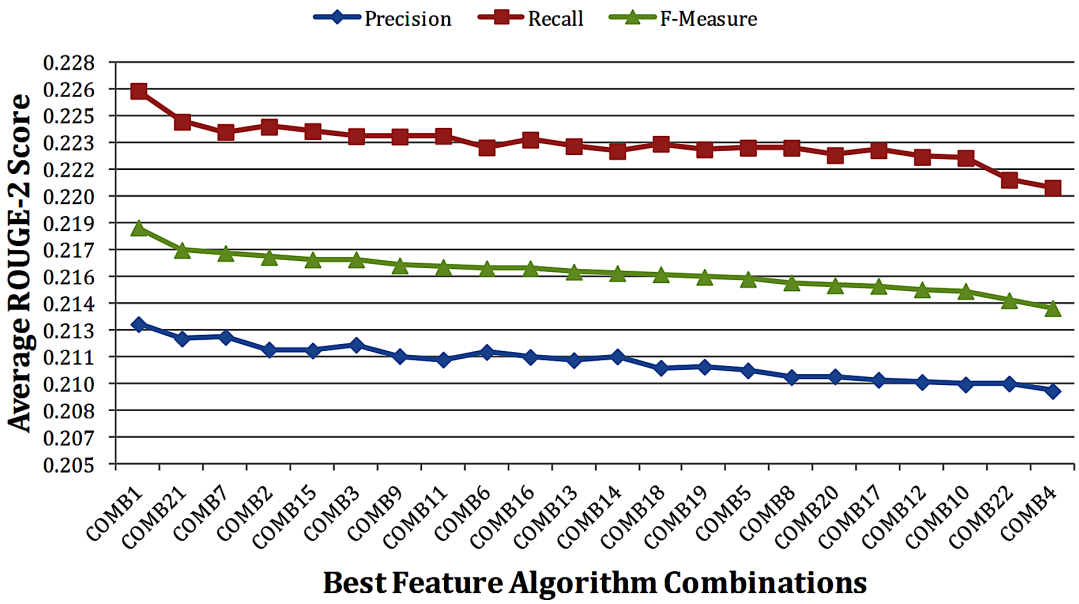


Figure 6.



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7.

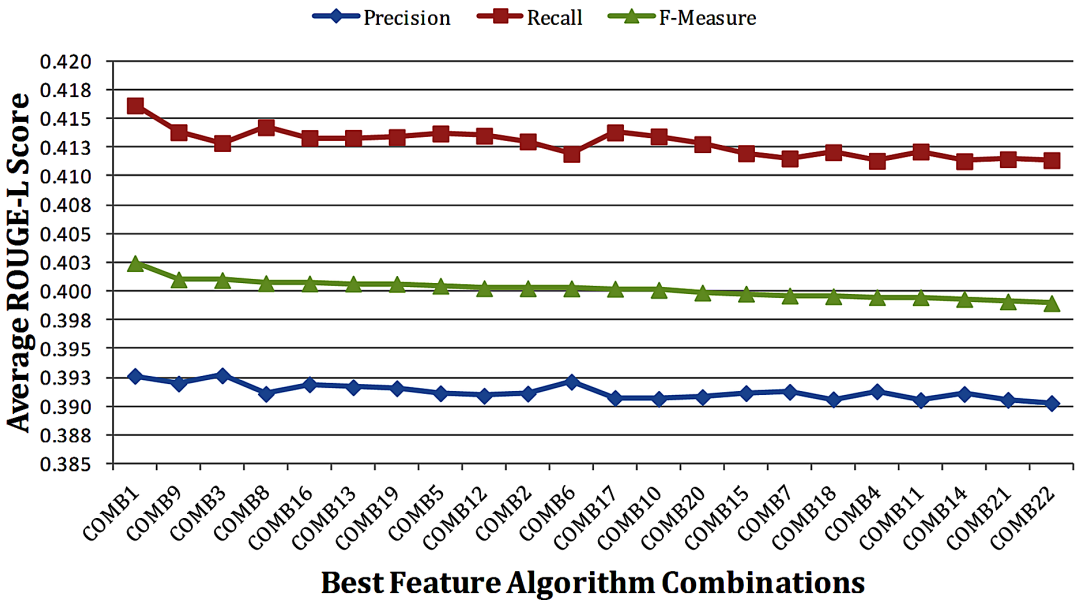
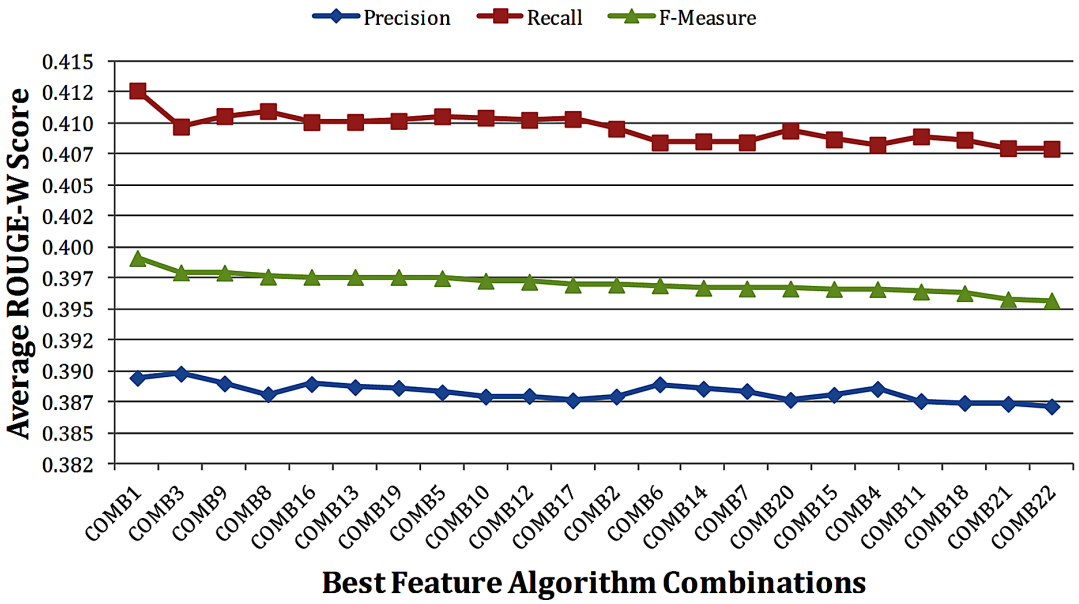


Figure 8.



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

All the best twenty two feature algorithm combinations differ slightly in ROUGE scores. Comparing with individual feature algorithms, these feature algorithm combinations performs better as shown in Figure 9. It is clearly visible that all best feature algorithm combinations performs better as compared to all individual algorithms. Similar to individual feature algorithms, it is not a good idea to use a single feature algorithm combination for summary generation. As different feature algorithm combinations have different inherent properties. In comparison with work carried out by Rafael et al. (Rafael et al., 2013; Rafael et al., 2014) reported best combination is COMB1 contains two graph level and one sentence level features. Moreover, no dependent features such as cue words (Edmundson, 1969), named entity (Nobata et al., 2002) etc. are present in the reported best combinations.

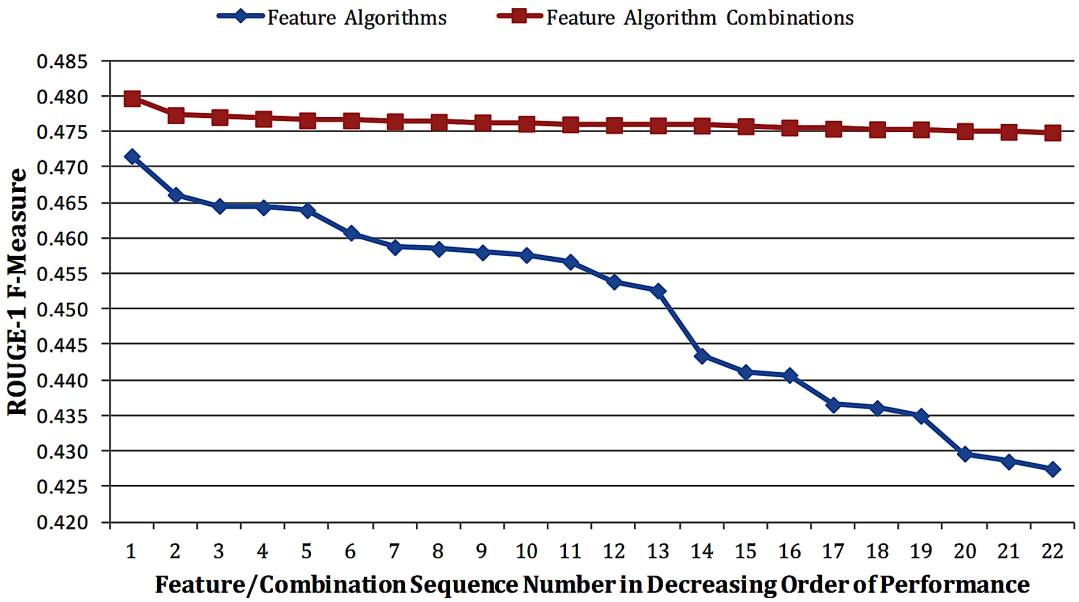
Comparisons in between ROUGE-1 scores for short size, medium size, large size and all set of documents are shown in Figure 10. It is observed from the figure that performance of summarization system decreases as the size of document increases. This degradation is due to increase in the number of sentences because as the number of sentences increases, probability of selection of informative sentences decreases.

ROUGE-2 scores comparisons for short size, medium size, large size and all set of documents are shown in Figure 11. Here also performance decreases with the increase in the size of input text document.

CONCLUSION

In this chapter performance of various individual features algorithms such as bushy path, cosine similarity, cue words etc. is tested using ROUGE evaluation measures. It is observed that single feature algorithms only use specific feature strengths that are not capable of generating efficient summaries. Stemming impact is also analyzed while pre-processing of input text document; it is observed that with stemming performance improves significantly. Thereafter, feature algorithm combinations are generated

Figure 9.



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10.

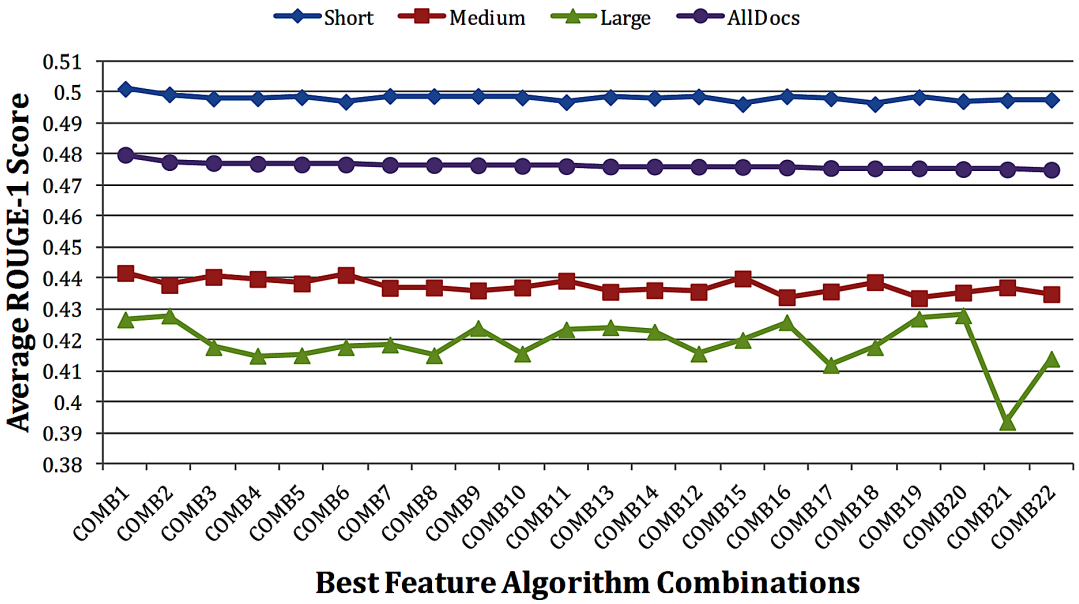
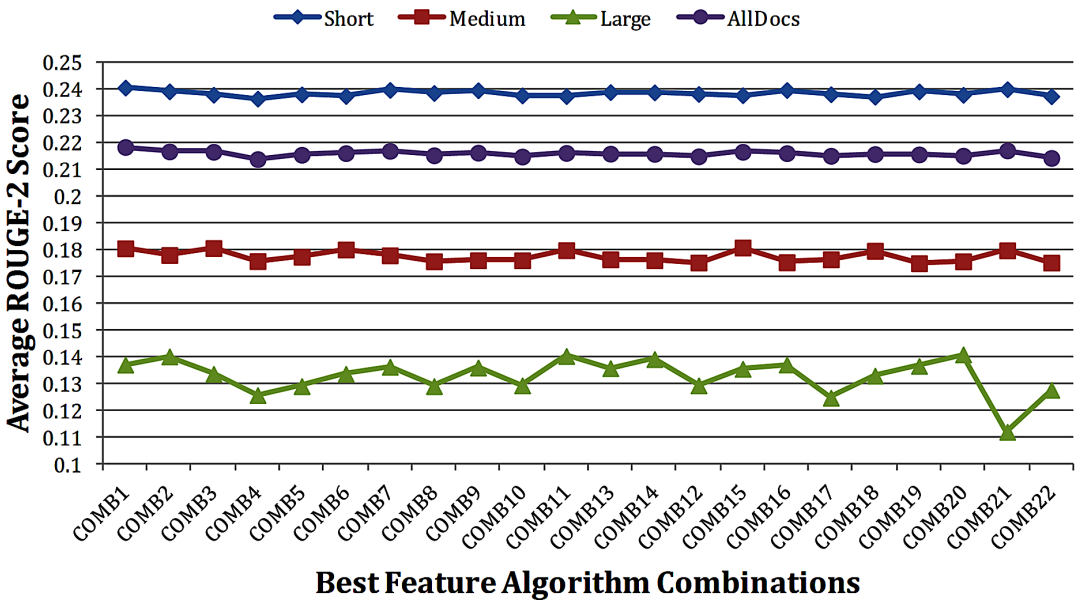


Figure 11.



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

using best performing single feature algorithms using ROUGE evaluation measures. A feature algorithm combination which performs well for all ROUGE evaluation measures are reported as best feature algorithm combinations. Their performance is investigated on short, medium and large size documents. It is observed that performance degrades as the size of document increases.

REFERENCES

Abuobieda, A., Salim, N., Binwahlan, M. S., & Osman, A. H. (2013). Differential evolution cluster-based text summarization methods. *Computing, Electrical and Electronics Engineering (ICCEEE), 2013 International Conference on*,. 10.1109/ICCEEE.2013.6633941

Abuobieda, N., Salim, A., Albaham, A., Osman, Y., & Kumar. (2012). Text summarization features selection method using pseudo genetic-based model. In *International Conference on Information Retrieval Knowledge Management (CAMP)*. 10.1109/InfRKM.2012.6204980

Aliguliyev, R. M. (2009). A new sentence similarity measure and sentence based extractive technique for automatic text summarization. *Expert Systems with Applications*, *36*(4), 7764–7772. doi:10.1016/j.eswa.2008.11.022

Baldwin, B., & Morton, T. S. (1998). Dynamic co reference based summarization. In *Proceedings of the Third Conference on Empirical Methods in Natural Language Processing (EMNLP-3)*.

Barzilay, R., & Elhadad, M. (1997). Using Lexical Chains for Text Summarization. In *Proceedings of the ACL/EACL'97 Workshop on Intelligent Scalable Text Summarization* (pp. 10–17).

Baxendale, P. B. (1958). Machine-made index for technical literature: An experiment. *IBM Journal of Research and Development*, *2*(4), 354–361. doi:10.1147/rd.24.0354

Binwahlan, M. S., Salim, N., & Suanmali, L. (2009). Swarm Based Text Summarization. In *Proceedings of the International Association of Computer Science and Information Technology - Spring Conference. IACSITSC '09* (pp. 145-150). 10.1109/IACSIT-SC.2009.61

Binwahlan, M. S., Salim, N., & Suanmali, L. (2010). Fuzzy swarm diversity hybrid model for text summarization. *Information Processing & Management*, *46*(5), 571–588. doi:10.1016/j.ipm.2010.03.004

Brandow, R., Mitze, K., & Rau, L. F. (1995). Automatic condensation of electronic publications by sentence selection. *Information Processing & Management*, *31*(5), 675–685. doi:10.1016/0306-4573(95)00052-I

Carpineto, C., Osinski, S., Romano, G., & Weiss, D. (2009). A survey of Web clustering engines. *ACM Computing Surveys*, *41*(3), 1–38. doi:10.1145/1541880.1541884

Church, K., & Gale, W. A. (1995). Inverse document frequency (idf): A measure of deviations from poisson. In *Proceedings of the Third Workshop on Very Large Corpora*.

Conroy, J., & O'leary, D. (2001). Text summarization via hidden Markov models. In *Proceedings of the 24th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 406-407). New Orleans, LA: ACM.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dehkordi, P.-K., Kumarci, F., & Khosravi, H. (2009). Text Summarization Based on Genetic Programming. *Proceedings of the International Journal of Computing and ICT Research*, *3*, 57–64.

Edmundson, H. P. (1969). New methods in automatic extracting. *Journal of the ACM*, *16*(2), 264–285. doi:10.1145/321510.321519

Fattah, M. A., & Ren, F. (2009). Ga, mr, ann, pnn and gmm based models for automatic text summarization. *Computer Speech & Language*, *23*(1), 126–144. doi:10.1016/j.csl.2008.04.002

Garcia, H. R., & Ledeneva, Y. (2013). Single Extractive Text Summarization Based on a Genetic Algorithm. In J. Carrasco-Ochoa, J. Martínez-Trinidad, J. Rodríguez, & G. Baja (Eds.), *Pattern Recognition* (Vol. 7914, pp. 374–383). Springer Berlin Heidelberg. doi:10.1007/978-3-642-38989-4_38

Gupta, V., & Lehal, G. S. (2010). A survey of text summarization extractive techniques. *Journal of Emerging Technologies in Web Intelligence, 2*(3), 258-268.

Hirao, T., Sasaki, Y., Isozaki, H., & Maeda, E. (2002). NTT's Text Summarization system for DUC-2002. In *Workshop on Text Summarization (In conjunction with the ACL 2002 and including the DARPA/NIST sponsored DUC 2002 Meeting on Text Summarization)*.

Hovy, E., & Lin, C.-Y. (1998). Automated text summarization and the summarist system. In *Proceedings of a Workshop on Association for Computational Linguistics*.

Jezek, K., & Steinberger, J. (2008). Automatic Text Summarization (The state of the art 2007 and new challenges). In Znalosti 2008 (pp. 1-12).

Jones, K. S. (1998). Automatic summarising: Factors and directions. In *Advances in Automatic Text Summarization* (pp. 1–12). MIT Press.

Kogilavani, A., & Balasubramanie, P. (2010). Clustering and Feature Specific Sentence Extraction Based Summarization of Multiple Documents. *International Journal of Computer Science and Information Technology*, *2*(4), 99–111. doi:10.5121/ijcsit.2010.2409

Kupiec, J., Pedersen, J., & Chen, F. (1995). A trainable document summarizer. In *Proceedings of the 18th annual international ACM SIGIR conference on Research and development in information retrieval*. Seattle, WA: ACM.

Lewandowski, D. (2005). Web searching, search engines and Information Retrieval. *Inf. Serv. Use, 25*(3-4), 137-147.

Lin, C.-Y. (2004). Rouge: a package for automatic evaluation of summaries. In *Proceedings of the ACL-04 Workshop on Text Summarization Branches Out* (pp. 74-81).

Lin, C.-Y., & Hovy, E. (1997). Identifying topics by position. In *Proceedings of the fifth conference on Applied natural language processing (ANLC '97)*. Association for Computational Linguistics. 10.3115/974557.974599

Litvak, M., Last, M., & Friedman, M. (2010). A new approach to improving multilingual summarization using a genetic algorithm. In *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics* (pp. 927-936). Uppsala, Sweden: Association for Computational Linguistics.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Liu, X., Webster, & Kit, C. (2009). An Extractive Text Summarizer Based on Significant Words. In *Proceedings of the 22nd International Conference on Computer Processing of Oriental Languages. Language Technology for the Knowledge-based Economy* (ICCPOL '09). Springer-Verlag.

Lloret, E., & Palomar, M. (2012, January). Text summarization in progress: A literature review. *Artificial Intelligence Review*, *37*(1), 1–41. doi:10.100710462-011-9216-z

Luhn, H. P. (1958). The automatic creation of literature abstracts. *IBM Journal of Research and Development*, *2*(2), 159–165. doi:10.1147/rd.22.0159

Mani, I., & Maybury, M. T. (1999). Advances in Automatic Text Summarization. MIT Press.

Marcu, D. (1998). Improving summarization through rhetorical parsing tuning. In *Proceedings of The Sixth Workshop on Very Large Corpora* (pp. 206-215).

Meena, Y. K., & Gopalani, D. (2014). Analysis of Sentence Scoring Methods for Extractive Automatic Text Summarization. *International Conference on Information and Communication Technology for Competitive Strategies (ICTCS-2014)*. ACM. 10.1145/2677855.2677908

Mendoza, M., Bonilla, S., Noguera, C., Cobos, C., & Leon, E. (2014). Extractive single-document summarization based on genetic operators and guided local search. *Expert Systems with Applications*, *41*(9), 4158–4169. doi:10.1016/j.eswa.2013.12.042

Mihalcea, R., & Tarau, P. (2004). Textrank: Bringing order in to texts. In Proceedinngs of EMNLP 2004. Barcelona, Spain: Association for Computational Linguistics.

Mori, T. (2002). Information gain ratio as term weight: the case of summarization of ir results. In *Proceedings of the 19th International Conference on Computational Linguistics*. Association for Computational Linguistics. 10.3115/1072228.1072246

Nenkova, A., & McKeown, K. (2012). A Survey of Text Summarization Techniques. In C. C. Aggarwal & C. Zhai (Eds.), *Mining Text Data* (pp. 43–76). Springer, US. doi:10.1007/978-1-4614-3223-4_3

Neto, J. L., Freitas, A. A., & Kaestner, C. A. A. (2002). Automatic text summarization using a machine learning approach. In Advances in Artificial Intelligence (LNCS), (vol. 2507, pp. 205-215). Springer Berlin Heidelberg. doi:10.1007/3-540-36127-8_20

Nobata, C., Sekine, S., Isahara, H., & Grishman, R. (2002). Summarization System Integrated with Named Entity Tagging and IE pattern Discovery. *Proceedings of Third International Conference on Language Resources and Evaluation (LREC 2002)*.

Nobata, C., Sekine, S., Murata, M., Uchimoto, K., Utiyama, M., & Isahara, H. (2001). Sentence extraction system assembling multiple evidence. In *Proceedings of the Second NTCIR Workshop Meeting*.

Ono, K., Sumita, K., & Miike, S. (1994). Abstract generation based on rhetorical structure extraction. In *Proceedings of the 15th conference on Computational linguistics* (Vol. 1, pp. 344-348). Kyoto, Japan: Association for Computational Linguistics. 10.3115/991886.991946

Pingali, J. J., & Varma, V. (2005). Sentence extraction based on single document summarization. In *Workshop on Document Summarization*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Pollock, J. J., & Zamora, A. (1975). Automatic Abstracting Research at Chemical Abstracts Service. *Chemical Information and Computer Sciences*, *15*(4), 226–232.

Porselvi, A., & Gunasundari, S. (2013). Survey on web page visual summarization. *International Journal of Emerging Technology and Advanced Engineering*, *3*, 26–32.

Porter, M. F. (1980). An algorithm for suffix stripping. *Program*, *14*(3), 130–137. doi:10.1108/eb046814

Prassad, R., Shardanand, & Kulkarni, U. (2010). Implementation and evaluation of evolutionary connectionist approaches to automated text summarization. Journal of Computer Science, 1366-1376.

Radev, D., Blair-Goldensohn, S., & Zhang, Z. (2001). Experiments in single and multi-document summarization using MEAD. In *First Document Understanding Conference*.

Rafael, F., Freitas, F., De Souza, C. L., Dueire, L. R., Lima, R., Franca, G., … Favaro, L. (2014). A Context Based Text Summarization System. *2014 11th IAPR International Workshop on Document Analysis Systems (DAS)*.

Rush, J. E., Salvador, R., & Zamora, A. (1971). Automatic abstracting and indexing. ii. production of indicative abstracts by application of contextual inference and syntactic coherence criteria. *Journal of the American Society for Information Science*, *22*(4), 260–274. doi:10.1002/asi.4630220405

Salton, G., Fox, E. A., & Wu, H. (1983). Extended Boolean information retrieval. *Communications of the ACM*, *26*(11), 1022–1036. doi:10.1145/182.358466

Salton, G., Singhal, A., Mitra, M., & Buckley, C. (1997, March). Automatic text structuring and summarization. *Information Processing & Management*, *33*(2), 193–207. doi:10.1016/S0306-4573(96)00062-3

Shen, D., Sun, J.-T., Li, H., Yang, Q., & Chen, Z. (2007). Document summarization using conditional random fields. In *Proceedings of the 20th international joint conference on Artifical intelligence* (pp. 2862-2867). Hyderabad, India: Morgan Kaufmann Publishers Inc.

Svore, K., Vanderwende, L., & Burges, C. (2007). Enhancing single-document summarization by combining RankNet and third-party sources. In *Proceedings of the EMNLP-CoNLL* (pp. 448-457).

Torres, M., Juan-Manuel, P. M., & Meunier, J.-G. (2002). Condenses de textes par des methodes numeriques. In *JADT* (Vol. 2, pp. 723–734). JADT.

Wong, K.-F., Wu, M., & Li, W. (2008). Extractive summarization using supervised and semi-supervised learning. In *Proceedings of the 22nd International Conference on Computational Linguistics* (Vol. 1, pp. 985-992). Manchester, UK: Association for Computational Linguistics. 10.3115/1599081.1599205

Xiaoyan, C., Wenjie, L., You, O., & Hong, Y. (2010). Simultaneous ranking and clustering of sentences: a reinforcement approach to multi-document summarization. In *Proceedings of the 23rd International Conference on Computational Linguistics* (COLING '10). Association for Computational Linguistics.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

Abstractive Summarization: Extraction of sentences by merging and fusing from the original text.

Extractive Summarization: Retrieval of sentences as its is from the original text.

Precision: Ratio of the relevant sentences retrieved divided by total sentences retrieved.

Recall: Ratio of the relevant sentences retrieved divided by total sentences relevant sentences.

Sentence Features: Any of various abstract entities that specify or combine to specify phonological, morphological, semantic, and syntactic properties of linguistic forms and that act as the targets of linguistic rules and operations.

Stemming: The process of reducing the words to its root form or word stem.

Stopword: Commonly or frequently used words (such as "is") that needs to be ignored at most of the times of processing of text either searching or retrieving.

Text Summarization: A process of condensing the text document automatically by computer system such that the produced summary is no longer than a specified threshold [typically one-third of the size of original text] and covers all the information contents of it.

This research was previously published in Enterprise Big Data Engineering, Analytics, and Management edited by Martin Atzmueller, Samia Oussena, and Thomas Roth-Berghofer; pages 126-144; copyright year 2016 by Business Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 31
Improving Collaborative Filtering Algorithms: Sentiment-based Approach in Social Network

Firas Ben Kharrat
University of Carthage, Tunisia

Aymen Elkhleifi
Paris-Sorbonne University, France

Rim Faiz
University of Carthage, Tunisia

ABSTRACT

This paper puts forward a new recommendation algorithm based on semantic analysis as well as new measurements. Like Facebook, Social network is considered as one of the most well-prominent Web 2.0 applications and relevant services elaborating into functional ways for sharing opinions. Thereupon, social network web sites have since become valuable data sources for opinion mining. This paper proposes to introduce an external resource a sentiment from comments posted by users in order to anticipate recommendation and also to lessen the cold-start problem. The originality of the suggested approach means that posts are not merely characterized by an opinion score, but receive an opinion grade notion in the post instead. In general, the authors' approach has been implemented with Java and Lenskit framework. The study resulted in two real data sets, namely MovieLens and TripAdvisor, in which the authors have shown positive results. They compared their algorithm to SVD and Slope One algorithms. They have fulfilled an amelioration of 10% in precision and recall along with an improvement of 12% in RMSE and nDCG.

DOI: 10.4018/978-1-7998-0951-7.ch031

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Nowadays, the continuous progress of data and communications technologies, the increasing use of digital technologies among the people and the development of the users' lifestyles and norms necessitate innovative types of systems. The success of these systems will rely upon their capability to operate as a unit, to supply services to their users in every place and at any moment as well as to accustom themselves to the changes surrounding their environments. Recommender systems help users to identify their interests and sets of choices by predicting the invalidity of an item or a group of items to these users. They are defined as a special type of information filtering that provide information about which items may be interesting to users. Recommender systems generate a personalized recommendation for users based on a set of previously rating items. However, according to (Baldoni et al., 2011) the social web principles, social network and platforms promote the participation of users in many ways, stimulating the expression of opinions about the contents inserted by other users, by 'Like', 'Share' annotations, star-rating systems... This huge amount of data is a precious information source about perceptions, trends and feelings, and a lot of research is being carried on to identify ways for extracting meaningful information from these data. In fact, the social network could be considered an important source of recommendation. This paper endeavors to combine two types of information respectively called social and semantic information. The main purpose is to compute recommendation lists through making a visit to a limited part of social network.

We have improved the Slope One and SVD algorithms by external knowledge and through some experiment results. The approach has been applied in Java and tested on the MovieLens and TripAdvisor data sets. The results demonstrate that the additional knowledge achieved by the semantic analysis contributes to the amelioration of the quality recommendation in terms of accuracy.

Hence, this study foregrounds two major issues. The first part seeks to gather as well as parse data, while the second one attempts to use external resources to come up with relevant suggestions.

We construct this paper as following: We start with some major research literature on recommendation systems. Next, we highlight our approach via applying user's sentiment in order to raise the value of recommendation algorithm. We clarify the experiment immediately followed by an estimation of our algorithm. In conclusion, we briefly summarize our proposed study providing a few remarks on future works.

EASE OF USE

According to Burke (Burke, 2002), Recommender systems have three fundamental categories: collaborative-filtering, content-based and hybrid. The hybrid recommender system combines the content-based and the collaborative-filtering regarding the kind of advocated data Recommendation types are explained in details with their limitations in this section:

- **Content-based:** It advocates items that are akin to those the users selected on time before (Adomavicius & Tuzhilin, 2005; Burke, 2002). Items like products, services or even people are identified by their related characteristics. User's favorite choices, which are preserved in the user's profile and taking into account those associated criteria, emerge in already estimated items. According to (Schafer et al., 1999) content-based technique is termed item-to–item affinity. We show some usual research on which these traits are revolved around this technique, which are:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- NewsWeeder (Lang, 1995) is newsgroup, an internet base-discussion, filtering system. It charges users to rank some not viewed news articles according to their preferences. The applications were built on schematic tree-shaped diagram, sensory network and graphics based on vector.
- Pazzani et al., 1997 suggests a system able to bring forth web sites based on issues that the users have to be involved in regarding to their typical interest. In this experimentation, the applied system is Bayesian classifier.
- Zhang et al., 2002 suggests to differentiate in the midst of queries appropriate files including new data and using Bayesian system.
- **Collaborative Filtering:** CF is also termed people-to-people correlation.
- Tapestry (Goldberg et al., 1992) extracts electronic script on a drafted issue by a particular user.
- Ringo (Shardanand & Maes, 1995) advocated recording of music and artists established by resemblances among user's tastes as well as others.
- Hoaks (Hill & Terveen, 1996; Terveen et al., 1997) proposes home pages from news messages. If followers are interested, the system will not only find the person who got in a touch but also identify the web page.
- GroupLens (Konstan et al., 1997; Resnick et al., 1994) makes an advance in a system that deals with extracting data that assesses newsgroups article.
- Jester (Goldberg et al., 2001) is an online funny recommending the system. It uses a Collaborative Filtering algorithm termed Eigentaste and it is adjacent to algorithm for the disconnected phase. Next, the author utilizes a repetitive rectangular classification method for the offline stage.
- **Knowledge-based:** Advocates items taken from the assumption or the opinion of the user's tastes and choices. Indeed, the user profile comprises useful conception organized and explained in harmony with the inference machine:
- Google (Brin & Page, 1998) proposes the most famous links including the questions given by the users. The system is characterized by probabilistic standard.
- CF is the widely successful recommendation approach and has become used in many e-commerce systems. However, collaborative-based filtering undergoes various drawbacks:
- **Data Sparsity:** Most users rate small numbers of items in the system (Lianga et al., 2006). Therefore, those users may not find similar users.
- **Cold start Problem:** This is referring to the problems occurring with a new item or a new user. On the one hand, a new user who has not rated any items yet will not find similar users to help in finding recommendations. On the other hand, when a new item is introduced into the system and no user has rated it yet, it will be impossible to recommend that item anyway.

In recent works, semantic web mining is used in order to improve recommendations (Blanco et al., 2008; Blanco et al., 2010; Moreno et al., 2010; Kim et al., 2011). This approach is a field of research where it converges the semantic web and the web mining, which are two areas that have been evolving very quickly. Both techniques are contributing to the latest challenges of the WWW in a complementary way. The shortage of structure of most of the web data can be just interpreted by human beings. Whereas, the huge amount of information can only be processed efficiently by machines. The semantic web addresses the first problem by enriching the web with machine-understandable information, while web mining addresses the second one by automatically extracting the useful knowledge hidden in web data (Stumme et al., 2006). Semantic web mining aims at improving the results of recommendation by exploiting the semantic structures in the web as well as building the semantic web making use of web mining techniques (Schroder et al., 2011).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In a social network context, the identification of the opinion of users is receiving an increasing attention. It is a striking indicator of the appraisal of topics, people, situations, re-sources and trends. Hence, the development of opinion Mining (Pang & Lee, 2008), of ontologies of emotions and of W3C markup language proposals like Emotion Markup Language (Schroder et al., 2010; Schroder et al., 2011). However, there are few applications that still use the most advanced results in web technology to deal with opinion. Moreover, most of the approaches use emotional ontologies where emotions are individual isolated units (e.g. WordAffect). Such considerations motivated our focus on the ontology of emotions in (Francisco et al., 2007), an OWL ontology where emotions are structured and organized in levels, trying to integrate the results of the most recent psychological models.

Furthermore, on account of the evolution of social network application, recommendation systems in the social network have become more common. Facebook group recommender is proposed in order to make each user join the group that is near to its interests (Baatarjav et al., 2008). The authors proposed a mechanism called Group Recommendation System 'GRS' which uses hierarchical clustering and decision tree techniques. These techniques classify the users according to their features and then provide the recommendation depending on this classification.

Slope One One (Lemire & Maclachlan, 2005) is a CF algorithm. Firstly, it is not sophisticated to put into function or to test it indeed. Secondly, it is updatable in the sense that new estimation may alter expectations immediately. Thirdly, it is useful when we want to question visitors. It is necessary that those queries must be swift to the benefit of storage. Ultimately, the Slope One algorithm expects a little assessment from the first visitors in order to produce authentic recommendations.

Slope One utilizes a measure of similar items akin to Item-based CF. This standard is quiet distinct from the variation of Item-base Collaborative Filtering just as the prediction calculation. In another word, we intend to keep the differential rating values of items in pair.

The following symbols or signs were used to describe schemes adopted by (Lemire & Maclachlan, 2005):

- u_i is the rating of a user given to item i.
- S(u) is the subset of items consisting of all items rated in u.
- Card(S) is the number of elements in S.
- The average of ratings in an evaluation u is denoted $\overline{u}$.
- The set Si(X) is the set of all evaluations u ∈ X as they contain item i (i ∈ S(u)).
- Given two evaluations u, v, we define the scalar product (*u,v*) as $\sum_{i \in S(u) \Delta S(v)} u_i v_i$.

Predictions, P (u), represent a vector where each component is the prediction corresponding to one item: predictions depend implicitly on the training set X.

The similarity in Slope One is defined as the average deviation of items i, j given by:

$$dev(i,j) = \sum_{u \in S_{i,j}(X)} \frac{u_i - u_j}{card\left(S_{i,j}(X)\right)} \tag{1}$$

When new ratings are entered, *dev*(*i,j*) is symmetric and computation time is halved. $dev(i,j) = u_j$ used as a prediction for u_i. Therefore, a usable predictor is the average of all predictions:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$P^{S1schema}\left(a\right)_i = \frac{1}{card\left(R_j\right)}\sum_{j\in R_j}\left(dev\left(i,j\right)+u_j\right) \tag{2}$$

With

$$R_i = \left\{j|j \in S\left(u\right), i \neq j, card\left(S_{i,j}\left(X\right)\right) > 0\right\}.$$

The authors (Lemire & Maclachlan, 2005) present an approximation based on a dense dataset, where almost all pairs of items i and j were rated, R_i=$S(u)$/(i). Also, in a dense dataset $\sum_{j\,/\,R_i}\frac{u_j}{card\left(R_j\right)}$ is approximately $\overline{u}$. The accurate prediction is:

$$P^{S1}\left(u\right)_j = \overline{u} + \frac{1}{card\left(R_j\right)}\sum_{i\in R_j} dev_{i,j} \tag{3}$$

According to (Koren et al., 2009), Matrix factorization models map both users and items to a joint latent factor space of dimensionality *f*, such that user-item interactions are modeled as inner products in that space. Accordingly, each item *i* is associated with a vector *qi* ∈ R, and each user u is associated with a vector *pu* ∈ R. For a given item *i*, the elements of *qi* measure the extent to which the item possesses those factors, whether positive or negative. For a given user *u*, the elements of *pu* measure the extent of interest the user has in items that are high on the corresponding factors, once again, positive or negative. The resulting dot product, $q_i^T p_u$, captures the interaction between user *u* and item *i.* The user's overall interest in the characteristics of the item. This approximates user u's rating of item *i*, which is denoted by $\hat{r}_{ui}$, leading to the estimate:

$$\hat{r}_{ui} = q_i^T p_u \tag{4}$$

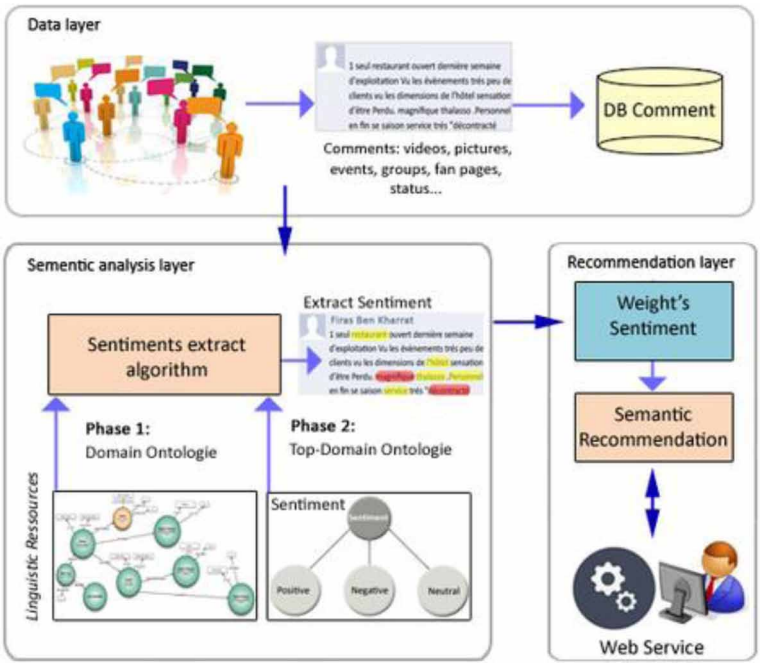
OUR APPROACH

General Architecture

Figure 1. displays the overall architecture of our proposed approach in this current paper. It is based on semantic analysis of user's comment. Our architecture connects the users via the collaborative social network in order to use sentiments in their comments. This system combines the social network analysis comments with the sentiment of item's concepts in one social recommendation algorithm. The suggested system consists of three parts. The first part deals with a comment posted by the user. This comment is stored in the database of comments. We utilize these data to create a user profile. The second part is established on semantic analysis methods that deploy semantic-based techniques. This consists of exploring the semantics of different sentiments, for each user's comment. Next, we calculate user's weight from the extracted sentiment. Finally, the recommendation layer, where the new input is sent to our algorithms.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Collaborative recommender systems architecture



Data

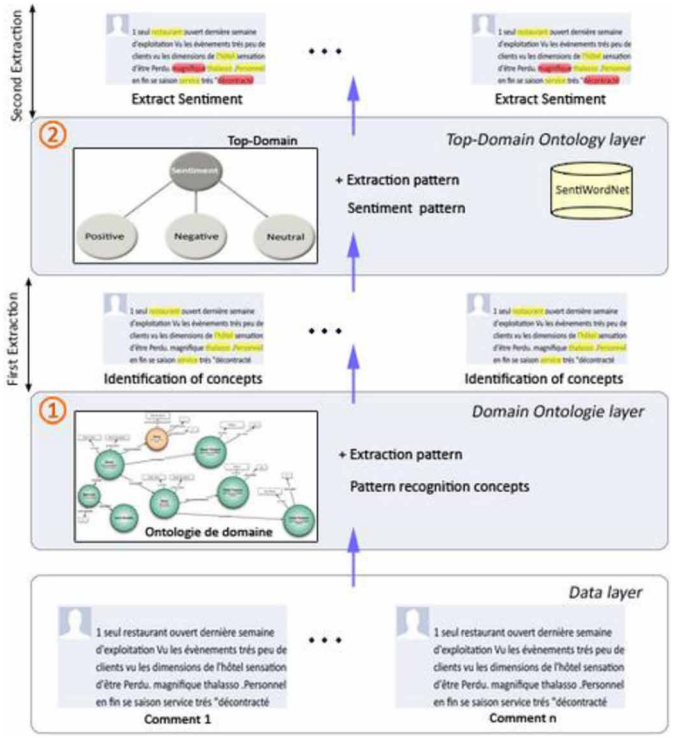
The MovieLens dataset is one of the gold-standards for testing collaborative filtering algorithms (Ekstrand et al., 2011). It contains 100,000 real ratings (rє{1,2,3,4,5}) from 943 users on 1,682 movies, which are collected through the MovieLens web site. For purposes of validation, the dataset is organized in five different splits, each containing a training set with 80,000 ratings and a test set with 20,000 ratings.

In addition, we chose to experiment our system in the tourism domain. We created one benchmark data composed by description and comments of many hotels. We developed an extract program based on 'jaunt API.' It's an open source API to draw forth data such as hotel's description, comments, localization, and so on and so forth from tourism website. TripAdvisor is a well-known online website that collects comments and reviews. It is also rated by customers who book hotels throughout the world.

The primary problem that we encountered was that of collecting information from social network. We chose the social network Facebook for our experiment. Then, we set up a script in java in order to gather all necessary comments using 'Facebook Graph API', which included comments related to the topic of movies. We stimulated all MovieLens' users with our profile collection by a simple matching. This matching is based on demographic information about the users like age, gender, occupation and country. For privaetness, both data sparsity and dynamicity were being taken into consideration when tackling with social network data. Furthermore, we have created an online website to collect users profile from Facebook. In the tourism domain, the users of TripAdvisor and Facebook were alike. Ultimately, all users profile was stored in the database.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Semantic analysis layer



Semantic Analysis

This semantic analysis contributes to a system that allows a comment with reference to a particular item as well as offers sentiment average to each concept in this item. The semantic analysis of the developed system is well elucidated in Figure 2. The suggested approach is divided into two stages. At first, (a) is the application of the domain ontology to filter concepts. At second, (b) is the sentiment analysis on a collection of comments characterized by the concepts and properties and also included in the Top-Domain ontology. These two stages are well depicted afterward.

- **Domain Ontology:** The most important contribution of ontologies is the key role that they have played in extracting concepts from comments. We extended the existing ontologies so as to use a domain of ontologies. Figure 2 and 3 demonstrate a set of concepts which are organized in a hierarchical structure (depending on the domain) to describe objects of items. Our produced system uses an ontology to represent items for description. As we show in Figure 3, the concept `Person' at level 1 is the most general concept in films' ontology, whereas the concept in deeper levels is more specific one.
- **Top-Domain Ontology:** The primary idea behind the proposed approach is to take advantage of a Top-Domain ontology for providing more elaborate sentiment scores regarding the notions contained in a user's comment. In order to rise the underlying semantics, the Top-Domain ontology is enriched with synonyms and hyponyms (SentiWordNet) of the detected attributes, Figure 5.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Films ontology

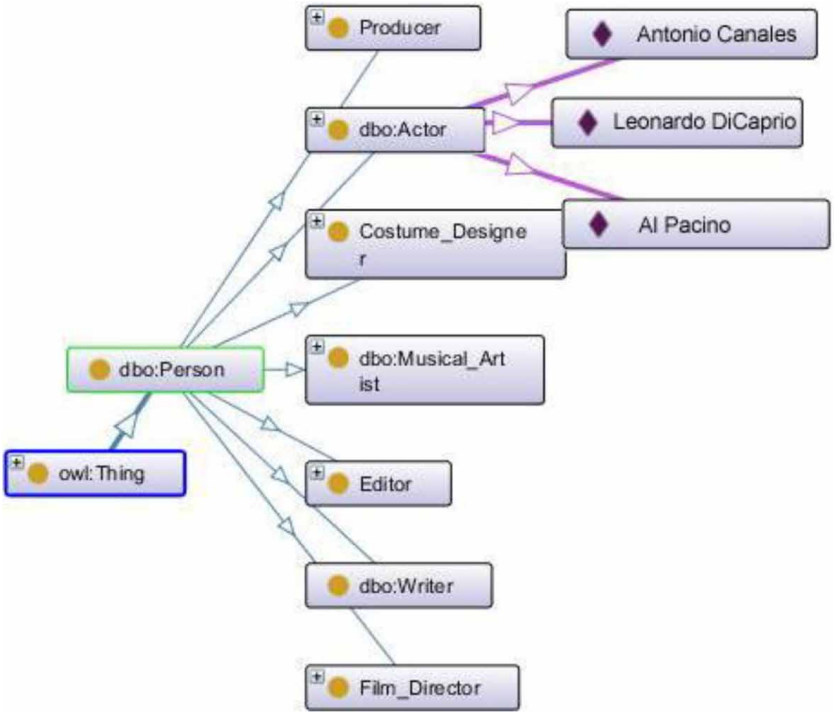
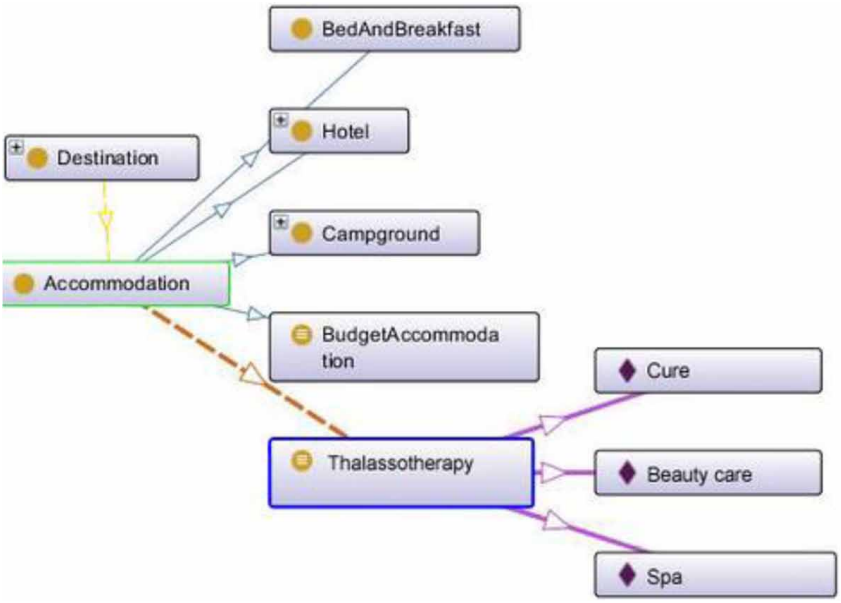


Figure 4. Tourism ontology



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Similarly, SentientiWordNet is defined as, a lexical resource for opinion mining. It contains different subsets of effective concepts that group together synsets which denote sentiment states. SentiWordNet assigns to each synset of WordNet three sentiment scores: positivity, negativity, objectivity.

- **Extraction Algorithm of Sentiment:** used to annotate all sentiments in comments. We have two phases of annotation. In the first phase, we annotated concepts. Then, in the second one, we elucidated sentiments on these concepts. Sometimes we found more than two sentiments in the same sentence. In this case, we have defined a priority for the selection of sentiment as follows: Negative, Neutral and positive. Figure 1 described the algorithm that annotated sentiments in our textual corpus, based on both the Domain ontology and Top-Domain ontology.

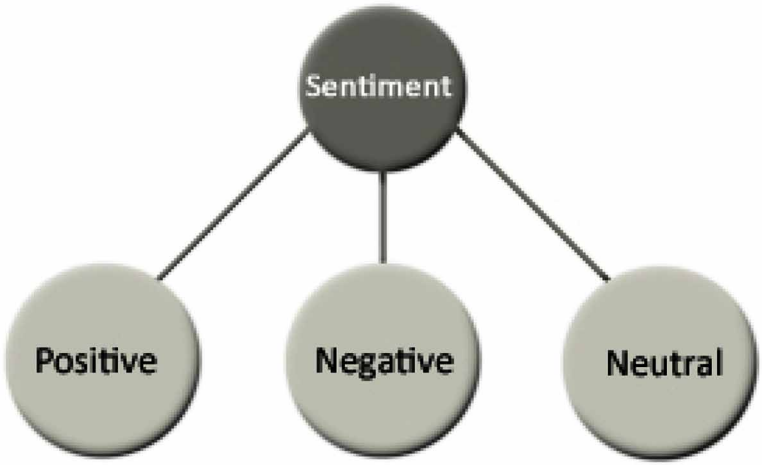
The users' semantic data is provided by the user's sentiment for distinct concepts of items. In the system, every user has a profile expressed by the sentiment of many concepts, demonstrated in Figure 6. The opinion embodies the user interests applying a comment's semantic analysis. Correspondingly, concepts of every item have a range of preferences regarding the significance of the concept. For instance, an actor in the film is more important than the sound production company. It is important to mention that the user may have different opinions on the same concept in different items.

Recommendation Algorithm

As already explained, the basic idea beyond the proposed approach is to take advantage of a semantic analysis for providing more elaborate semantic scores regarding concepts contained in a comment. We combine the social network analysis measures with the sentiment of the users in one semantic recommendation algorithm.

We divide the step of recommendation into two phases: (a) is the creation of sentiment score (Table 1), and (b) is the integration of sentiment dimension onto recommendation algorithms. These two phases are more described hereunder.

Figure 5. Semantic map of Top-Domain Ontology



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Algorithm 1. Definition

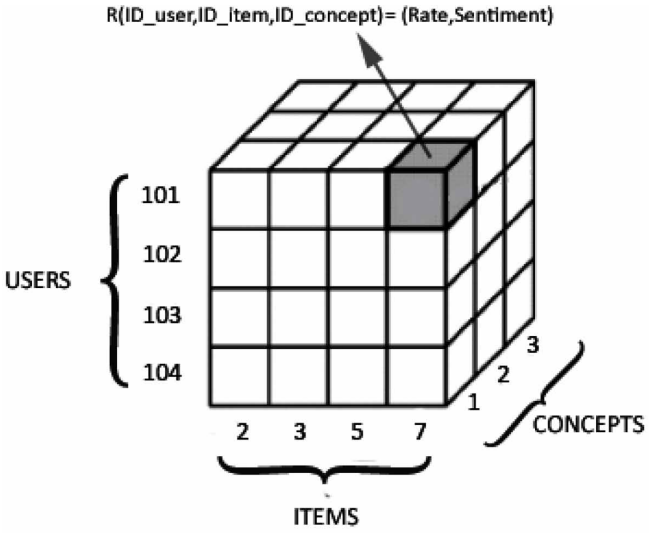
1 : $\boldsymbol{CSt}$: Ontology
2 : $\boldsymbol{R}$ a set of defined rules, based on categories $\boldsymbol{CSt}$
3 : with $\boldsymbol{Ri} = \{\boldsymbol{Ii}, \boldsymbol{Cgi}, \boldsymbol{Cdi}, \boldsymbol{C'gi}, \boldsymbol{C'di}\}$ an extraction rule
4 : $\boldsymbol{Ii}$ a set of indicator
5 : $\boldsymbol{Cgi}$ and $\boldsymbol{Cdi}$ are respectively left and right indice.
6 : $\boldsymbol{C}$ is a Corpus formed of a set of comment $\boldsymbol{D}$
7 : $\boldsymbol{S}$ is the set of segments that can be formed from $\boldsymbol{D}$
8 : $S = \{S_i, \bigcup S_i = D\}$with $\boldsymbol{Si}$ = {phrase, title, paragraph, article}
9 : $T = \{T_i, \bigcup T_i = S\}$with $\boldsymbol{Ti}$ is a $\boldsymbol{Si}$ segment term
10 : $X_P = \{\bigcup X_{Pi} = P\}\boldsymbol{P}$ is the set of terms in positive concept
11 : $X_N = \{\bigcup X_{Ni} = N\}\boldsymbol{N}$ is the set of terms of negative concept
12 : $X_{Ne} = \{\bigcup X_{Nei} = Ne\}\boldsymbol{Ne}$ is the set of terms of neutral concept

Algorithm 2. AnnotS

1 : ***AnnotS***(Ontology $\boldsymbol{CSt}$, Rules $\boldsymbol{R}$, Corpus $\boldsymbol{C}$) **{**
2 : ***for each*** $S_i \in S$ ***do*** **{**
3 : ***if*** $(\exists T_i \in I)$ ***then***
4 : $F = \{R_i, T_i = I_i\}$all Rules with $\boldsymbol{Ti}$ indicator
5 : $\boldsymbol{CL}$, $\boldsymbol{CR}$ the left and right context segment $\boldsymbol{Si}$
6 : ***for each*** $R_i \in F$ ***do*** **{**
7 : $left = false$; $right = false$; $annotation = false$;
8 : ***if*** $\neg((\exists T_j \in C_L) \vee (\exists T_j \in C_R))$ ***then***
9 : ¦ ***if*** $(\exists C_{di} \wedge (\exists T_j \in C_{di}) \wedge (\exists T_j \in C_R))$***then***
10 : ¦ $left = true$ ***end if***
11 : ¦ ***if*** $(\exists C_{gi} \wedge (\exists T_j \in C_{gi}) \wedge (\exists T_j \in C_L))$***then***
12 : ¦ $right = true$ ***end if***
13 : ¦ ***// Concepts extraction***
14 : ¦ ***if*** $(\exists C_{di} \wedge \exists C_{gi} \wedge right \wedge left)$***then***
15 : ¦ ¦ $Annotate(S_i, R_i)$ and $annotation = true$
16 : ¦ ¦ ***else if*** $(\exists C_{di} \wedge right)$***then***
17 : ¦ ¦ ¦ $Annotate(S_i, R_i)$and $annotation = true$
18 : ¦ ¦ ¦ ***else if*** $(\exists C_{gi} \wedge left)$***then***
19 : ¦ ¦ ¦ $Annotate(S_i, R_i)$ and $annotation = true$ ***end if***
20 : ¦ ¦ ¦***end if***
21 : ¦ ¦***end if***
22 : ¦ ***// Sentiments extraction***
23 : ¦ ***if*** $(annotation)$***then***
24 : ¦ ¦ ***if*** $(\exists T_j \in X_N)$***then*** $Annotate2(S_i, I_i, T_j)$
25 : ¦ ¦ ***else if*** $(\exists T_j \in X_{Ne})$***then*** $Annotate2(S_i, I_i, T_j)$
26 : ¦ ¦ ***else if*** $(\exists T_j \in X_P)$***then*** $Annotate2(S_i, I_i, T_j)$
27 : ¦ ¦ ***end if end if***
28 : ¦ ***end if end if***
29 : ***end if end for***}
30 : ***end if end for***}}

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. User's profile



For every concept with sentiment, a relevant score is submitted to this concept via table score, described in Table 1. We are also interested in calculating the score for other users of the system. Therefore, we assign a specific score to their concepts that carry sentiment information.

The sentiment of users is given by user profiles, Figure 6. Each user in the system has a profile described by means of concept. The concept represents the item description using the domain Ontology.

Figure 7. shows an example of item's concept and user's sentiment. User 1 has a positive sentiment about the actor 'Leonardo Dicaprio' in the film 'Titanic'. We want to anticipate another film to user 1 with the same actor. We calculated the concept's sentiment to check if we can recommend him the film 'Aviator'.

Table 2. displays two sample pre-processed comments retrieved during this phase as well as their corresponding sentiment score generated from the average of weight.

The formalization of weight for the user *i* and item *j* is as follows:

$$w1_{i,j} = \frac{\sum_{k=1}^{n} \left(\alpha_k \times g_k\right)}{n} \tag{5}$$

Table 1. Sentiment scores

Predict User		Current Users	
Sentiment	**Weight**	**Sentiment**	**Weight**
Positive	2	Positive	1
Neutral	0	Neutral	-1
Negative	-1	Negative	-2

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. An example of concepts of sentiment (Films' domain)

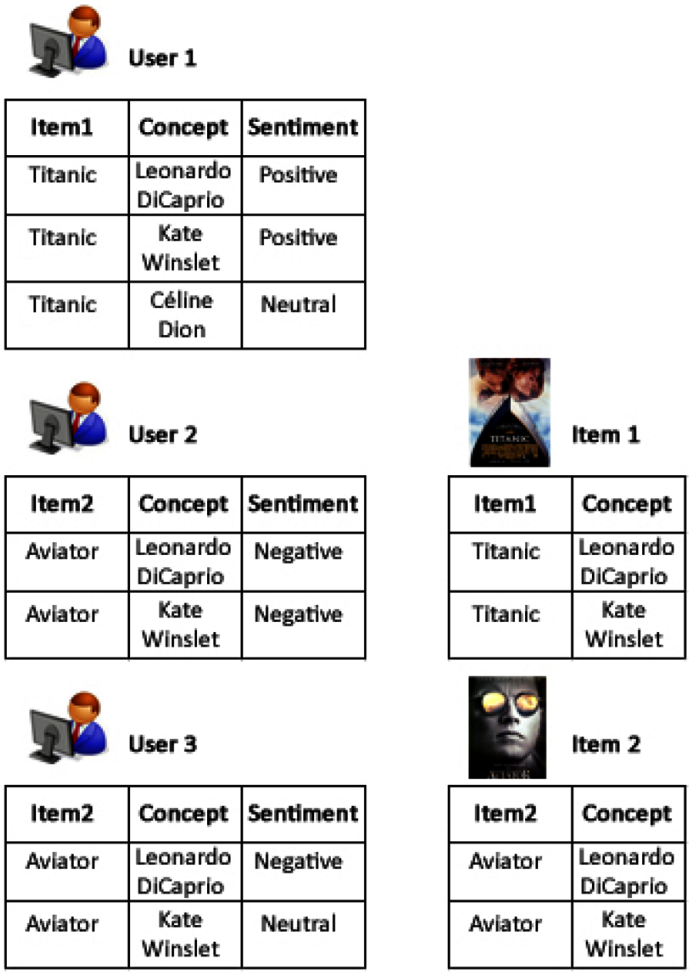


Table 2. An example to calculate the weight of items and concepts

ITEM j

	User1	...	UserN	
Concept1	Positive (+1)		Neutral (-1)	Average of weight
⋮				
Concept K	Negative (-2)		Negative (-2)	

USER i

	Item1	...	ItemM
Concept1	Neutral (-1)		Negative (-2)
⋮			
Concept L	Positive (+1)		Negative (-2)
	Average of weight		

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

With: α_k is the weight of the concept's sentiment *k*. *n* is the number of concept. And g_k is the grade of the concept *k;* $g_k \in]0,1[$ depends on the importance of the concept.

The formalization of weight for system users and item j is:

$$w2_j = \frac{\sum_{k=1}^{n} (\beta_k \times g_k)}{n} \tag{6}$$

With: β_k is the weight of the concept's sentiment *k*, returned from all sentiments of current users.

The proposed recommendation algorithm is based on user's sentiment and the concepts of item.

The new algorithms are defined by an improved Slope One algorithm, described in Equation (7) and an improved SVD algorithm, described in Equation (9).

$$P^{SS1}\left(u\right)_j = \frac{\sum_{i \in R_j} (dev_{i,j} + u_i)}{card\left(R_j\right)} + \log\left(1 + \frac{w1_{u,j} \times w2_j}{1 + w1_{u,j} + w2_j}\right) \tag{7}$$

With:

$$dev\left(i,j\right) = \sum_{u \in S_{i,j}(X)} \frac{(u_i - u_j) \times \left(1 + w1_{u_i,j}\right)}{card\left(S_{i,j}\left(X\right)\right)} \tag{8}$$

The matrix factorization model should integrate all signal sources with enhanced user representation. The idea is to minimize the squared error over all ignored rate:

$$\hat{r}_{ui} = \left[q_i^T + \log\left(\frac{w2_j}{w1_{u,i} + w2_j}\right)\right]\left[p_u + \log\left(\frac{w1_{u,i}}{w1_{u,i} + w2_j}\right)\right] \tag{9}$$

EXPERIMENTATION AND RESULTS

We used the MovieLens dataset with 1k ratings by 943 users on approximately 1682 items. Each user rated at least 20 movies. Then, we used the TripAdvisor dataset with 500 ratings, in which only users who have rated at least 20 hotels are included in our database. We also used the LensKit recommender toolkit for our experiments (Ekstrand et al., 2011). LensKit evaluations are defined with Groovy scripts. Evaluations are organized around tasks such as crossfold (to partition data for cross-validation) and train Test (to run a train-test evaluation over one or more datasets). Tasks can optionally be contained within targets. In this case, their execution is deferred until the target is executed allowing a single evaluation script to define multiple evaluation capabilities.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

We used the following algorithms by identifying parameters revolved around prior results in the literature research and experience tuning LensKit for both the MovieLens and TripAdvisor datasets. Therefore, we will compare Slope One (Lemire & Maclachlan, 2005) algorithm with our improved one, 'SemSlope One,' against Simon Funk's SVD (Funk, 2006) with our enhanced one, 'SemSVD'.

The effectiveness of a given CF Systems established by evaluative Slope One or SVD algorithms can be measured precisely. As an attempt to assess our given approach, we have utilized four standards metrics (Dixit et al., 2013) such as RMSE, nDCG, precision and recall measurements (Fellbaum et al., 1998).

RMSE and nDCG estimate the prediction accuracy of every advocated item. We have observed that these two standards act quite alike in our assessment. In order to create training and tests collections for distinct number of ratings from 1 to 20, we first made set of experimentation and training partitions. Thereby, test users have 20 training items when applying the crossfold method. Next, to produce any other number of ratings, we haphazardly selected a model of a training set of 20 items. This paved us the way to create training collections with a given number of ratings for test users. For example, each profile size is assessed with the same test set.

RMSE is the square root of the mean/average of the square of all of the error. The use of RMSE is very common and it makes an excellent general purpose error metric for numerical predictions. Compared to the similar Mean Absolute Error, RMSE amplifies and severely punishes large errors.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{10}$$

Normalized discounted cumulative gain (NDCG) measures the performance of a recommendation system based on the graded relevance of the recommended entities. It varies from 0.0 to 1.0, with 1.0 representing the ideal ranking of the entities. This metric is commonly used in information retrieval and to evaluate the performance of web search engines.

$$DCG_k = \sum_{i=1}^{k}\frac{2^{rel_i} - 1}{log_2\left(i+1\right)} \tag{11}$$

With k is the maximum number of entities that can be recommended. $IDCG_k$ is the maximum possible (ideal) DCG for a given set of queries.

$$nDCG_k = \frac{DCG_k}{IDCG_k} \tag{12}$$

DISCUSSION

This study introduces our algorithm, which is termed SemSlope One. We attempted to examine the anticipation errors of our algorithm. Using the LensKit toolkit (Ekstrand et al., 2011), we divided 'SemSVD' in comparison with the original algorithms 'Slope One' and 'SVD'. We begin the task of separating the collection of data in 5 folds cross-validation configuration. Users are divided into 5 groups. In a random

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Results of RMSE, nDCG metrics for algorithms: SemSVD and SVD (MovieLens)

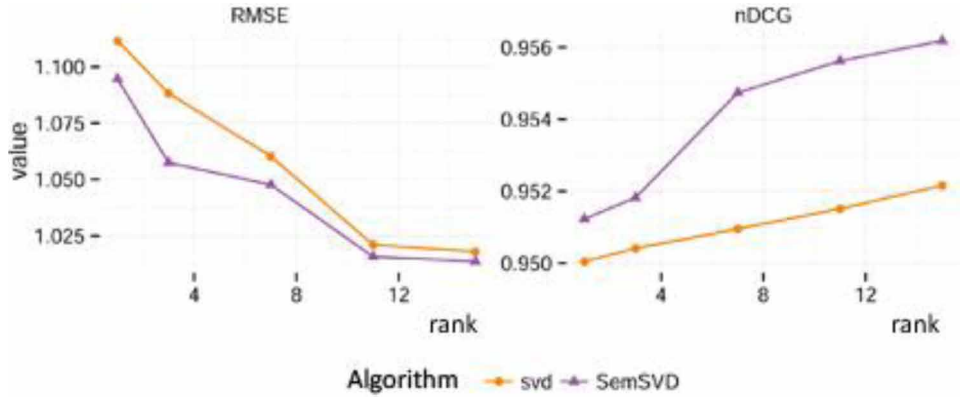


Figure 9. Results of Precision, Recall metrics for algorithms: SemSVD and SVD (MovieLens)

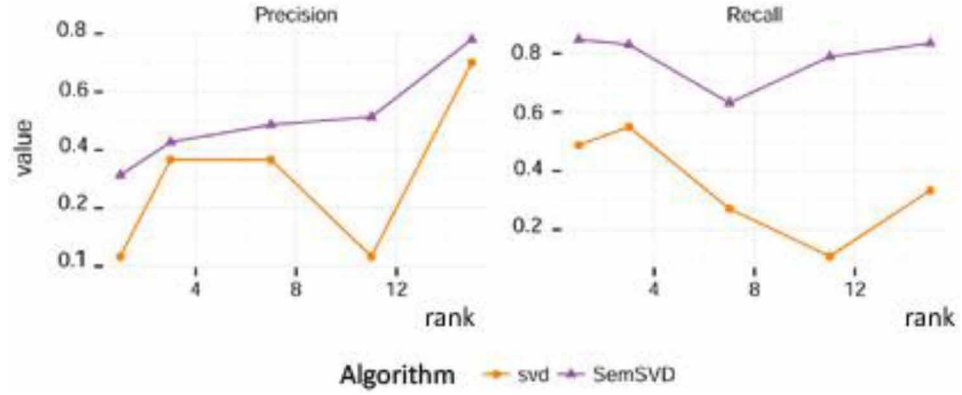
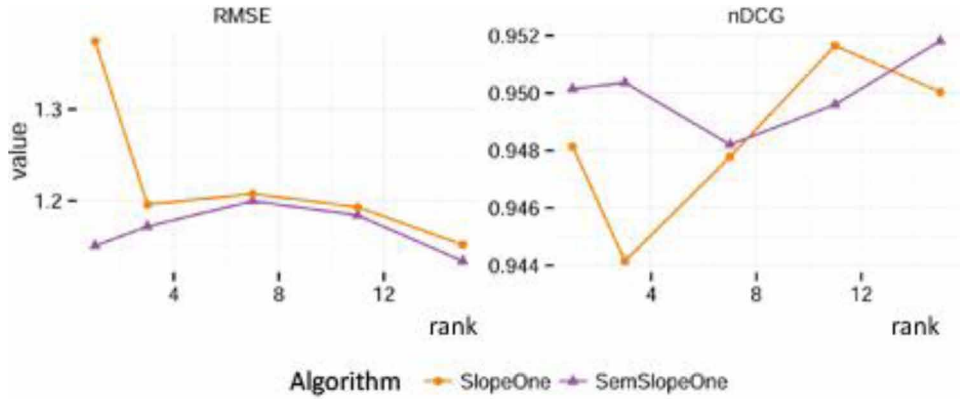


Figure 10. Results of RMSE, nDCG metrics for algorithms: SemSlope One and Slope One (MovieLens)



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Results of Precision, Recall metrics for algorithms: SemSlope One and Slope One (MovieLens)

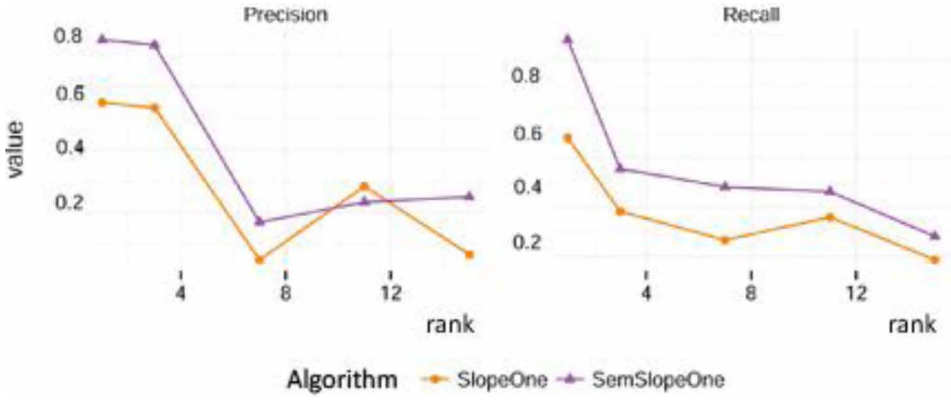


Figure 12. Results of RMSE, nDCG metrics for algorithms: SemSVD and SVD (TripAdvisor)

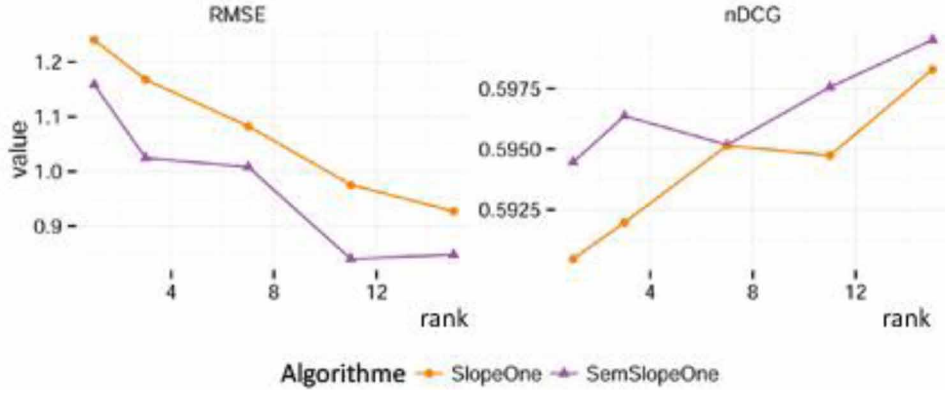
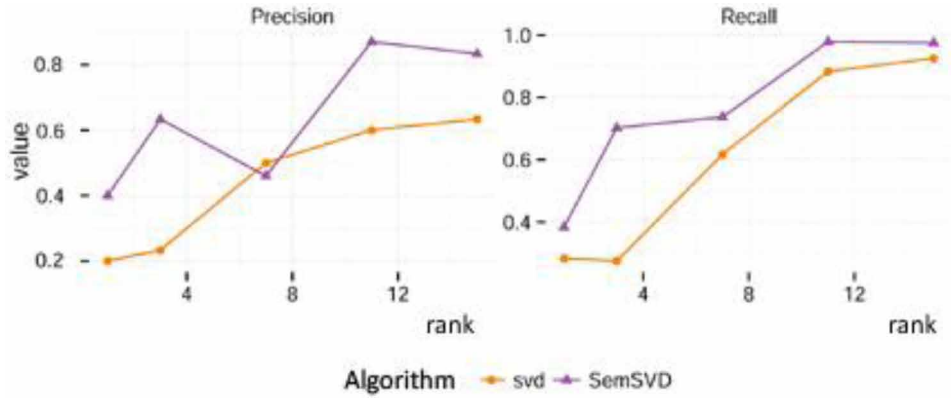


Figure 13. Results of Precision, Recall metrics for algorithms: SemSVD and SVD (TripAdvisor)



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 14. Results of RMSE, nDCG metrics for algorithms: SemSlope One and Slope One (TripAdvisor)

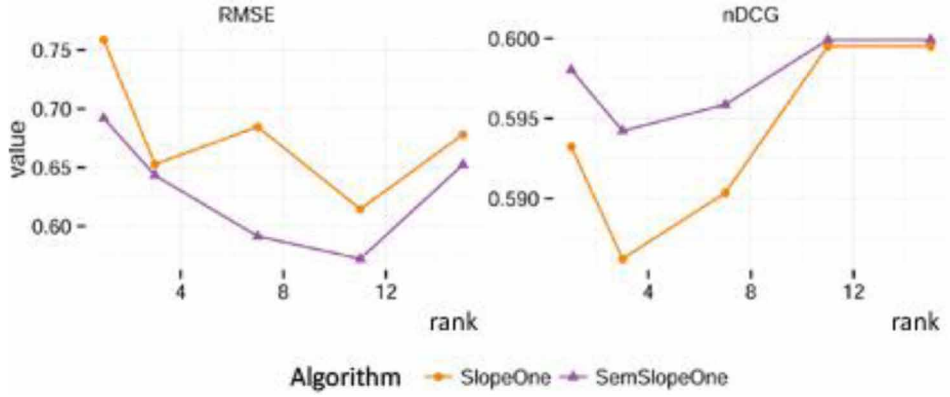
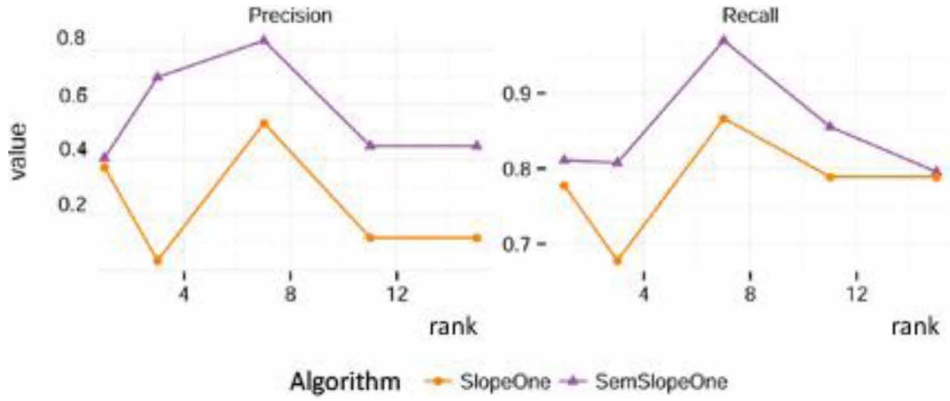


Figure 15. Results of Precision, Recall metrics for algorithms: SemSlope One and Slope One (TripAdvisor)



manner, we picked out 20 of their ranges according to the test to each user in each division. The residual ratings from users combine to create training set. Next, we put into function five recommender algorithms on the data, and recorded the predictions of each algorithm used to each test estimation rating. Thereupon, we compared Slope One algorithm with our enhanced Slope One. But, the task of comparing 'SemSVD' based on matrix factorization with SVD, is rather important. The high-level results from our evaluation are summarized in Figures 8, 9, 10 and 11 using MovieLens dataset and in 12, 13, 14 and 15 applying TripAdvisor dataset. When the number of rates is between 0 and 4, all our improved algorithms perform similarly well. To note, when the number of rating increases, the 'SemSVD' recommender algorithm is the optimal predictor, and it is better than the others. To thoroughly appreciate the importance of our improvement over existing algorithms, it is worth noting that the average progress of our algorithm over the Slope One recommender is 10% and that over SVD is 12%. These are the major progress mainly when compared to the differences that could be attributed to the implementation details, which are shown by the variation in performance between Slope One, SVD and our algorithms.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION AND FUTURE WORK

What we propose in this paper is a fruitful study of a new approach for the recommendation systems based on users' sentiments. This suggested approach is divided into three main stages with an obvious intention to recommend the items.

In the first phase, we selected the comments of the users from the social network using a program called preprocessor. In the second one, we detected sentiment from users' comments. Then, we deduced both user's and item's weight using sentiment from certain selected concepts. The study concludes that the recommendation system is predicted with two algorithms based on Slope One and SVD. The achieved results are promising compared to those reached by the Slope One and SVD algorithms.

The future of algorithm is, indeed, very promising. In the future, we propose to elaborate user's profile using a demographic information. We also aim to extend our study of algorithm to search various connected communities. Besides, we will try to extract data from other resources such as discussion forms, blogs and semantic information in order to improve our assumption about the users in various social networks. Furthermore, we will attempt to extract semantic information from IMDb database to enrich knowledge about items for movies dataset.

REFERENCES

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems. In *IEEE Transactions on Knowledge and Data Engineering* (pp. 26–41).

Baatarjav, E.-A., Phithakkitnukoon, S., & Dantu, R. (2008). Group recommendation system for Facebook. In *On the Move to Meaningful Internet Systems: OTM 2008 Workshops* (pp. 211–219).

Baldoni M., Baroglio, C., Patti, V., & Schifanella, C. (2013). Sentiment Analysis in the Planet Art: A Case Study in the Social Semantic Web. In *New Challenges in Distributed Information Filtering and Retrieval, SCI* (Vol. 439, pp. 131-149).

Blanco-Fernández, Y., Pazos-Arias, J. J., Gil-Solla, A., Ramos-Cabrer, M., López-Nores, M., Garcia-Duque, J., . . . Bermejo-Muñoz, J. (2008). A flexible semantic inference methodology to reason about user preferences in knowledge-based recommender systems. *Proceedings of the 17th international conference on World Wide* (pp. 92-101).

Blanco-Fernández Y., Pazos-Arias J.J., Gil-Solla, A., Ramos-Cabrer M., López-Nores M., Garcia-Duque J., Fernádez-Vilas A., Díaz-Redondo R.P., & Bermejo, M., (2010). Incentivized provision of metadata, semantic reasoning and time driving filtering: making a puzzle of personalized e-commerce. *Expert Systems with Applications: An International Journal,* 2010, 102-114.

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, *30*(1-7), 107–117. doi:10.1016/S0169-7552(98)00110-X

Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, *12*(4), 331–370. doi:10.1023/A:1021240730564

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dixit, V. S., Bedi, P., & Mehta, H. (2013). Generation of web recommendations using implicit user feedback and normalised mutual information. *Int. J. Knowledge and Web Intelligence*, *4*(2/3), 113–141. doi:10.1504/IJKWI.2013.056362

Ekstrand, M. D., Ludwig, M., Konstan, J. A., & Riedl, J. T. (2011). Rethinking the recommender research ecosystem: reproducibility, openness, and lenskit. *Proceedings of the fifth ACM Conference on Recommender Systems* (pp. 133–140). 10.1145/2043932.2043958

Fellbaum, C., Grabowski, J., & Landes, S. (1998). Performance and confidence in a semantic annotation task. Cambridge, Massachusetts, USA: MIT Press.

Francisco, V., Gervas, P., & Peinado, F. (2007). Ontological reasoning to configure emotional voice synthesis. *Proc. of RR*, *LNCS* (*Vol. 4524,* pp. 18-33). Springer. 10.1007/978-3-540-72982-2_7

Funk, S. (2006, December 11). Netflix Update. Retrieved from http://sifter.org/~simon/journal/20061211.html

Goldberg, D., Nichols, D., Oki, B. & Douglas, T. (1992). Using collaborative filtering to weave an information tapestry. *Communications of the ACM,* 35(12), 61–70.

Goldberg, K., Roeder, T., Gupta, D., & Perkins, C. (2001). Eigentaste: A constant time collaborative filtering algorithm. *Information Retrieval*, *4*(2), 133–151. doi:10.1023/A:1011419012209

Hill, W., & Terveen, L. (1996). Using frequency-of-mention in public conversations for social filtering. *Proceedings of the ACM conference on Computer supported cooperative work*, New York, USA (pp. 106–112). 10.1145/240080.240229

Kim, H. N., Alkhaldi, A., ElSaddik, A., & Jo, G. S. (2011). Collaborative user modeling with user-generated tags for social recommender systems. *Expert Systems with Applications*, 2011, 39–51.

Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L., & Riedl, J. (1997). Grouplens: Applying collaborative filtering to usenet news. *Communications of the ACM*, *40*(3), 77–87. doi:10.1145/245108.245126

Koren, Y., Bell, R., & Volinsky, C. (2009). Matrix Factorization Techniques for Recommender Systems. *Computer*, 42(8), 30-37.

Lang, K. (1995). Newsweeder: learning to filter netnews. *Proceedings of the 12th International Conference on Machine Learning* (pp. 331–339).

Lemire, D., & Maclachlan, A. (2005). Slope One prediction for online rating-based collaborative filtering. Proceedings of SIAM Data Mining (SDM'05), Newport Beach, California, USA (pp. 21-23).

Lianga, J., Rakesh, K., & Keith, W-R. (2006). The FastTrack overlay: A measurement study. *Journal of Computer Networks*, 50(6), 842-858.

Moreno, M. N., Pinho, J., López, V., & Polo, M. J. (2010). Multivariate Discretization for Associative Classification in a Sparse Data Application Domain. Proceedings of the International Conference on Hybrid Artificial Intelligence Systems, LNCS (Vol. 6076, pp. 104-111). Berlin, Germany: Springer.

Pang, B., & Lee, L. (2008). Opinion Mining and Sentiment Analysis. Now Publishers Inc.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Pazzani, M., & Billsus, D. (1997). Learning and revising user profiles: The identification of interesting web sites. *Journal of Machine Learning*, 27(3), 313–331.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). Grouplens: an open architecture for collaborative filtering of netnews. *Proceedings of the ACM* (pp. 12-22). 10.1145/192844.192905

Schafer, J., Konstan, J., & Riedl, J. (1999). Recommender systems in e-commerce. *Proceedings of the 1st ACM conference on Electronic commerce EC '99* (pp. 158-168). 10.1145/336992.337035

Schroder M., Baggia P., Burkhardt F., Oltramari A., Pelachaud C., Peter C., & Zovato E. (2010). Emotion Markup Language (EmotionML) 1.0.

Schroder, M., Pirker, H., Lamolle, M., Burkhardt, F., Peter, C., & Zovato, E. (2011). Representing emotions and related states in technological systems. In *Emotion-Oriented Systems* (pp. 369–387). Springer. doi:10.1007/978-3-642-15184-2_19

Shardanand, U., & Maes, P. (1995). Social information filtering: algorithms for automating word of mouth. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 210–217). 10.1145/223904.223931

Stumme, G., Hotho, A., & Berendt, B. (2006). Semantic web mining: state of the art and future direction. Journal of Web Semant, 4(2), 124-143.

Terveen, L., Hill, W., Amento, B., McDonald, D., & Creter, J. (1997). Phoaks: A system for sharing recommendations. *Communications of the ACM*, *40*(3), 59–62. doi:10.1145/245108.245122

Zhang, Y., Callan, J., & Minka, T. (2002). Novelty and redundancy detection in adaptive filtering. *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval* (pp. 81–88). 10.1145/564376.564393

This research was previously published in the International Journal of Knowledge Society Research (IJKSR), 7(3); edited by Miltiadis D. Lytras and Linda Daniela; pages 99-118; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 32
Knowledge-Based Support to the Treatment of Exceptions in Computer Interpretable Clinical Guidelines

Alessio Bottrighi
University of Piemonte Orientale, Italy

Luca Piovesan
Università degli Studi di Torino, Italy

Giorgio Leonardi
University of Piemonte Orientale, Italy

Paolo Terenziani
University of Piemonte Orientale, Italy

ABSTRACT

Clinical guidelines are one of the major tools that have been introduced to increase the rationalization of healthcare processes, granting both the quality and the standardization of healthcare services, and the minimization of costs. Computer interpretable clinical guidelines (CIGs) are widely adopted in order to assist practitioners in decision making, providing them evidence-based recommendations based on the best available medical knowledge. However, a main problem in CIG adoption is the fact that, in the medical context, some degree of uncertainty is often present. Thus, during guidelines executions on specific patients, unpredictable facts and conditions (henceforth called exceptions) may occur. A proper and immediate treatment of such exceptions is mandatory, but most of the current software systems coping with CIGs do not support it. In this paper, the authors describe how the GLARE system has been extended to deal with this purpose. They identify different types of exceptions, considering their "pre-locability" and "pre-plannability". On the basis of such parameters, the authors propose different treatment modalities, consisting of both data structures to model the different types of exceptions, and the algorithms to treat them. The resulting methodology is an innovative one, integrating different Artificial Intelligence techniques (ranging from planning to ontology-based reasoning). Finally, they also discuss how they implemented their system-independent methodology on top of GLARE, and describe its application in the ROPHS project, considering the management of the severe trauma guideline.

DOI: 10.4018/978-1-7998-0951-7.ch032

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. INTRODUCTION

1.1. Context

The rationalization of the healthcare system is important to grant both the quality and the standardization of healthcare services, and the minimization of costs. *Knowledge* is the cue factor to promote such a rationalization: it may support both strategic *decisions* at the organization level, as well as diagnostic/therapeutic *decisions* concerning specific patients. *Evidence-based* medicine has gained a major role in this context. Clinical Practice Guidelines (CPGs) are one of the major tools that have been introduced to provide *evidence-based medical decision support*. CPGs are, in the definition of the USA Institute of Medicine, "*systematically developed statements to assist practitioner and patient decisions about appropriate health care in specific clinical circumstances*" (Institute of Medicine, Committee on Quality Health Care in America, 2001). In short, CPGs provide evidence-based recommendations to suggest to physicians the set of diagnostic/therapeutic actions to be executed to cope with a specific disease. Notably, the last decision on the activities to be performed on the patient is always left to the user-physician. Thousands of CPGs have been developed in the last decade, or so. For instance, the Guideline International Network (http://www.g-i-n.net) groups 100 organizations of 48 countries, and provides a library of more than 6500 CPGs. Despite such a huge effort, CPGs have not provided all the expected advantages to clinical practice yet. Recent research has shown that Computer Science can help to drastically improve the impact of CPGs in clinical practice, by developing tools to manage computer-interpretable clinical guidelines (**CIG**s henceforth), defined as formal representations of CPGs in some computer format. The adoption of CIGs, and of computerized approaches to acquire, represent, execute and reason with provides crucial advantages to:

1. Patients, enabling them to receive the best quality medical treatments;
2. Physicians, providing them with a standard reference which they may consult, with a way of certifying the quality of their activity (e.g., for insurance or legal purposes), and, above all, with advanced support for their decision-making activity;
3. Hospitals, and healthcare centers, providing them with tools to enable the quality and the standardization of their services, as well as with a tool to evaluate quality, and to take strategical knowledge-based decisions for optimizing costs and resources.

In recent years, the research about CIGs has reached an important role within the Medical Informatics community, and many different approaches and projects have been developed to create domain-independent computer-assisted tools for managing, acquiring, representing and executing them. See e.g. the systems Asbru (Shahar, Miksch, & Johnson, 1998), EON (Musen, Tu, Das, & Shahar, 1996), GEM (Shiffman et al., 2000) GLARE (Terenziani, Molino, & Torchio, 2001) (Terenziani., Montani, Bottrighi, Molino, & Torchio, 2008), GLIF (Peleg et al. 2000), GUIDE (Quaglini, Stefanelli, Lanzola, Caporusso, & Panzarasa, 2001), PROforma (Fox, Johns, & Rahmanzadeh, 1998), and the collections (Gordon, & Christensen, 1995; Fridsma, 2001; Ten Teije, Miksch & Lucas, 2008; Peleg, 2013). One of such approaches is GLARE (Guideline Acquisition, Representation and Execution), which started from 1997 in a long-term cooperation between the Department of Computer Science of the University of Eastern Piedmont Alessandria, Italy, and the Azienda Ospedaliera San Giovanni Battista in Turin (one of the largest hospitals in Italy). Besides supporting CIG acquisition, representation, storage and

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

execution, GLARE is characterized by the adoption of advanced Artificial Intelligence and Temporal Database formal techniques to provide advanced supports for different tasks, including reasoning about temporal constraints (Anselma, Terenziani, Montani & Bottrighi, 2006), the treatment of periodic data (Stantic, Terenziani, Governatori, Bottrighi & Sattar, 2012), guideline versioning (Anselma, Bottrighi, Montani & Terenziani, 2013), model-checking verification (Bottrighi et al., 2010), decision support (based on Decision Theory) (Montani, & Terenziani, 2006; Anselma et al., 2011), CIG contextualization (Terenziani et al., 2004).

1.2. Handling Exceptional Cases

One of the main goals of CPGs (and of CIGs) is to capture medical evidence and to put it into practice. However, from one side, evidence is essentially a form of statistical knowledge, and is used to capture the generalities of classes of patients, rather than the peculiarities of a specific patient. From the other side, demanding to expert committees the elicitation of all possible executions of a CPG on any possible specific patient in any possible clinical condition is an infeasible task. Thus, two conditions are usually implicitly assumed by experts/committees building a CPG:

1. Ideal patients, i.e., patients that have just the single disease considered in the CPG (thus excluding the concurrent application of more than one CPG), and are "statistically relevant" (they model the typical patient affected by the given disease), not presenting rare peculiarities/side-effects;
2. Ideal context of execution, so that all necessary resources are available.

On the other hand, when a specific physician applies a given CPG to a specific patient, the patient and/or the context may not be "ideal". Hence, a CPG cannot be interpreted as a protocol which has to be executed exactly as it is (without any possible modification), and the actions prescribed by CPGs cannot be interpreted as "must do" actions. Informally speaking, given a patient *X* to which a CPG *G* has to be applied in a context *C*, *G* has to be interpreted as a set of default prescriptions: whenever *X* and *C* fit with *G*'s prescriptions, they must be executed, unless some exceptional situation arises. Specifically, *X* (or *C*) may have peculiar features, and unpredictable facts and conditions (henceforth called ***exceptions***) may occur which are not explicitly covered by *G*. In such a case, recommendations in the CIG cannot be followed any more, and additional *knowledge* must be exploited to provide appropriate treatments to the patient.

Ex.1: A patient treated for chronic atrial fibrillation, suddenly presenting massive gastrointestinal bleeding. The **CIG** for atrial fibrillation prescribes to perform anticoagulant treatment. However, the anticoagulant treatment is not performed. Since an unexpected problem threatening the patient's life has arisen, the execution of the CIG is suspended, and the massive gastrointestinal bleeding is treated first.

In general, especially in case of unexpected life threatening problems, the physician has to suspend/stop the "standard" execution of the CG, and to start to cope with such new problems. Henceforth, such problems will be called "exceptions", since they are exception to the "standard" execution of a CIG.

Despite the fact that supporting the treatment of "exceptions" is of dramatic importance for the practical applicability of software tools managing CIGs, only a limited attention has been devoted to such a

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

challenging task until now (see the Related Work Section). In the rest of the paper, we describe an approach to extend current CIG systems to cope with exceptions. Although the methodology we propose is mostly general, in this paper we will focus specifically on the extensions we made to the GLARE system.

The paper is organized as follows. In the second Section, we propose a preliminary discussion of the main features of GLARE. In the third section, we describe our ontology of exceptions, and our representation of exceptions within GLARE. In the fourth section we focus on the treatment of a specific type of exceptions, external exceptions, which involves limited forms of re-planning. In the fifth section, we discuss how GLARE execution engine has been extended to cope with exceptions. In the sixth section, we consider a practical example (the CIG about severe trauma). In the seventh section we discuss related work in the literature and, finally, in the eighth section we propose discussions and conclusions.

2. MAIN FEATURES OF GLARE

GLARE is a long-term project started in 1997 to provide a user-friendly and domain-independent tool to support CIG acquisition, representation and execution, and to provide the physician with a wide range of facilities, helping her/him in decision support or CIG analysis tasks. The main goals of the GLARE systems are:

1. To be domain-independent. In particular, GLARE has been already applied to deal with a wide range of CIG, ranging, e.g., from asthma to ischemic stroke.
2. To be user-friendly. GLARE is intended to be a tool to support physicians (in particular, regarding decision making), not at all to replace them. Such a goal has several implications, including the facts that (i) CIGs representation language must be as close as possible to the physicians' way of looking at it, and (ii) the "technological" complexity of the system must be hidden to the user-physicians, through a user-friendly and easy-to-use interface.
3. To be easily extendable. Indeed, physicians should be provided with plenty of facilities concerning CIG treatment. Thus, the system architecture must be modular and easily extendable.

2.1. GLARE Representation Language

In GLARE, a CIG can be represented as a hierarchical graph, where nodes are the actions to be executed, and arcs are the control relations linking them. GLARE distinguishes between *atomic* and *composite* actions (called *plans*), where atomic actions represent simple steps in a CIG, and plans represent actions that can be defined in terms of their components via the "has-part" relation. To be as physician-friendly as possible, a limited and "physician oriented" number of types of atomic actions have been identified:

- **Work actions**, i.e. actions that describe a procedure which must be executed at a given point of the CIG; *pharmacological actions* are specific subtype of work actions, specifying a drug (or drug category, of active principle) to be administered to the patient, and the posology.
- **Decision actions**, used to model the selection among different alternatives. In GLARE has been introduced a distinction between:
- **Diagnostic decisions**, used to make explicit the identification of the disease the patient is suffering from, among a set of possible diseases, compatible with her findings;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- **Therapeutic decisions**, used to represent the choice between paths in a GL, where each path represents a particular therapeutic process. The choice can be made by evaluating a fixed set of parameters (effectiveness, cost, side-effects, compliance and duration);
- **Query actions**, i.e. requests of information (typically patient's parameters), that can be obtained from the outside world (physicians, databases, patient visits or interviews).
- **Conclusions**, which explicitly identify the output of a decision action.

Each type of action is described by a set of properties. For instance, a *pharmaceutical action* is described by: *Action Type*, which for each action of this category takes the value "Administration of Medicament", *Substance*, which describes the administered drug, *Posology*, which describes the dosage and the modality of assumption.

Actions in a CIG are connected through ***control relations***. Control relations establish which actions can be executed next, and in what order. In particular, the ***sequence*** relation explicitly establishes what is the following action to be executed; the ***alternative*** relation describes which alternative paths stem from a decision action, the ***constrained*** relation establishes temporal constraints between actions to be executed (e.g., one action *during* the other), and the ***repetition*** relation, states that an action has to be repeated several times As well as the modality and periodicity of the repetitions (Anselma, et al., 2006).

2.2. Architecture of the Kernel of GLARE

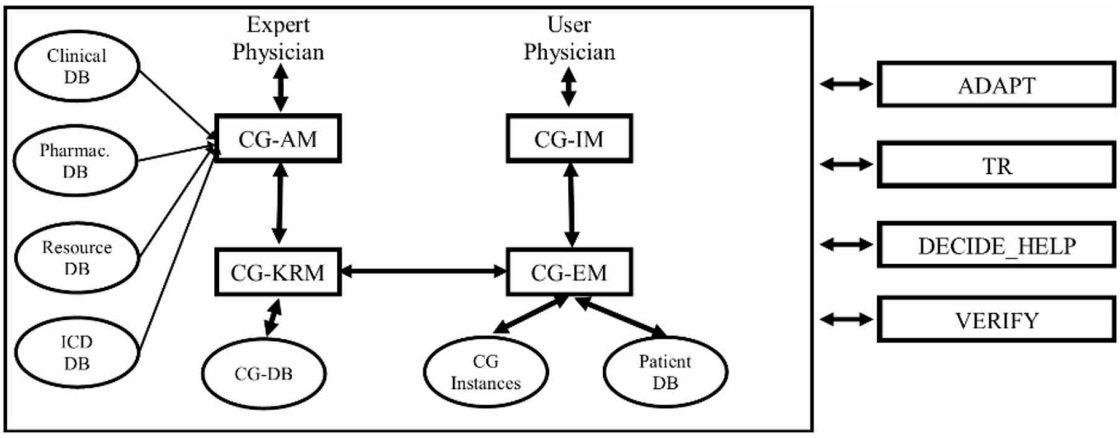
The core of GLARE (see box on the left of Figure 1) is based on a modular architecture. CG-KRM (Clinical Guidelines Knowledge Representation Manager) is the main module of the system: it manages the internal representation of CIGs, and operates as a domain-independent and task-independent knowledge server for the other modules; moreover, it permanently stores the acquired CIG in a dedicated Clinical Guidelines Database (CG-DB). The Clinical Guidelines Acquisition Manager (CG-AM) provides expert-physicians with a user-friendly graphical interface to introduce the CIG into the CG-KRM and describe them. It may interact with four databases: the Pharmacological DB, storing a structured list of drugs and their costs; the Resources DB, listing the resources that are available in a given hospital (it is therefore used to represent the context-dependent version of a CIG); the ICD DB, containing an international coding system of diseases; the Clinical DB, providing a "standard" terminology to be used when building a new CIG, and storing the descriptions and the set of possible values of clinical findings.

The execution module (CG-EM) executes a CIG for a specific patient, considering the patient's data (retrieved from the Patient DB). Roughly speaking, GLARE execution module starts the execution starting from the first action in the CIG, and goes on, following the control flow described by the control arcs in the CIG. Composite actions (plans) are simply decomposed in their atomic components. The execution of query actions demands for an access to the Patient DB. The execution of diagnostic decisions takes advantage of the data in the patient DB and of the decision criteria, to suggest one or more alternatives. However, the last decision, also in the case of therapeutic decision, is always left to the user physician. CG-EM directly supports the concurrent execution of (possibly temporally constrained) multiple actions, through the *agenda technique* (Terenziani, Mastomonaco, Torchio, & Molino, 2000).

CG-EM stores the execution status in another DB (CG Instances) and interacts with the user-physician via a graphical interface (CG-IM). A more accurate discussion of GLARE execution engine is reported in (Terenziani et al., 2000).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Architecture of GLARE. Rectangles represent computation modules, and ovals represents data/ knowledge bases. The left box contains the "core" of GLARE, while additional modules in the right of the figure depict advanced AI support for several tasks.



Through the CG-AM and CG-IM modules, GLARE provides user-friendly interfaces to user, achieving the goal of "hiding" them in the internal complexity of the system.

2.3. Extended Architecture

To enhance extendibility, GLARE's architecture is open: new modules and functionalities can be easily added to the system if\when necessary. In recent years, new modules and\or methodologies has been added in order to cope with automatic resource-based contextualization (ADAPT module, (Terenziani et al., 2004)), temporal reasoning (TR, (Anselma et al., 2006) and (Stantic et al., 2012)), decision-making support (DECIDE_HELP, (Montani, & Terenziani, 2006) and (Anselma et al., 2011)), and model-based verification (VERIFY, (Bottrighi et al., 2010)).

3. ONTOLOGY AND REPRESENTATION OF EXCEPTIONS

The management of exceptions within a software system managing CIGs arises several difficult challenges. For instance, in general, plans to cope with exceptions cannot be simply represented as new actions in CIGs, for the simple reason that exceptions are unpredictable, so that they can occur at any moment during the execution (or may not occur at all). Thus, an independent (from CIGs) representation of CIGs must be provided, as well as a methodology to combine the treatment of exceptions with the CIG execution. In this section, we focus on representation. First, we identify the "ontology" of exceptions. Second, we propose a formalism to represent the exceptions, and the plans to treat them. Our mechanism to integrate the treatment of exceptions with the "standard" execution of CIGs is then discussed in the following sections.

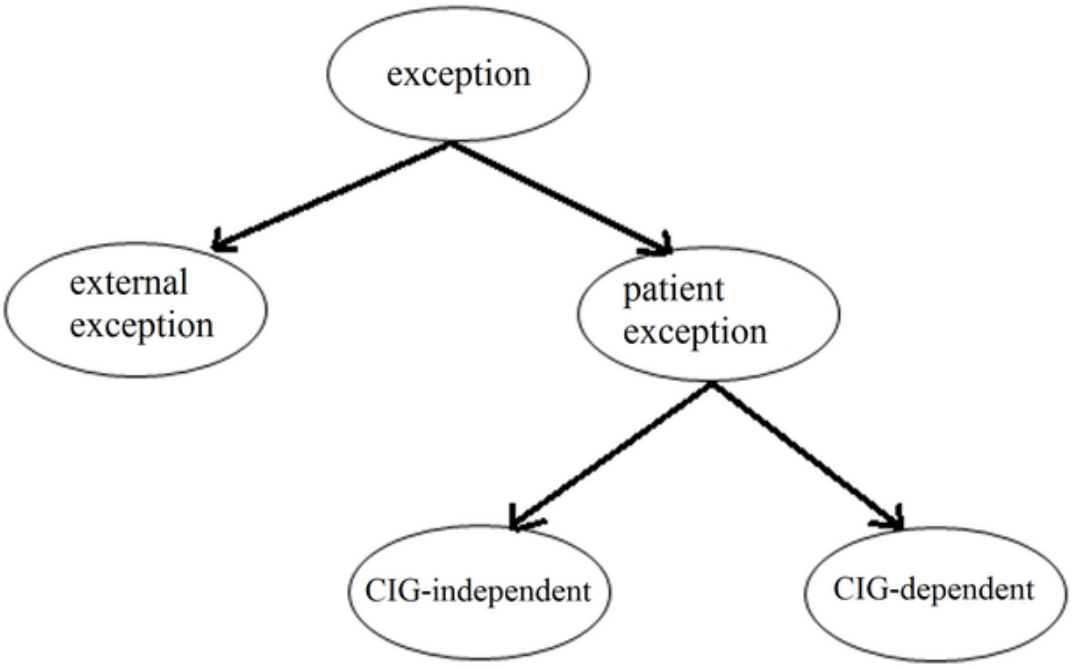
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.1. Ontology: Different Types of Exceptions

At least two different types of exceptions can be distinguished: exceptions related to the state of the execution of the actions in a CIG, and exceptions related to the state of the patient (see Figure 2). The first type (external-exception henceforth) arises whenever, for any contextual reason, the execution of the current action fails. For instance, the failure may be due to the absence of a resource needed to carry on the action (typically, an instrument). Another typical failure occurs when the deadline for the execution of the current action expires. The second type of exceptions (patient-exception henceforth) arises when the standard execution of a CIG cannot be carried on, due to an unexpected condition in the status of the patient. Indeed, we further distinguish into two subtypes of patient-exceptions: CIG-dependent and CIG-independent. CIG-dependent patient-exceptions are those exceptions that, although are not frequent enough to be explicitly managed in the CIG, can be foreseen to occur at a specific point in the execution of the CIG. For instance, an action of drug administration can raise an exception, in case a patient is allergic to such a drug. On the other hand, CIG-independent patient-exceptions are totally unpredictable, and may occur at any time during the execution of CIGs. For instance, an ischemic stroke may occur, requiring an immediate treatment, regardless of what is the CIG and the specific action in the CIG currently in execution.

The distinction among external-exceptions, CIG-dependent-patient-exceptions and CIG-independent-patient-exceptions is important also from an operational point of view. There are two orthogonal aspects to be considered here: "*pre-locability"* and "*pre-plannability"*. Such aspects deeply affect the treatment of exceptions, both from the point of view of their representation, and from the point of view of how they can be managed.

Figure 2. Exception taxonomy



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

With the term "pre-locability" we intend the possibility, for the experts defining a given CIG, to find (one or more) proper locations of the exceptions, at specific points (actions) of the CIG. In this sense, External-exceptions and CIG-dependent-patient-exceptions can be "pre-located". This fact affects the representation of the CIG (since the triggering conditions for such exceptions must be stored within the appropriate actions in the CIG) and the CIG execution engine (since the rise of pre-located exceptions must be verified when such exceptions are encountered, during the execution of CIG actions). On the other hand, CIG-independent-patient-exceptions cannot be "pre-located", since they are totally unpredictable. Thus, triggering conditions will need to be stored in a data structure (library) which is external to the CIG, and must be checked "continuously", by an independent process which may interrupt the execution of the current CIG when a trigger is satisfied.

With the term "pre-plannability" we intend the a-priori availability of a "plan", expressed by a CIG, for the treatment of the exception. Since patient-exceptions (both CIG-dependent and CIG-independent) correspond to general conditions of patients, it is reasonable to assume that expert physicians can provide proper guidelines to cope with them. In this work, we envision the development of a wide body of medical knowledge, consisting of a domain/guideline-independent library of guidelines, coping with patient-exceptions. On the other hand, external-exceptions have a different nature. They do not concern the state of the patient, but depend on "external" contextual situations that are neither under the control of physicians, nor depends on the patient state. Assuming that also in case of External-exceptions expert physicians provide a library of guidelines seems to us quite unrealistic. We thus suggest a more general approach to cope with such exceptions, based on general medical knowledge about the goals of guideline actions, and on a limited form of "planning" on such a knowledge.

3.2. Representation of Exceptions

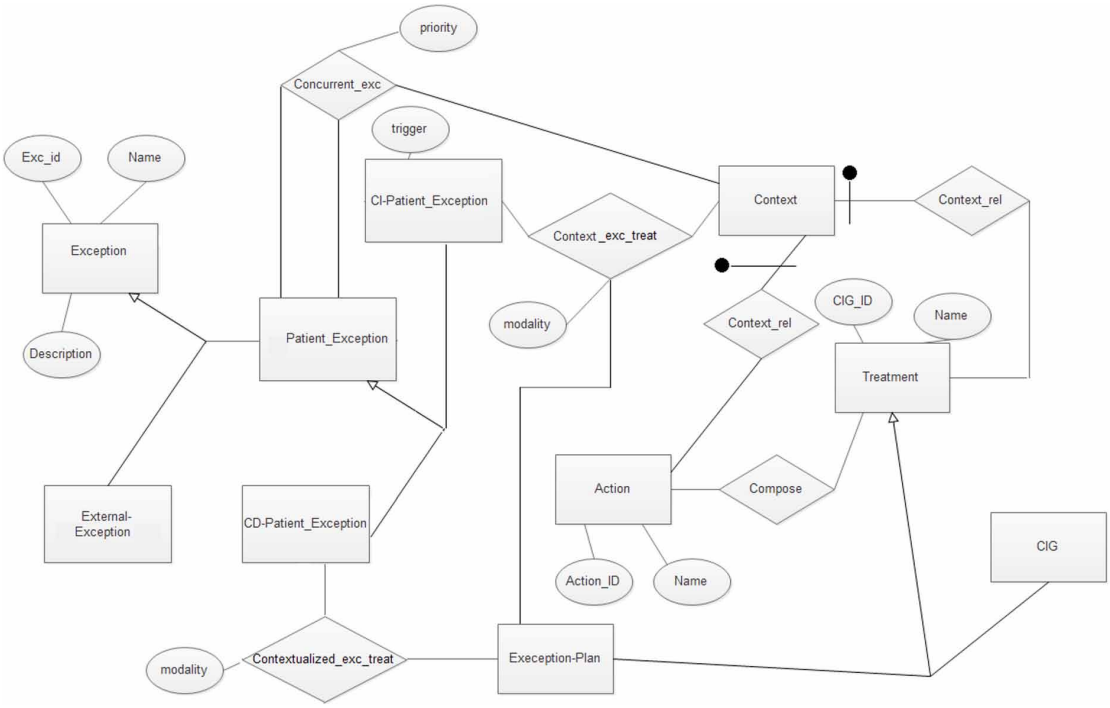
Until now, we have distinguished among three main types of exceptions. We now focus on their representation. In the following, we propose the representation of exceptions we adopt in our approach. Such a representation involves different parts of the system: CIG representation, Library of Exceptions, and Exception DB.

3.3. CIG Representation

First of all, the description of actions in the CIG has been modified, in order to represent "pre-locable" exceptions. External exceptions are modelled as an additional property (the property <Completed_Execution>) of *work actions*. When executing a work action, the property <Completed_Execution> leads to a request to the user-physician, who has to indicate whether the action has been successfully executed in the external world or not. Of course, in the case of failure, the treatment of external exceptions will be launched. Also, the description of actions is augmented with the <Exception_List> property, which is optional and, if present, indicates the identifiers (<Exc_ID> in the *Exception DB,* see also Figure 3) of the CIG-dependent patient-exceptions that may occur at that point.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Entity relationship model of Exception DB



3.4. Library of Exception Treatments

In our proposal, the plans to cope with "pre-plannable" exceptions are stored in a library (the Exception Treatment Library) which complements the library of available CIGs. As discussed above, also exception treatments can be described using GLARE formalism. This grants for generality, and makes exception execution easier (since exception treatments can be directly managed by the CIG execution module).

3.5. Exception DB

In our approach, all the other pieces of information describing the exceptions are stored into a new database (to be added to the ones shown in Figure 1), the *Exception DB*. In the following, we adopt the Entity-Relationship notation (see Figure 3), to represent the content of such a DB, since, from one side, it grants for the generality of the description and, on the other side, it supports the internal logical and physical representation in the relational model.

Each exception (<Exception> in Figure 3) can be identified by an attribute <Exc_ID>, and is characterized by textual attributes <Name> and <Description> (see the left corner of Figure 3). No additional attribute is needed for the description of external-exceptions (<External-Exception> in Figure 3). Indeed, as discussed above, such exceptions are not "pre-plannable". This means that the actions to recover from such exceptions are not fixed a-priori, but will be determined when needed, on the basis of the available medical knowledge. The recovery actions will be executed instead of the failed action, and, after their execution, the standard execution of the current CIG will be carried on.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

On the other hand, several different aspects must be modelled, considering patient-exceptions: (1) triggers, (2) recovering plan, (3) synchronization, (4) priorities, and (5) context-dependence.

1. **Triggers:** Patient-exceptions arise when specific well-known conditions characterize the state of the patient. Thus, a set of triggering conditions must be associated with each patient exceptions (see the attribute <Trigger> associated with <CI-Patient-Exception> in Figure 3). When such conditions become true, the exception arises for the given patient.
2. **Recovering Plan:** Since patient-exceptions are "pre-plannable", they have an associated recovery plan (<Exception-Plan> in Figure 3). It is worth noticing that, thanks to the expressiveness of GLARE's formalism, exception plans can be represented using the same formalism we adopt in order to specify CIGs (<CIG> in Figure 3). Thus, GLARE supports a library containing both CIGs and Exception-plans. As we will see, this solution allows us to deal in a homogeneous way (through the same execution engine) with both the execution of CIGs and the execution of plans coping with exceptions.
3. **Synchronization:** While in the case of external-exceptions the recovering actions have to substitute the failed action in the CIG, in the case of patient-exception a wider range of possibilities must be taken into account. The attribute <Modality> in Figure 3 specifies what must be the interplay between the execution of the exception plan and the execution of the current CIG. For the sake of generality, we envision five different types of interplay
    a. *Concurrent*. The plan for the exception must be executed concurrently with respect to the current CIG;
    b. *Suspend*. The CIG execution is suspended, and will be resumed after the end of the execution of the plan for the exception;
    c. *Abort*. The current CIG is aborted. Only the plan for the exception is executed;
    d. *Abort&Goto(X,*CIG_i*)*. The execution of the current CIG is aborted. The plan for the exception is executed. After this execution, the execution engine executes the action X of the guideline CIG_i.
    e. *Ignore*. The Exception is ignored; so that the standard execution of the current CIG continues (this modality is useful in cases when the execution of the CIG is crucial for the patient health, while the treatment of the exception is somehow not relevant, given the current CIG).

Notice that the methodology above easily supports the ***nesting*** of CIG-dependent patient-exceptions. The cue idea for facilitating the treatment of nested exceptions is to represent exception plans as CIG themselves. Thus, from the point of view of our system, there is no difference whether an exception arises during the execution of a CIG, or during the execution of an exception plan (which, recursively, may arise during the execution of a CIG, or of an another exception). In all cases, the recovering plan is activated, and is synchronized with the process activating it. Specifically, depending on the chosen modality, such nested exceptions may be run concurrently, or may be suspended, or may abort the execution of the exception plan in which they are nested.

4. **Priorities:** Unluckily, more than one CIG-independent patient-exception might be triggered together, for the same patient. Unfortunately, there seems to be no general rule to identify a unique correct behaviour, covering all possible situations. We thus allow the possibility to specify, for chosen pairs <Ei,Ej> of patient-exceptions, the priority between them (attribute <Priority> of the relationship

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

<Concurrent_exc>; such a relationship will be explained later on). We support different values for such an attribute: (1) E_i may *override* E_j (so that only E_i is executed) or vice-versa, or (2) E_i may be executed *before* E_j (or vice-versa), or (3) E_i and E_j may be executed *concurrently*. Since not all the co-occurrences are known a-priori for each CIG, it is possible to define a minimum set of known general rules as the default behaviour, leaving guideline authors free to override these rules, or to add new ones, to provide different selection strategies for specific CIGs. Furthermore, these rules can be maintained incrementally: after defining a first set, new rules can be added when new exceptions are added to the repository, or each time a new co-occurrence has been experienced by the CIG users and defined as a new rule.

5. **Context-Dependence:** Though the above description of exception is, to the best of our knowledge, the most comprehensive one in the CIG literature, it still neglect a major practical problem: the fact that, unfortunately, the clinical procedures adopted to cope with patient-exceptions are not "absolute", but may change depending on the context (and, in particular, on the specific CIG or action in the CIG which is currently under execution).

3.6. Example

Let us consider an episode of sudden systolic blood pressure decrease (i.e. the value of systolic pressure falls under 90 mmHg). The procedure to cope with this exception during a hemodialysis session suggests a nurse to administer osmotic drugs (e.g. mannitol) to re-establish proper blood pressure levels, which allow the patient to terminate the session. In the severe trauma guideline, instead, different treatments are specified, which are furthermore highly dependent on the current action in the guideline (see the section about the sample CIG implementation).

We propose an innovative approach to the treatment of exceptions, characterized by the fact that we support the possibility of specifying different plans of treatment and different modalities of execution for the same exception, depending on the context when it occurs (i.e., depending on which action of which CIG is being executed when the exception arise). This attention to the context is a major departure with respect to the treatment of exceptions in object oriented programming, as well as in the current approaches to exceptions in the CIG literature. However, this step is extremely important, to cope with practical medical applications.

To cope with such cases, we allow the possibility of specifying context-dependent plans and modalities for patient-exceptions.

Indeed, CIG-dependent patient-exceptions (entity <CD-Patient_Exception>) are, by definition, "contextualized" (since they are associated with a specific action in a specific CIG). They are thus directly associated, through the relationship <Contextualized_exc_treat>, to their <Exception-Plan>. Notice that the <Modality> attribute is associated with such a relationship, to model the synchronization between the current CIG and the exception plan.

On the other hand, several entities and relationships have been introduced in Figure 3 to cope with context-dependence. First of all, <Context> is a weak entity and is identified via the two relationships <Context_rel> connecting it to the entities <Treatment> (which is a super-entity of <CIG> and <Exception-plan>) and <Action> (which represents the actions composing the <Treatment>). In such a way, we define specific contexts of execution by specifying an action, and a treatment containing it. The core notion to represent context-dependence of treatments is the relationship <Context_exc_treat>. It is a ternary relationship connecting (i) a CIG-indepentent patient-exception (entity <CI-Patient-Exception>),

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

a context (<Context>), and an execution plan (<Exception-Plan>). The <Modality> attribute is associated to such a relationship, to model the fact that the treatment of CIG-indepentent patient-exceptions depends on the context, and also the modality of synchronization. Of course, at the implementation level, defaults can be specified, in case the same treatment (and or modality) is applicable to an entire CIG (independently of the current action), or to all the CIGs (context-independent treatment of an exception).

Moreover, the priorities between two exception treatments could depend on the <Context> and thus it is modelled by the attribute <Priority> of the ternary relation <Concurrent_exc>, which involves two exceptions (i.e. the super-entity <Patient_exception> of CI-Patient_Exception and CD-Patient_Exception) and a <Context>.

4. DEALING WITH EXTERNAL EXCEPTIONS

Differently from the other types of exceptions, external exceptions are related to outer factors that, independently of the state of the patient, may lead to a failure of the current action in the CIG. Given their nature, we cannot expect that the treatment of such exceptions can be pre-planned in the library. On the other hand, we aim at proposing a comprehensive approach, supporting user-physician also in the case of external exceptions. To such a purpose, we propose a knowledge-based and goal-based approach, in which a limited form of planning is adopted. The core idea is quite simple. External exceptions arise due to the failure of *work-actions* (see the subsection "GLARE representation language" above). Obviously, each one of such actions is present in a CIG in order to achieve some goals (e.g., decrease blood pressure). If the action fails, it can thus be substituted by other actions that allow one to reach the same goals.

We thus propose a typical Artificial Intelligence approach, in which a knowledge base represents general medical knowledge, specifying actions and their goals, and a goal-oriented form of reasoning is adopted to find the actions to cope with a failure (i.e., to achieve the goals of the failed action). In the following, we first describe our simple ontology of actions and goals, and then we describe our reasoning algorithm.

4.1. A Medical Ontology of Actions and Goals

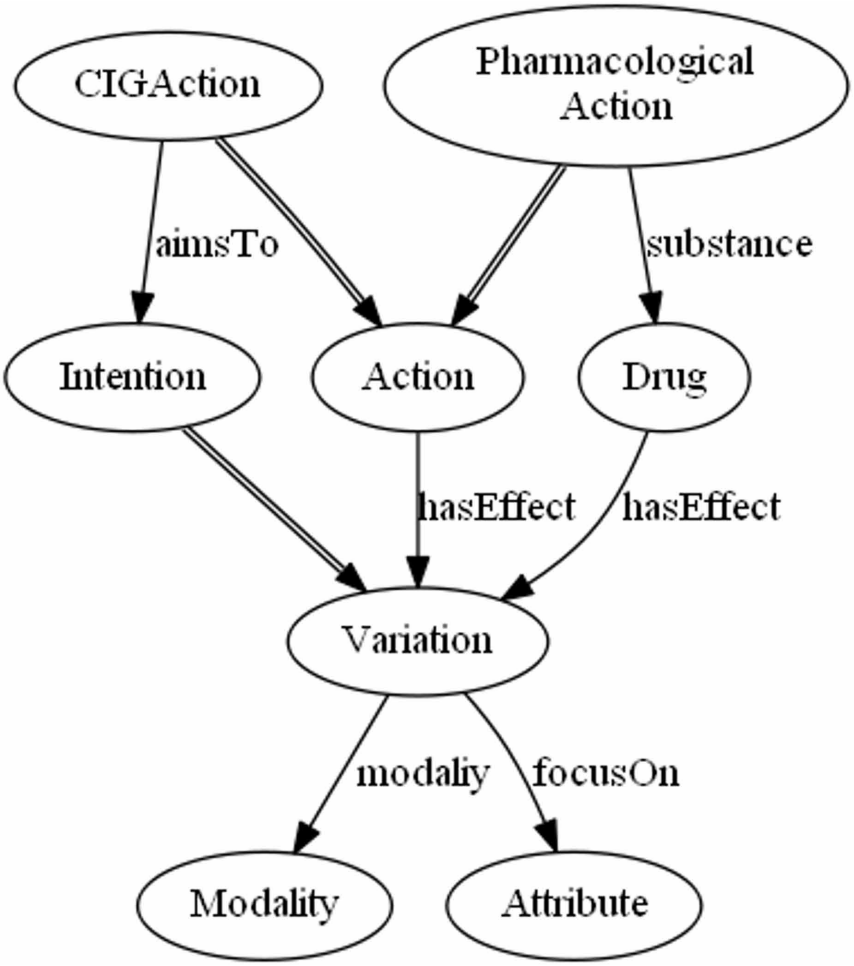
In the medical domain, decisions are often driven by the synthesis of several knowledge sources. Modelling and exploiting such a knowledge are open problems of the Medical Informatics community. In the ontological extension of GLARE (Piovesan, Molino, & Terenziani, 2014), we decided to represent at least a part of such a medical knowledge through the development of an ontological model of clinical actions and, consequently, also of CIG actions. Our model is an extension of consensus medical ontologies (in our approach, we used SNOMED CT (International Health Terminology Standards Development Organisation)).

The idea underlying our model is that medical actions are not "primitive" elements, but entities that can be described through their relations with, for instance, their goals, their effects, and through their classification in a taxonomy. In the following, we will describe a part of our ontology, which is particularly relevant for this work.

In the ontological extension of GLARE (Piovesan et al., 2014), we characterize clinical actions (concept *Action* in Figure 4) through several properties, including its effects (*hasEffect* arc). We characterize such effects as variations (concept *Variation*), with certain modalities (concept *Modality*; e.g., increase,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Ontological representation of CIG actions



decrease) of patient state parameters (concept *Attribute*). For instance, the effect "urine alkalinization" (characterizing, for example, the Calcium Carbonate administration) is modelled through the variation "increase" of the attribute "urine pH".

Actions in a CIG (*CIGActions*) have an additional attribute (*aimsTo*) relating them to the goals or intentions (concept *Intention*) they are intended to achieve in the specific CIG. Such intentions are still variations of the state of the patient. However, notice that there is subtle distinction between the (CIG-specific) intentions of CIG actions and the effects of actions "per se" (i.e., independently of any CIG in which they can be considered). For example, while the urine alkalinization is a general effect of the administration of calcium carbonate, in the CIG for the treatment of the gastroesophageal reflux such an administration has the (CIG-specific) goal of decreasing the gastric acidity.

In our model, variations (and intentions) are hierarchically organized in an IS-A and a PART-OF taxonomy. For instance, the intention of "block of calcium channels" is a sub intention (IS-A relation) of "decrease blood pressure", since the first one is a more specific way of achieving the second one. PART-OF relations allows to decompose "composite" intentions into their subcomponents.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

For instance, the intention of "improve lifestyle" can be decomposed into "improve diet", "increase physical activity" and "stop smoking". Intentions are associated with CIG actions at each level of abstraction of the CIGs.

Pharmacological actions (i.e., drug administrations) are of paramount importance in the medical domain. Pharmacological actions are clinical actions (and, of course, may be also CIG actions) that are related with the drug (arc *substance*) they prescribe. Drugs are hierarchically organized following the ATC classification (WHO Collaborating Centre for Drug Statistics Methodology). Also drugs are related to their effects (arc *hasEffect*) and, for integrity purpose, the action of administering a drug has the same effects (or superconcepts of them) of the drug it recommend.

4.2. The Planner for Coping with External-exceptions

In the following, we describe how the previous model can be exploited, in order to identify plans to recover from an external exception. However, before describing it, we have to make some considerations about the adopted solution. Many approaches are present in Artificial Intelligence (AI) literature for the adaptation of plans that failed due to action failures or due to unsatisfied preconditions. However, given the peculiarities of CIGs, only the most conservative approaches can be adopted in real world situations. Indeed, CIGs are developed by experts following the medical evidence, and the solutions proposed by decision support systems must adhere as much as possible to the existing plans. In addition, our experience in the medical domain suggested that physician would not trust "black box" tools that simply output plans without explaining to the user (in an "understandable" way) how such conclusions have been reached.

Considered such desiderata, we decided to face the recovering from external exceptions through a *mixed initiative* (Horvitz, 1999) goal-oriented plan repair strategy. Exploiting our ontology, we developed a tool highly collaborative with the user physicians, supporting her/him in the repairing of the current CIG. It is worth stressing that such approach could potentially be used also for replanning from scratch in case a large part of the CIG fails. However, we intentionally limited it to local (in most cases single-action) repairing, to propose solutions as close as possible to the original (evidence-based) plan.

Our algorithm helps physicians in the navigation of our ontology, finding out the relevant alternatives for the failed action. We distinguish between two possible cases: the recovering involves

1. only one (not decomposable) intention, or
2. more than one intention.

First, let us focus on the first (basic) case (i.e., case (i) above), which occurs when one CIG action failed, and it has only one intention (not decomposable -through the PART-OF taxonomy- into a set of sub-intentions).

For the basic case, we need to retrieve all the actions (*actset*) that satisfy (*hasEffect* arc) the intention *int* of the failed action *a*.

```
algorithm basic_af:
input: CIGAction a;
output: Set<Action>;
Intention int ← aimsTo(a);
Set<Action> actset ← inverseof  hasEffect^noTrans(int);
```


EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

```
for each action act ∈ actset
     issuit ← the system asks the
               user if act is a suitable alternative in the specific situation
     if issuit == 'no'
          actset ← actset \ {act}
return actset;
```

The algorithm for the management of the basic case algorithm *basic_af* (basic action finding) is divided in three phases. First, the intention *int* of the failed action a is retrieved from the CIG model. Then, the set *actset* containing all the actions[1] which have as effect (*hasEffect* arc, followed backward) *int* is retrieved from the ontological knowledge. Notice that the set *actset* also includes composite actions (i.e., actions whose components have been specified through the PART-OF taxonomy). Finally, *actset* is pruned by asking the user to discard uninteresting actions, in order to maintain only the relevant results. At the end of the algorithm, *actset* represents the set of possible alternatives to the failed action a, satisfying its intention and considered relevant by the physician.

In case *actset* is empty, there is no direct way to recover from the exception. However, a more flexible treatment of such an exception can be tried, in case the failed action *a* is part of a composite action *a'* in the CIG. In such a case, the basic algorithm can be applied considering *a'* as the failed action. Such a procedure can be recursively applied, until either one (or more) recovery action is found, or the highest level of abstraction in the CIG is reached.

Case (ii) involves the management of more than one intention. There are three different cases in which the treatment of more than one intention is required: (a) if the failed action has itself two or more intentions or (b) if the intention of the failed action is a composite one (so that all its sub-intentions have to be satisfied), or (c) if two or more actions failed, each one with its own intention(s).

The management of case (ii) requires some modifications in the basic algorithm, since two or more intentions must be satisfied by the proposed solution. For the sake of simplicity, we only describe how our algorithm manage the case of two intentions int_1 and int_2

When two intentions int_1 and int_2 have to be satisfied, the procedure of action finding in the ontology described for the basic case is applied separately for the two intentions. Two resulting sets of actions are obtained, named actset1 and actset2. The system then returns the Cartesian product of $actset_1$ and $actset_2$ as possible solutions. Notice that the result of Cartesian product between two sets is a set of sets

$$A = \{a_1, a_2, \ldots, a_n\}, B = \{b_1, b_2, \ldots, b_m\}$$

$$A \times B = \{\{a_1, b_1\}, \{a_1, b_2\}, \ldots, \{a_n, b_m\}\}$$

If two actions $a_i \in A$ and $b_j \in B$ are equal, such couple will produce in the result the element $\{a_i\}$ so that the result does not contain duplicates.

The development of solutions containing two or more actions raises two problems: on one side, interactions between the actions can occur (e.g., drug-drug interactions)[2]; on the other side, no ordering constraints are returned by the system. However, such a kind of issues are very common in the management of the concurrent execution of two or more CIGs (e.g., for the treatment of comorbid patients). Thus, for instance, the techniques presented in (Piovesan, Anselma, & Terenziani, 2015) to cope with concurrent execution of CIGs can be easily adapted for this use.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

5. MANAGING EXCEPTIONS DURING CIG EXECUTION

The execution engine of CIGs (CG-EM in Figure 1; henceforth called *Executor* for the sake of clarity) must be widely extended in order to cope with exceptions. Different extensions are needed, to cope with the different types of exceptions pointed out in the previous section. Indeed, roughly speaking, the *Executor* executes the actions in the CIG, following the flow of control described by the control arcs in the CIG. "Pre-locable" exceptions can be managed at specific points during the flow of execution. Therefore, such exceptions can be directly managed by the Executor, which must be properly extended.

In particular, to cope with **external-exceptions**, the Executor simply checks, at the end of each action, whether it has been accomplished with success, or whether it has failed (in the current version, success and failure are simply signaled to GLARE by the user-physician executing the CIG; i.e. the property <Completed_Execution> of *work action*). In the case of success, the Executor simply continues its standard execution, moving to the next action in the CIG. On the other hand, in the case of failure, the planner for external-exceptions is invoked, having in input the goals of the failed action. The new plan returned by the planner (navigating the ontological knowledge and interacting with the user-physicians, as discussed in the previous section) is then executed by the Executor. After the execution of the new plan, the standard execution is restarted, from the next action in the CIG.

The treatment of **CIG-dependent patient-exceptions** is easier. The basic *Executor* has been modified in such a way that, before the execution of each CIG action, the triggering conditions of the CIG-dependent patient-exceptions associated to the current action (if any) are checked. If an exception is triggered, the proper exception-plan is searched in the library of exception plans, considering the context (current action and current CIG in execution), as well as its synchronization modality. The Executor loads the related exception plan and its modality of execution (as explained in subsection "Representation of exceptions", the modality may be unique for the given exception, or may depend on the current CIG and/or action). Then, depending on the modality, the following actions are performed:

Figure 5. Creation of new operands and associated SQL query using the graphical editor. The names in the right figure refer to real patient's data defined in the RoPHS (Report on the Piedmont Health System) project



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. In the case of "*Concurrent*" modality, the Executor inserts in the agenda also the first action in the plan for the exception, and re-start its standard agenda-based execution (it is worth remembering that (1) GLARE's original *Executor* supports the concurrent execution of the actions in the Agenda, and (2) since the plan for the exceptions are represented using the guideline formalism, the CIG Executor itself can be used in order to execute such plans).
2. In the case of "Suspend" modality, the first action of the exception plan is put on top of the Agenda. Thus, the Executor executes this plan first, and then returns to the other actions in the Agenda, recovering the execution of the suspended CIG.
3. In the case of "Abort" modality, the Agenda is first made empty (i.e., all the actions in it are deleted). Then, the first action of the exception plan is put on top of the Agenda. Thus, the Executor executes this plan first, and then terminates (since the Agenda becomes empty).
4. The "Abort&Goto(X,CIG_i)" modality is similar to "Abort", except that, after the completion of the exception plan, the action X of the guideline CIG_i is inserted in the Agenda, forcing the *Executor* to jump to that specific point of the CIG.
5. In the case of "Ignore" modality, the standard execution of the current CIG is simply carried on.

On the other hand, **CIG-independent patient-exceptions** are not "pre-locable": they can arise at any moment, during the execution of a CIG. For such a reason, they cannot be properly associated with any specific point in CIG. On the contrary, their triggering conditions should be continuously monitored, considering the state of the patient (as stored in the patient DB).

For such a reason, the proper treatment of CIG-independent patient-exceptions demands the definition of a new engine, that we call ***Exception-monitor***. Of course, the *Exception-monitor* and the *Executor* run concurrently. The Exception-monitor can interrupt the Executor whenever an exception is triggered. At that point, The *Executor* loads the related exception plan and its modality of execution (they may depend on the current CIG/action, i.e. the contest) and goes on with the execution of the CIG and/or the exception plan, depending on the given synchronization modality (as discussed above for the case of CIG-dependent patient-exceptions).

Before concluding this section, it is worth mentioning two features of the approach we propose.

First of all, the methodology above easily supports CIG-dependent exception nesting. Depending on the chosen modality, such nested exceptions may be run concurrently, or may be suspended, or may abort the execution of the exception plan in which they are nested.

Second, we also support cases in which two CIG-independent exceptions, defined in the exceptions library, are triggered together, through the management of (possibly context-dependent) priorities (see the discussion in the subsection "Representation of exceptions").

5.1. Implementation

This subsection presents a brief description of the exception handling system we developed and its integration with the GLARE system. In the next Section, real world examples, implemented in GLARE using the functionalities of this new exception manager, are provided in the context of severe trauma management, to show the applicability of our strategy in a real context.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

5.2. System Components Implementation

We implemented the concepts exposed in an exception management system, which has been integrated and tested in GLARE. The exception management system is implemented in Java and is composed by two main components: (1) a graphical editor for user-friendly definition and maintenance of the exceptions library, and (2) proper extensions to the GLARE Executor to allow the CIG-independent patient-exceptions handling. Each exception of this type is associated to a triggering condition, which defines a particular (problematic) patient's state. Exceptions are triggered when the associated triggering conditions become true, independently from the status of execution of the actions in the CIG. To access the patient-related information (e.g. the blood pressure or the body temperature) needed to periodically verify these triggering conditions, the system relies to SQL queries to retrieve the most up-to-date information from the patient's data base.

5.3. Graphical Editor and Exception Repository Maintenance

The exception management system offers a user-friendly graphical editor, which permits the CIG author to: (1) build new triggering conditions combining operators, operands and constant values; (2) create new exceptions and eventually customize the modality of execution of the exception plan for a particular CIG plan (see "Representation" above), and (3) define the rules to be applied by the Executor when two or more exceptions are triggered at the same time (see "Integration of Exception and CIG Executions" above).

In particular, the graphical editor helps the user in performing these three tasks providing the following features. Considering the task (1), the editor permits the user to define new triggering conditions (for example the condition " 'systolic pressure' <= '90 mmHg' AND 'cardiac frequency' < '45 bpm') or to edit an existing one by combining operators (i.e. arithmetic, logic and comparison), operands (such as 'systolic pressure') and constants (e.g. '90 mmHg') by means of a guided procedure. For each operand (see Figure 5), the editor can automatically build the SQL query which will be used run-time to retrieve the most up-to-date values associated to the considered operand from the patients database, each time this condition must be evaluated. Note that, since in Figure 3 trigger is an attribute of <CI-Patient-Exception>, at the implementation level we have defined a triggering library in order to allow the reuse of a triggering condition, which can be associated with a unique identifier (condition_ID).

Task (2) helps the user in defining new exceptions or in modifying existing ones. An exception is defined as follows: <Exc_ID, name, description, Context_ID, condition_ID, exceptionPlan_ID, modality>, where:

- exc_ID: unique identifier of the exception
- name, description: name and natural language description of the exception
- Context_ID: identifies the context, for which this exception must be checked, see section "Exception DB"
- condition_ID: identifier of the triggering condition to be verified to raise this exception
- exceptionPlan_ID: identifier of the exception plan to be executed to manage this exception
- modality: modality of execution of the exception plan (Concurrent, Suspend, Abort, Abort & Goto, Ignore)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Once the user has indicated the Exc_ID (or accepted an auto-generated identifier), she/he must: (1) enter name and description of the associated exception, (2) choose a condition from the triggering condition library and an exception plan from the exception plans library. The standard execution modality (if defined) for the chosen plan is automatically loaded, but the user has the possibility to override it, by changing the modality or by customizing it to indicate, for example, to go to a particular action of the CIG plan, in case of Abort & Goto.

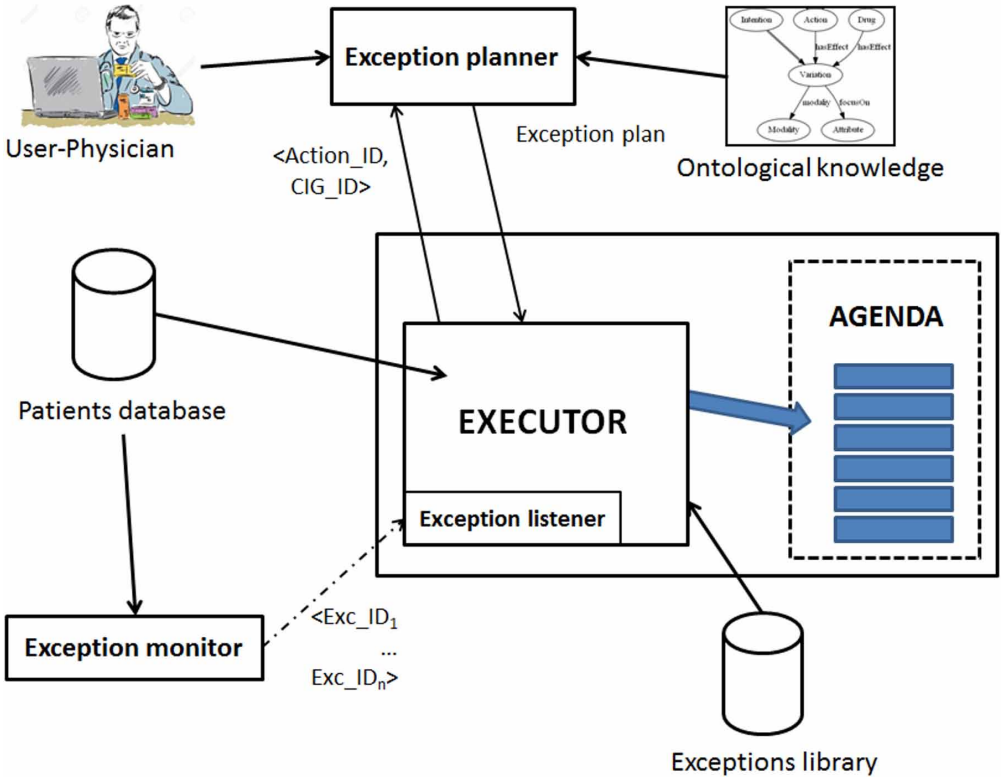
Finally, task (3) permits to define, for each pair $<E_i, E_j>$ of exceptions for one Context_ID, the rules to be applied (i.e. one of them overrides or must be executed before or after the other one, or both of them must be executed concurrently) when they happen at the same time.

All the exceptions defined by the user are permanently stored in the Exceptions library, in order to be automatically loaded and used by the Executor, when a new instance of a CIG plan is created and run. This repository can be updated at any time using this graphical editor, adding, deleting or modifying triggering conditions, exceptions and co-occurrence rules.

5.4. CIG Executor Modifications

The CIG Executor of GLARE has been modified and extended to cope with exceptions at run-time. The new version of the Executor now loads and exploits the exceptions saved in the Exceptions library when exceptions arise during the execution of a CIG on a specific patient.

Figure 6. The exception management system's architecture



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

CIG-dependent patient-exceptions are directly managed by the Executor, as discussed in the previous section. On the other hand, the treatment of external-exceptions and CIG-independent patient-exceptions require an extension to the architecture of GLARE. In particular, the architecture extensions are illustrated in Figure 6: while a CIG is executed on a patient the Exception monitor, responsible for checking and raising the CIG-independent patient-exceptions runs concurrently and asynchronously with the CIG Executor, sending messages to the latter as soon as one or more exceptions are raised.

On the other side, the Executor can receive messages from the Exception monitor by means of an integrated Exception listener. Any time a message is received, the Executor stops immediately the execution of the current CIG plan, loads the proper exception plan(s) from the exceptions library and applies the proper steps to allocate it run-time on the agenda, in order to manage the exception(s) risen from the monitor. Finally, the Executor continues the execution of the CIG, augmented with the allocated exception plan(s). This architecture allowed us to obtain a system which is able to react run-time to the CIG-independent exceptions, without modifying the existing GIC Executor (CG-EM in Figure 1).

Additionally, if an external-exception is risen by the failure of the execution of a CIG action, the external-exception planner is invoked. It takes in input the failed action (the couple <Action_ID, CIG_ID> in Figure 6), and navigates the ontology, possibly interacting with user-physicians, in order to identify a new plan, which is sent to the CIG Executor to recover the failed action (Exception-Plan in Figure 6).

The exception monitor has been implemented as a Java thread (subclass of the *java.lang.Thread* class), and its constructor permits to receive:

1. the set of exceptions to be checked (the set E of tuples <exc_ID, name, description, CONTEXT_ID, trigger, exceptionPlan_ID, modality>, where CONTEXT_ID matches the current context of execution and exceptionPlan_ID identifies the treatment (see CIG_ID in figure 3) to be executed to manage this exception, and
2. the time interval T which must elapse between the check of two successive exceptions. T may also be set to zero, to enforce continuous monitoring of triggering conditions.

In particular, the exception monitor executes the following operations (implemented by overriding the method *run()* of the *Thread* class):

1. Wait for the period T to elapse
2. For each exception *exc* **ϵ** E:
   a. Let P the triggering condition identified by: *exc.conditionID*, associated to the exception *exc.exceptionID*
   b. For each operand *op* **ϵ** P, execute the query *op.query* on the patients database to retrieve the most up-to-date values
   c. Solve the triggering condition P and store its result R (true/false) as a pair <exc_ID, R>
3. If at least one of the couples <exc_ID, R> has R = true, then:
   a. Build a message M containing all the couples <exc_ID, R>, where R = true
   b. Send the message M to the Executor
4. Jump to point 1

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

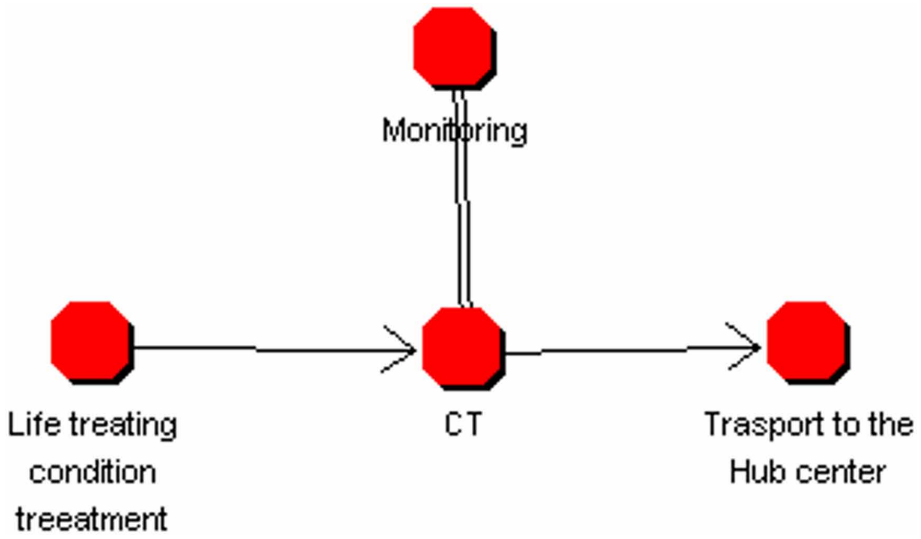
The exceptions to be risen are defined as all the exc_IDs for which the corresponding value of R is set to true. This set of exceptions will be sent to the Executor through a message containing them, in order to be managed by the Executor itself. The operations from point 1 to point 4 form an infinite cycle, which will be interrupted by the Executor (by interrupting the thread implementing the monitor) only when the treatment of the patient (CIG and –possibly- exception plan executions) reaches its end.

On the other side, the Engine has been modified to receive messages from the Exception Monitor, thanks to an instance of a custom class implementing the MessageListener Java Interface.

The treatment of a patient starts with the execution of the proper CIG. A specific instance of the CIG is created by the Executor, and started on the patient. Meanwhile, the exception monitor starts monitoring the exception rules and the exception listener allows the Executor for continuously listening to potential asynchronous messages arriving from the monitor. While no messages are received from the exception monitor, the Executor continues with no interruptions. As soon as one or more exceptions are detected by the exception monitor, a message is sent to the Executor which, after receiving the message, stops the execution of the current CIG, performs all the operations needed to load the proper exception plans(s) in the Agenda and restarts the execution. The following pseudo-code illustrates this process in detail. Each time the Exception listener receives a message M from the exception monitor, the Executor performs the following steps:

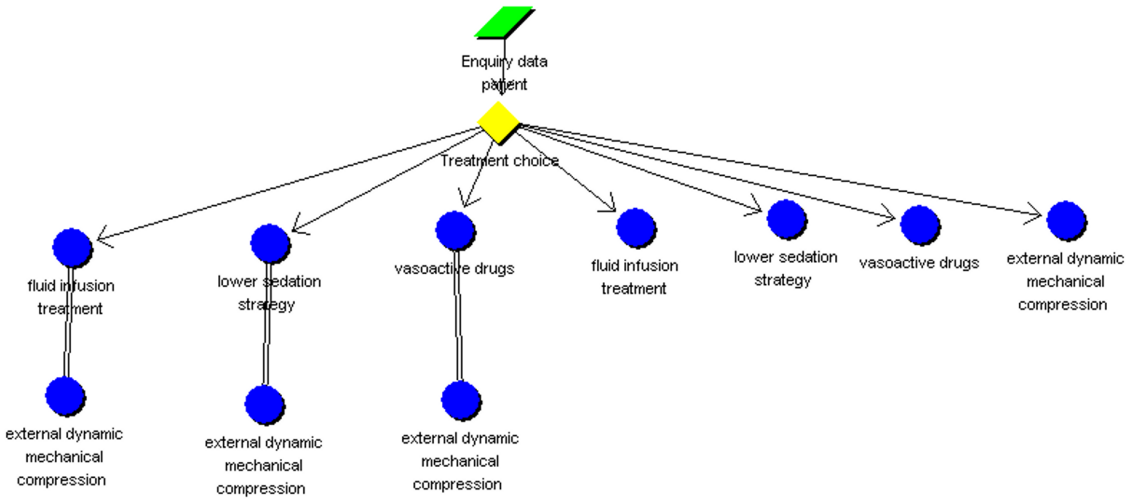
1. Stop executing the current CIG
2. Unpack the message M to retrieve the set of exceptions X= {exc_ID_1, …, exc_ID_n} in M
3. If X contains two or more exceptions, then apply the priority rules to decide what exception E must be managed among all those in X; otherwise, E = the only exception In X
4. Retrieve from the Exceptions library the exception plan(s) to manage the exception X
5. Update the Agenda with the plan(s) retrieved at point 5, accordingly to the policies to be applied to manage the modality of execution associated to the(se) plan(s)
6. Restart the execution of the actions in the Agenda

Figure 7. Fragment of the severe trauma guideline acquired in GLARE



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. The treatment plan for systolic blood pressure decrease exception meanwhile "Life treating condition treatment" plan execution



The steps 1 and 6 are necessary to stop and to restart the execution of the CIG before and after the Agenda has been altered in the steps 4 and 5. After having performed all these operations, the new plan(s) are thus seen by the Executor as part of the (augmented) CIG. The process described continues until the treatment (CIG instance and, possibly, exception plans) reaches its end and, consequently, the Executor shuts down both the exception monitor and the engine.

It is worth noting that through this technique, we obtained a run-time and asynchronous management of the CIG-independent exceptions without modifying the core of the existing Executor. Furthermore, the extensions described are totally transparent to the user, since this important additional service allows the system to execute CIGs in a flexible and more true to life way in a completely automatic fashion. In the case where no exceptions are defined for a particular CIG, the latter will be executed such as no CIG-independent patient-exceptions management system exists.

6. AN EXAMPLE OF CIG IMPLEMENTATION

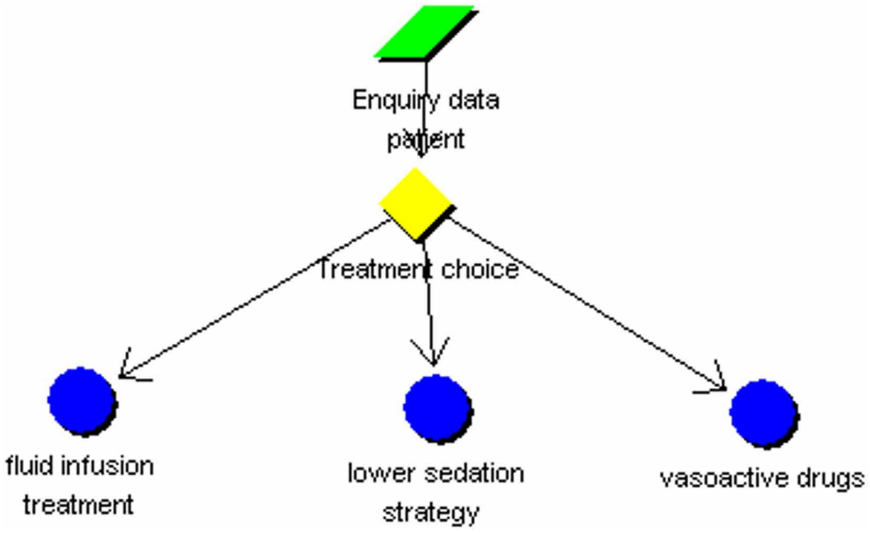
In this section we describe the application of our approach with real examples in the context of severe trauma treatment. The severe trauma guideline, briefly described in this section, is developed under the RoPHS (Report on the Piedmont Health System) project, funded by the Regione Piemonte.

This guideline is based on the evidence of the need of a synergistic organization to reduce the time of intervention and the excessive delay before hospital admission. In particular, the guideline is based on the following principles:

1. an appropriate clinical out-of-hospital evaluation of trauma patients for a prompt identification of lesions and therapeutic priorities;
2. the ability to perform some effective life-treating actions to support cardiac and pulmonary functions;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. The treatment plan for systolic blood pressure decrease exception after "Life treating condition treatment" has been executed



3. the correct clinical decision must be made to refer the patient to an hospital offering a competent treatment of the trauma, not necessarily the nearest, and
4. the optimum operative linking between the services for the acute management and the rehabilitation hospital.

Moreover, the guideline exploits the "Hub and Spoke" strategy, in order to emphasize the integration of services throughout the various service areas to ensure efficacy of treatment of the trauma patients. The "Hub and Spoke" model relies on a leader center (i.e. the hub), where the trauma patients may be treated for severe neurologic, thoracic, abdominal and vascular traumas. Spoke centers are strategically placed in the territory to ensure an appropriate treatment in the geographical area and are closely linked with their hubs through the emergency system (e.g. the 118 service, in Italy or the 999 and 112 services, in United Kingdom). In the context of the RoPHS (Report on the Piedmont Health System) project, the guideline is developed and applied on Eastern Piedmont. In this area, the Hub center is located in Novara, and the Spoke centers are hospitals located in Borgosesia, Borgomanero, Domodossola, Verbania, Vercelli and Biella.

Different exceptions are defined for this guideline: exceptions which can occur during the execution of the entire guideline, and exceptions which can occur only in specific sections of the guideline. For example, an exception applied on the whole guideline regards the heart failure. In the case that the patient has a heart failure, the current action and the guideline are immediately aborted, (since the <Modality> of this exception is set to Abort), and the treatment plan for heart failure exception is executed.

Considering exceptions occurring in a specific section of the guideline, shown in Figure 7, the "Life treating condition treatment" plan is followed by two concurrent plans: (1) "CT" plan (defining all the actions concerning the choice and the execution of the proper kind of Computed Tomography), and (2) "Monitoring" plan, where the patient is monitored with a multi-parametric monitor to gather information about his/her heart rate, blood pressure, pulse oximetry and saturation. Finally, the patient will be

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

transported in a specialized hospital, called Hub center. During this part of guideline, exceptions on monitored data can happen. For example, an exception to cope with the problem of sudden systolic blood pressure decrease is defined. This exception captures episodes where the systolic pressure falls under 90 mmHg. If the exception monitor, while checking the patient's systolic pressure, verifies that it is lower than 90 mmHg, it raises this exception. The execution engine will finally capture the exception and will start the proper treatment (exception) plan. The treatment plan is dependent from the action executed at the time of the exception, since its management in the severe trauma guideline is context dependent:

- If the current action(s) belong(s) to the "Life treating condition treatment" plan, this plan will be executed concurrently with the exception (i.e. <Modality> is set to Concurrent). This exception is managed by choosing one or more of the following actions: fluid infusion treatment, vasoactive drugs, and lower sedation strategy. These actions can be performed as unique treatments, or can be associated to external dynamic mechanical compression, in order to manage an eventual arterious external bleeding. Moreover, also the external dynamic mechanical compression can be performed as a unique treatment (see Figure 8).
- Otherwise, if the exception is captured after the conclusion of the "Life treating condition treatment" plan, the exception plan, even if executed concurrently with the guideline such as the previous one, relies on a different strategy. The actions to be performed in this plan are: fluid infusion treatment, vasoactive drugs, and lower sedation strategy. Note that there is not external dynamic mechanical compression among them (neither unique treatment nor in combination with one other treatment) since, after the "Life treating condition treatment", the patient's conditions are more stable and arterious external bleeding cannot occur (see Figure 9).

The system GLARE, extended as described above, allowed to cope with all the exceptions needed to properly manage the severe trauma guideline, representing and managing them during the execution of this CIG.

7. RELATED WORK

Considering the generality and relevance of the phenomenon, in the following we discuss the more relevant approaches in the literature, covering exceptions representation and management.

PROforma language (Sutton, & Fox, 2003) allows one to specify exceptions, while the system can set an Exception flag if an abnormal event has occurred in the processing of a guideline operation. However, in our approach, exceptions are hierarchically organized and we distinguish among three different types of exceptions. Moreover, for every exception, it is possible to define a specific treatment plan to manage it.

Tu & Musen (1999) differentiate between normal flow and exceptional flow and contemplate the specification of scenario-based exceptions handlers. Tu & Musen (1999) only consider exceptions regarding the execution of actions in the guideline. In their approach only the current action(s) can generate exceptions (such as our CIGs dependent exception). Our approach is more general, since we also allow to define exception on the basis of the patient's data. Moreover, Tu & Musen (1999) provides only two modalities of interactions between CIGs and exceptions: (1) exceptions leading a patient back to a scenario covered by the guideline (as our Abort&Goto modality), and (2) exceptions where the patient is managed outside the guideline (Abort modality in our approach).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hierarchies of exceptions are managed by (Grando, Peleg, & Glasspool, 2010). They distinguish between "hazards" (concept somehow close to our CIG-independent exceptions) and "obstacles" (as our CIG-dependent exceptions), mainly focusing on what we call External-exceptions, allowing hierarchical refinement. They take into account some of the modalities of interactions between CIGs and exceptions we also pointed out in our approach: parallelExecuting (close to our Concurrent modality), suspending (Suspend in our approach) and discarding (Abort in our approach). However, in their approach, exception treatment is always goal oriented: corrective actions are chosen accordingly to the substitution of the default goal expressed in the CIG. In our approach, such a goal-oriented treatment of exception is adopted just in the case of external exceptions, and is based on ontological knowledge. Additionally, we focus on context-dependent treatments of exceptions, and on both "pre-locable" and not "pre-locable" exceptions.

Quaglini et al. (2001) managed the flexibility of their guideline-based careflow system introducing a classification of exceptions which is similar to our ontology. Exceptions can be expected or unexpected, synchronous or asynchronous and GL-related or unrelated. Our approach differs in the exception handling, since in (Quaglini, et al., 2001) the triggers are requested by the physician who, at a certain point of the CIG execution, asks the system to delay the current action, to substitute the current action with another one or to add a new non-planned action. The system will then manage this trigger to cope with the physician's request. Our system, instead, triggers exceptions when an abnormal patient's state is detected, and automatically loads from the repository a plan to cope with the triggered exception, without the need of human intervention. This difference is justified, because manages exceptions in a careflow management system, where the physician must have the experience (and must be free) to introduce the proper deviations under her/his responsibility. Our system is a CIG-based decision support system, therefore it is responsibility of the system to suggest not only the proper actions to be performed in a normal situation, but also to recognize abnormal states of the patient and to suggest what must be done to treat the patient accordingly.

Peleg, Somekh, & Dori (2009) propose a methodology for eliciting and modelling exceptions. They describe a conceptual model in which exceptions can be expected or unexpected, and are triggered as asynchronous events. Triggers start a synchronous exception management branch during the CIG execution. Peleg et al. (2009) define two types of exception causes: (1) HumanCause, e.g. human errors, malicious actions or non-compliance, and (2) Non-humanCause, e.g. organizational causes, work item failures, deadline expirations. However, they do not distinguish between CIG-dependent and CIG-independent exceptions, and the exception management, briefly sketched, describes an approach similar to our Suspend modality, but goal driven: after the execution of the corrective actions to manage an exception, a set of goals should fulfilled to resume the CIG normal execution otherwise the CIG is aborted. On the other hand, our approach defines a complete exception management system, where different modalities of execution can be defined for each exception plan.

8. DISCUSSION AND CONCLUSION

Computer interpretable clinical guidelines (CIGs) are widely adopted to assist physicians in decision making, providing them evidence-based recommendations based on the best available medical knowledge. However, a main problem in CIG adoption is the fact that, during CIG executions on specific patients, unpredictable facts and conditions (exceptions) may occur. Decision support in the case of exceptions

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

is mandatory, but most current software systems coping with CIGs do not provide it. In this paper, we proposed a general approach to cope with exceptions, showing how software systems coping with CIGs can be extended (in the representation formalism, as well as in the execution module). Moreover, we have demonstrated the feasibility of our approach extending the GLARE system (Terenziani et al., 2008) and applying it the severe trauma guideline developed under the RoPHS (Report on the Piedmont Health System) project.

As discussed in the "Related work" section, some CIG approaches have already faced the treatment of exceptions. However, our approach is quite comprehensive and innovative, distinguishing between different types of exceptions, and proposing a wide range of methodologies to cope with them. In particular, we distinguish between exceptions than can be "pre-planned" or not, "pre-located" or not, considering a wide range of modalities of interactions between CIG and exception execution, and supporting the context-dependence of exception treatments. We also support the goal-based knowledge-based planning of new actions in substitution of CIG actions failed for "external" reasons.

Despite the wide range of phenomena being addressed, several interesting aspects deserve further analysis and work.

The availability of a large library of exception-plans is critical for the practical applicability of our approach. For the moment, we only cope with a very limited number of exception plans in our library. This is because the main goal of our work was that of devising a general methodology, and, only as a proof of concept, of implementing a prototypical version of it. However, an extensive adoption of our methodology in the real world would require the development of an extensive library of "pre-planned" CIGs coping with exceptions. This would require a substantial amount of work of the medical community, to devise evidence-based plans to cope with exceptions, and of the medical informatics community, to encode them into CIG formats. However, given the increasing adoption of computer tools in the medical field, and the increasing focus on CPGs and on exceptions and personalized medicine, it is hopeful that such a wide library can be developed, in the middle\long term.

In the presented methodology, we have proposed a limited adoption of *planning* techniques, to cope only with the substitution of "failed" actions during CIG execution (*external-exceptions*). Indeed, such knowledge-based and goal-based planning techniques could be used much more extensively, to cope also with other forms of exceptions (e.g., to build new complex plans –CIGs– coping with patient-exceptions not covered in the library of "pre-planned" exceptions yet). Such a solution would mitigate the problem of "manually" developing a large library of computerized exception plans. Therefore, such a new line of research is probably worth future investigation. However, our twenty-year experience of cooperation with physicians has taught us that physicians are quite reticent to accept new treatment plans automatically identified by software systems. There are two main reasons for that. First of all, many physicians are naturally diffident towards software tool "fully taking their place" in the definition of medical treatments. But, even more important, the main reason for the adoption of CPGs (and CIGs) is the desire to promote evidence-based treatments. Indeed, automatically generated plans, though "technically" correct and knowledge-based, cannot be evidence-based, since evidence can only be gathered through the definition and study of clinical trials. As a consequence, we try to limit as much as possible to delegate work to the planner, and suggest to exploit as much as possible libraries of evidence-based exception plans.

A further issue concern the treatment of patients affected by multiple diseases (the so-called *comorbid* patients). The treatment of comorbid patients is one of the main challenges for modern healthcare, also due to population aging and the increase of chronic diseases. For example, most of the patients suffering from such chronic diseases, are affected by two or more of them (Gijsen, Hoeymans, & Schellevis,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2001). Comorbid patients are not "ideal" patients, since two (or more) CIGs must be combined. Although in many cases one can suppose that all the relevant CIGs to be combined are available, the problem of combining them is a very critical and challenging one, since the interactions between the actions in the different CIGs must be properly analysed and managed. We have recently started to address such a challenging issue, proposing a knowledge-based mixed-initiative approach to support user-physicians in the detection and the analysis of interactions (Piovesanet al., 2014)

However, substantial work will be needed to integrate such a proposal with the current approach, to provide a comprehensive methodology coping with many different types of "exceptional" patients.

ACKNOWLEDGMENT

The work described in this paper has been partially developed within the RoPHS (Report on the Piedmont Health System) project, funded by the Regione Piemonte, and within the GINSENG project, funded by Compagnia San Paolo. The authors are very grateful to. Francesco Della Corte (Department of Translational Medicine, Univ. Piemonte Orientale "A. Avogadro", Novara, Italy) and Dr. Antonio Messina (Azienda Ospedaliera "Maggiore della Carità", Novara, Italy), for their support in the development of the CIG about severe trauma.

REFERENCES

Anselma, L., Bottrighi, A., Molino, G., Montani, S., Terenziani, P., & Torchio, M. (2011). Supporting knowledge-based decision making in the medical context: The GLARE Approach. *International Journal of Knowledge-Based Organizations*, *1*(1), 42–60. doi:10.4018/ijkbo.2011010103

Anselma, L., Bottrighi, A., Montani, S., & Terenziani, P. (2013). Managing proposals and evaluations of updates to medical knowledge: Theory and applications. *Journal of Biomedical Informatics*, *46*(2), 363–376. doi:10.1016/j.jbi.2012.12.004 PMID:23380684

Anselma, L., Terenziani, P., Montani, S., & Bottrighi, A. (2006). Towards a comprehensive treatment of repetitions, periodicity and temporal constraints in clinical guidelines. *Artificial Intelligence in Medicine*, *38*(2), 171–195. doi:10.1016/j.artmed.2006.03.007 PMID:16766167

Bottrighi, A., Giordano, L., Molino, G., Montani, S., Terenziani, P., & Torchio, M. (2010). Adopting model checking techniques for clinical guidelines verification. *Artificial Intelligence in Medicine*, *48*(1), 1–19. doi:10.1016/j.artmed.2009.09.003 PMID:19864118

Fox, J., Johns, N., & Rahmanzadeh, A. (1998). Disseminating medical knowledge: The PROforma approach. *Artificial Intelligence in Medicine*, *14*(1-2), 157–181. doi:10.1016/S0933-3657(98)00021-9 PMID:9779888

Fridsma, D. B. (Ed.). (2001). Special Issue on Workflow management and clinical guidelines. *Journal of the American Medical Informatics Association*, *22*(1), 1–80. PMID:11141509

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Gijsen, R., Hoeymans, N., Schellevis, F. G., Ruwaard, D., Satariano, W. A., & van den Bos, G. A. M. (2001). Causes and consequences of comorbidity: A review. *Journal of Clinical Epidemiology*, *54*(7), 661–674. doi:10.1016/S0895-4356(00)00363-2 PMID:11438406

Gordon, C., & Christensen, J. P. (Eds.). (1995). *Health Telematics for Clinical Guidelines and Protocols*. IOS Press.

Grando, A., Peleg, M., & Glasspool, D. (2010). A goal-oriented framework for specifying clinical guidelines and handling medical errors. *Journal of Biomedical Informatics*, *43*(2), 287–299. doi:10.1016/j.jbi.2009.11.006 PMID:19944779

Horvitz, E. (1999) Principles of Mixed-Initiative User Interfaces. *Proc. CHI 1999*. 159-166.

Institute of Medicine, Committee on Quality Health Care in America. (2001). Crossing the quality chasm: A new health system for the 21st century. Washington. USA: National Academy Press. International Health Terminology Standards Development Organisation. Retrieved from http://www.ihtsdo.org/snomed-ct

Musen, M. A., Tu, S. W., Das, A. K., & Shahar, Y. (1996). Synthesis of research: EON: A component-based approach to automation of protocol-directed therapy. *Journal of the American Medical Informatics Association*, *3*(6), 367–388. doi:10.1136/jamia.1996.97084511 PMID:8930854

Peleg, M. (2013). Computer-interpretable clinical guidelines: A methodological review. *Journal of Biomedical Informatics*, *46*(4), 744–763. doi:10.1016/j.jbi.2013.06.009 PMID:23806274

Peleg, M., Boxwala, A. A., Ogunyemi, O., & Zeng, Q. T. Tu, S.W., Lacson, R., Bernstam, E.V., Ash, N., Mork, P., Ohno-Machado, L., Shortliffe, E.H., & Greenes, R.A. (2000). GLIF3: the evolution of a guideline representation format. *Proc. AMIA 2000* (pp. 645-649).

Peleg, M., Somekh, J., & Dori, D. (2009). A methodology for eliciting and modeling exceptions. *Journal of Biomedical Informatics*, *42*(4), 736–747. doi:10.1016/j.jbi.2009.05.003 PMID:19442762

Piovesan, L., Anselma, L., & Terenziani P. (2015). Temporal detection of guideline interactions. *Proc. Health Informatics 2015*.

Piovesan, L., Molino, G., & Terenziani, P. (2014). An ontological knowledge and multiple abstraction level decision support system in healthcare. *Decision Analysis*, *1*(8), 1–24.

Quaglini, S., Stefanelli, M., Lanzola, G., Caporusso, V., & Panzarasa, S. (2001). Flexible guideline-based patient careflow systems. *Artificial Intelligence in Medicine*, *22*(1), 65–80. doi:10.1016/S0933-3657(00)00100-7 PMID:11259884

Shahar, Y., Miksch, S., & Johnson, P. D. (1998). The Asgaard project: A task-specific framework for the application and critiquing of time-oriented clinical guidelines. *Artificial Intelligence in Medicine*, *14*(1-2), 29–51. doi:10.1016/S0933-3657(98)00015-3 PMID:9779882

Shiffman, R. N., Karras, B. K., Agrawal, A., Chen, R., Marenco, L. N., & Nath, S. (2000). Model formulation: GEM: A Proposal for a more comprehensive guideline document model using XML. *Journal of the American Medical Informatics Association*, *7*(5), 488–498. doi:10.1136/jamia.2000.0070488 PMID:10984468

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Stantic, B., Terenziani, P., Governatori, G., Bottrighi, A., & Sattar, A. (2012). An implicit approach to deal with periodically repeated medical data. *Artificial Intelligence in Medicine*, *55*(3), 149–162. doi:10.1016/j.artmed.2012.03.002 PMID:22503730

Sutton, D. R., & Fox, J. (2003). Application of Information Technology: The syntax and semantics of the PROforma guideline modeling language. *Journal of the American Medical Informatics Association*, *10*(5), 433–443. doi:10.1197/jamia.M1264 PMID:12807812

Ten Teije, A., Miksch, S., & Lucas, P. (Eds.). (2008). *Computer-based medical guidelines and protocols: A primer and current trends* (Vol. 139). Studies in Health Technology and Informatics.

Terenziani, P., Mastromonaco, F., Molino, G., & Torchio, M. (2000) Executing clinical guidelines: temporal issues. *Proc. AMIA'2000* (pp. 848-852).

Terenziani, P., Molino, G., & Torchio, M. (2001). A modular approach for representing and executing clinical guidelines. *Artificial Intelligence in Medicine*, *23*(3), 249–276. doi:10.1016/S0933-3657(01)00087-2 PMID:11704440

Terenziani, P., Montani, S., Bottrighi, A., Molino, G., & Torchio, M. (2008). Applying artificial intelligence to clinical guidelines: the GLARE approach. In A. Ten Teije, S. Miksch, & P. Lucas (Eds.), *Computer-based medical guidelines and protocols: A primer and current trends, Studies in health technology and informatics* (Vol. 139, pp. 273–282).

Terenziani, P., Montani, S., Bottrighi, A., Torchio, M., Molino, G., & Correndo, G. (2004). A context-adaptable approach to clinical guidelines. In M. Fieschi, E. Coiera, Y.-C.J. Li, (Eds.), *Proc. Medinfo 2004 World Congress on Medical Informatics* (pp. 169–73). IOS Press.

Tu, S. W., & Musen, M. A. (1999). A flexible approach to guideline modeling. *Proc AMIA 1999* (pp. 420–424).

WHO Collaborating Centre for Drug Statistics Methodology. (n. d.). Retrieved from http://www.whocc.no/atc_ddd_index/

ENDNOTES

1 The keyword "inverseof" in the algorithm states that the arc *hasEffect* is followed backward. The superscript "noTrans" in the algorithm states that we do not consider the transitivity of the IS-A property for such query. In other words, only those actions that are directly linked to the effect *int* are returned, while their sub categories do not. Such a choice is adopted in order to limit the dimension of the output, making it manageable by user physicians.

2 Indeed, depending on the nature of the exception (and, thus, of the exception plan), the problem of interacting actions might arise also in the case of exceptions, since new actions (the actions in the exception plan) have to be "merged" with the action in the CIG. If the exception plan is specific of a given CIG (or, even better, of a specific point of a specific CIG –as in the case of "pre-locable" exceptions–), it is reasonable to assume that the possible interactions have been already considered during the definition of the exception plan, so that there is no need to check them when the excep-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

tion arise. On the other hand, CIG-independent-patient-exceptions cannot be "pre-located". Thus, one cannot know a priori at what point of the execution of a given CIG their exception-plan has to be executed. In such a context, the analysis of the possible interactions between the current actions in the CIG and the one in the exception plan may be of paramount importance. Such an analysis may be carried on through the techniques we have proposed in (Piovesan et al., 2015)

This research was previously published in the International Journal of Knowledge-Based Organizations (IJKBO), 6(3); edited by John Wang; pages 1-27; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 33
Statistical Relational Learning for Collaborative Filtering a State-of-the-Art Review

Lediona Nishani
University of New York Tirana, Albania

Marenglen Biba
University of New York in Tirana, Albania

ABSTRACT

People nowadays base their behavior by making choices through word of mouth, media, public opinion, surveys, etc. One of the most prominent techniques of recommender systems is Collaborative filtering (CF), which utilizes the known preferences of several users to develop recommendation for other users. CF can introduce limitations like new-item problem, new-user problem or data sparsity, which can be mitigated by employing Statistical Relational Learning (SRLs). This review chapter presents a comprehensive scientific survey from the basic and traditional techniques to the-state-of-the-art of SRL algorithms implemented for collaborative filtering issues. Authors provide a comprehensive review of SRL for CF tasks and demonstrate strong evidence that SRL can be successfully implemented in the recommender systems domain. Finally, the chapter is concluded with a summarization of the key issues that SRLs tackle in the collaborative filtering area and suggest further open issues in order to advance in this field of research.

INTRODUCTION

With the dramatic spread of Internet, many connections and interactions have taken place in the networks. Therefore, information relying on people reactions has experienced a dramatic growth. Social Networks, which are named due to the crucial role of humans, are concentrated in the social relations that people have among them. In order to exploit these large amounts of data, it is required designing new algorithms and appropriate techniques for gathering valuable information. The field of expertise,

DOI: 10.4018/978-1-7998-0951-7.ch033

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

which deals with the aforementioned challenges, is Social Network Mining *(SNM).* Its major purpose is to extract information about network objects, behaviors and activities that cannot be yielded if the entities were to be examined individually.

A social network represents a social structure of entities known as nodes, which are related by one or more specific types of interdependency, such as friendship, common interests or same opinions. Social network analysis considers social relationships based on the network theory perspective; nodes represent people, whereas ties represent relations between them. Together they form a graph-based structure, which experiences significant complexity. Social networks can be categorized in four significant types due to various domains on which they operate:

1. **Citation Networks:** Deal with storing scientific papers, the relation of their subjects across their paper and corporate authors or joint paper reference
2. **SNS (Semantic Network Service) Websites**: Concern the organization of people and their related friends or professional contact for example: Facebook, LinkedIn, Instagram, twitter, etc
3. **Social Shopping Websites**: Related to e-commerce and opinion mining exploring products
4. **Social Media Websites**: Concern the suggestions about movies, music relying on user preference

Social Network Analysis has appeared to be one of the major techniques in social networks (Wasserman & Faust, 1994) intending to establish a model that can map entities and the relations between them. The probability of two entities to build a link between them may depend on the likelihood of having similar attributes. On the other hand, two entities that have developed a connection or link between them have good probability to have similar attributes.

One of the most compelling techniques of Social Network Analysis is the Statistical Relational Learning, which have emerged when traditional statistical approaches could not treat relational and complex data. Traditional approaches consider only independent and fully distributed data; therefore, they do not take into account dependent and complex relational structure of data. Furthermore, traditional Inductive Logic Programming (ILP) and relational learning do not model noisy data, hence do not perform in the presence of uncertainty. Nowadays, systems generate real world datasets, which represent multi–relational, heterogonous and semi structured data. Data mining systems deal everyday with noisy data and uncertainty. Hence, statistical relational learning approaches help to overcome these issues. Statistical relational learning can be implemented in different domains. One important task of the SRL is dealing with the recommender systems problem. Recommender systems are personalized information tools that generate recommendations, suggestions and predictions. They are employed in different areas where users might want suggestions in order to select among many various objects. For example, online portals of news provide headlines where the reader can focus. Moreover, in e-commerce perspective, online sites generally display a short list of the products in order for the user to have an overview of what the online shop provides. Subsequently, the user can select the product and explore it in more details either purchasing it.

A web page is linked with the web server, which contains a database with items available from the site. Due to the large number of items in the database, it is deemed appropriate to show the object list, which can be selected by the user. Recommender systems employ various algorithms in order to perform recommendation tasks. The algorithms aggregate items from like-minded users, thus by eliminating the item, which have been selected before and suggest the items, which have not been rated by the user. The most well-known algorithms deployed in recommender systems are collaborative filtering and content-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

based filtering. Collaborative Filtering is concentrated on figuring out similar items; for each user that has rated items, the algorithm finds out the similar items, subsequently it aggregates and suggests them to the user.

This review chapter presents a comprehensive scientific survey from the basic and traditional techniques to the-state-of-the-art of SRL algorithms implemented for collaborative tasks. Authors throughout this chapter aim to disclose the most prominent methods and body of work related to Statistical Relational Learning combined with collaborative filtering techniques. These approaches combined together intend to alleviate the shortcomings of previous techniques. It is claimed that the research work can serve as a helpful roadmap for further researches and practice area.

The research chapter is structured as follows. First, it is provided an introduction to Statistical Relational Learning approaches (SRL). Along the chapter, the recommender systems are examined by focusing specifically at one of the most popular recommender system method: the collaborating filtering. In addition, authors have deepened their survey with the most significant research works that have combined SRL with Collaborating Filtering (CF). Finally, authors summarize and identify open issues in this area of research while providing significant findings for the aforementioned collaborative filtering systems.

BACKGROUND

Statistical Relational Learning

Statistical Relational Learning (SRL) is a new branch of machine learning for representing and learning a joint distribution over relational data (Getoor & Taskar, 2007). Getoor and Mihalkova (2011) have defined that Statistical Relational Learning are: "Methods that combine expressive knowledge representation formalisms such as relational and first-order logic with principled probabilistic and statistical approaches to inference and learning". There are some strong reasons why Statistical Relational Learning approaches have experienced a dramatic success and have become an emerging research area in the research community. SRLs constitute an important role in data mining. These algorithms are introduced with a set of candidate features and a model of selection process. The process of feature is decupled from modeling. Usually it can be performed manually. This comprises the most important challenge to the fully integrated application in real world practice where data is recorded in a relational way. Therefore, it is of paramount importance to employ statistical modeling techniques to explore richer data structures and to identify new and complex sources of evidence.

Statistical methods can effectively handle large quantities of noisy data. SRLs are widely exploited from the research community; this fact has given rise to a wide variety of different formalisms, models and probabilistic programming languages. Probability logic-based formalism determines probabilities in two kind of different methodologies: direct and indirect techniques. In the direct technique, probabilities are provided clearly for each probabilistic fact. In the indirect approach, formalisms are not associated to their probability. In this case, the probability is defined from an associated real-valued parameter.

As learning is a fundamental part of the SRL formalism-based system, authors have deemed appropriate to introduce herein two major problems that might be taken into account: parameter estimation and structure learning (Esposito et al., 2009). When the observed data are obtained from an unknown distribution or behavior, the aim is to make sense of this novel distribution. Usually both structure and model parameters must be learned. The parametric SRL approach concentrates on probabilistic model

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

with many finitely parameters by selecting a single model that outperforms the others. There are other nonparametric techniques, which employ probabilistic models with infinitely many parameters and multi-relational process models.

The major difference between traditional statistical approaches and statistical relational learning techniques is the fact that SRL deals with relational data while traditional models deal with propositional data. Therefore, data can be categorized as follows:

- **Propositional Data:** the data mining approaches systems mostly have focused on propositional data and rather in relational data. Propositional data (Khosravi & Bina, 2010) constitutes identically structured entities that are supposed to be independent.
- **Relational Data:** comprise different types of entities and each entity is distinguished by different attributes sets. They are complex and reflect better our real world large-scale systems where data are modeled in multiple related databases.

State-of-the-Art in Statistical Relational Learning

Probabilistic Graphical Models (PGMs) comprise joint probability distributions throughout graph tools and encode a complex distribution over high dimensional space (Chulyadyo & Leray, 2013). Regarding Probabilistic Graphical Models, nodes define the random variables in domain while edges define the probabilistic relationship among variables. Bayesian Networks and Markov Networks are two different classes of PGMs. While Bayesian networks incorporate directed acyclic graph and represent the joint probability distribution as a product of conditional probabilities, Markov Networks employ undirected graph by representing the joint probability distribution as a product of potentials.

Bayesian Networks (BNs)

The structure of a Bayesian Network (BN) (Kersting & De Raedt, 2007) is composed of directed acyclic graph and its nodes constitute a set of random variables. A Bayesian Network has a set of parameter values that determine the probability distribution of each node based on a conditional probability. The conditional probabilities are defined in a conditional probability table (Khosravi & Bina, 2010) for each variable. The assignment's joint probability is computed by multiplying the conditional probabilities of each node value assignment when it is provided the value assigment of its parent. Bayesian Logic Programming (BLPs) are models, which underlie the Bayesian Networks. They incorporate logic programmming. BLPs mitigate the propositional character of BNs, therefore, they are treated as one of the most succcesful model employed in SRL. The major principle of BLPs is to set one to one mapping between ground atoms in the least Herbrand model together with the random variables (Kersting & De Raedt, 2007). The least Herbrand model and its direct relation of influence embody the Bayesian Network. BPLs derive the same benefit of the Bayesian networks and the advantage of definite clause logic. BLPs can facilitate the graphical representation due to the strict separation. In addition, they can easily model any kind of Bayesian Network including those with pure Prolog program.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Probabilistic Relational Models (PRM)

Bayesian Networks have been very successful due to their simplicity. However, one of their major limitations is having difficulty in establishing and maintaining a model in large-scale systems. Complex systems are mostly made of repetitive patterns and they can be viewed in terms of object and relations. Bayesian model lacks the concept of objects and relations; in addition, it cannot model relational data. Hence, Bayesian model is intended to work with single tables of independent and identically distributed data. The task of converting relational real world data into flat propositional data may experience statistical skew and loss of significant data. In order to learn statistical models, PRMs were emerged as substitution of BNs for SRL.

Probabilistic Relational Models represent the first successful model belonging to the group of Statistical Relational Learning (SRLs). They represent a combination of frame-based systems with Bayesian Networks (Friedman et al., 1999). PRMs represent an extension of Bayesian Networks aiming to handle and explore relational data.

Markov Logic Networks (MLNs)

Markov Logic Networks have been one of the most successful techniques attempted to perform statistical relational learning. From the syntax perspective, they widen the first-order logic and establish weight for every formula. On the other hand, from the semantic point of view, they embody a probability distribution for possible worlds by incorporating formulas and their respective weights.

First-order knowledge base (KB) is a set of sentences or formulas in the first order logic (Genesereth & Nilsson, 1987). Formulas are created by incorporating four types of symbols: constants, variables, functions and predicates. A first-order KB is assumed to be a set of hard contraints on the set of possible worlds: if a world breaks just one formula, it is assigned to have zero probability. The motivation behind MLN was to mitigate these contraints. Thus, when a world has broken one formula in the knowledge base first-order, it has low probability, however, it is not impossiple to happen. Formulas are linked with weight in order to reveal the constraint strength.

Domingos and Richardson (2004) attempted to propose Markov logic as a unifying framework due to the need of facilitating knowledge transfer through tasks and approaches. Syntactically speaking, Markov logic can be differentiated from the first-order logic only from the formula to be weighted. In their research work, they exemplify the way certain distinct approaches like probabilistic relational model, knowledge-based construction and stochastic logic programs can be formulated into Markov Logic. Moreover, they have developed algorithms for learning and inference purposes.

Singla and Domingos (2005) have introduced a discriminative approach aiming to train MLN. Their research work intended to optimize the conditional likelihood of query predicates. They have carried out experiments regarding entity resolution and link prediction, which indicate the benefits of their approach if compared to the generative MLN training and even when compared to pure probabilistic models and logical ones.

Wong (2015) has addressed the limitation that Markov Logic Networks pose: the demand for substantial amount of training data. This downside leads to time consuming efforts from the human kind to develop and prepare the required training examples. This derives from the fact that MLN pertain to the supervised approaches. This paper attempts to come up with a semi-supervised mechanism for MLN. Researchers aim to maximize the likelihood function of the observation rather than maximizing the

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

likelihood function of the labeled training data; the latter is broadly used in the MLN supervised learning. Throughout various experiments, this novel approach appears to perform better than the existing frameworks, which exploit labeled training examples.

Relational Markov Networks (RMNs)

A Relational Markov Network identifies a conditional distribution over all entities labels given the relational structure and their content attributes (Taskar et al, 2002). It indicates the cliques, potentials and atributes of entities that are related. Hence, a single model specifies a logic and a relational distribution for any collection of instance. A relational clique template consists of cliques constructed in an instantiation. RMNs are mostly spread to many relational field of interest specifically in social networks. They give extensive information regarding people relations and certain links. Relational Markov Networks incorporate conjuctive database queries dubbed as clique template. Every possible state of cliques must have a feature aimed to reduce the complexity of dependecies. Markov logic undertakes first-order logic in order to describe features. Markov logic allows the presence of uncertainty against arbitrary relation and thus making RMNs more general. Generally, RMNs have been trained in a discriminative way.

An interesting piece of work regarding Relational Markov Network is proposed in (Taskar et al, 2002). Their work introduces a technique that establishes Markov Networks by addressing the shortcomings of undirected models. Undirected models do not impose acyclicity constraints and their limitations are well approppriate for discriminative training. Researchers have enhanced the conditional likelihood of the labels when features are provided. This work gives strong evidence of how to make use of approximate probabilistic inference for the collective classification task of entities related in multiple way with each other. Finally, they show that classification accuracy can be enhanced through incorporating these algorithms.

Assis Costa and Oliveira (2014) have tackled the problem integrating different data set in Linked Data Cloud, which comprises challenges due to incomplete and incosistent data containing outliers. Therefore, they have presented a novel method relying on the relational learning algorithms by incorporating statistical approximation techniques. They have viewed the problem as a relational machine learning task to be solved. This approach is founded in a sparse structure that has the significant capability to scale effectively large datasets.

RECOMMENDER SYSTEMS

A recommender system consists of a personalized information agent, which provides recommendations and suggestions for objects and items that one user might prefer (Burke, 2007). The difference between the recommender systems and information retrieval systems stands in the semantics of user interactions. A result from a recommender system is taken as a suggestion while in information retrieval systems the result is presented as the best match to the respective query. A recommender system varies from an information retrieval system even in terms of personalization and agent (Burke, 2007). A recommender system customizes its responses to a particular user according to user preferences. Rather than providing matches to user queries, it aims to embody an information agent. One compelling property of recommender systems is to alleviate the information overload by suggesting personalized recommendation for

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

each user due to user's previous likes and dislikes. There are many approaches proposed for recommendation systems. The recommendation method can be distingueshed based on the knowledge of the source. Previous research works have categorized recommender systems techniques as below (Burke, 2002):

- **Collaborative Techniques:** The recommender system operates by providing suggestions only via rating profiles of various users; they generate the items to the like-minded users by locating them with the same preference profile.
- **Content-Based Techniques:** The recommender system recommends an item to a user based upon a description of the item and the interest profile of the user (Pazzani & Billsus, 2007).
- **Knowledge Based:** The recommender system is based on knowledge, proposes items based on the inference about the preference of what a user needs (Burke, 2002).
- **Demographic Techniques:** This type of recommender systems works by generating recommendations through the demographic user profile. These methods have the capacity of indicating cross-genre niches and can withdraw the user to overcome and to try new items or products.

All of the four aforementioned techniques have their advantages and shortcomings; collaborative recommender systems have their own limitations, too. The most important challenges that collaborative techniques face when generating recommendations are:

- **New-User Problem:** In order to make sure that the generated recommendation is relevant to users, the system must learn the preference of user from the rating data that user generates by rating various items (Adomavicius & Tuzhilin, 2005). When the user is new, he does not have a rating preference, and the system is not able to recommend any items. This problem can be tackled by using strategies based on item popularity, item entropy, user personalization, etc (Rashid et al, 2002).
- **New-Item Problem:** This happens when an item cannot be recommended when no user has rated it before. Generally, new items face difficulties in being recommended to users. Collaborative systems only make use of the user's preference to develop recommendations. Hence, until a considerable number of users can rate the new item, the recommender system cannot develop a recommendation.
- **Sparsity:** Generally, the number of ratings obtained from users is much smaller than the ratings that must be predicted from the system. Recommender Systems do not perform well when the rating data of product is sparse. Mostly, users lack in rating items. Therefore, the matrix containing the user rating for each item is generally sparse. Hence, the systems have difficulty on finding out a set of users with similar rating. This challenge is of paramount importance when the system is in the first step of usage. It is crucial for systems to effectively predict rating from small number of samples.

Zhu et al., (2015) introduced an approach, which deals with recommender top-N queries in the relational databases. This methodology yields the top-N results for queries conducted by various users. Afterwards, the mechanism determines the linked keywords by computing respective weights of the related keyword. Thus, the approach proposed in this paper can be able to figure out new query words, which are connected to the prior queries and therefore can come up with a recommender query. The associated information of tuples is stored in a knowledge base database in order to evaluate and to

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

examine the recommender query. The advantage of this approach relies on the improving the classic ranking strategy in Information retrieval. The knowledge base, associated interchangeably as the index is the key component of this novel approach.

State-of-the-Art in Collaborative Filtering

The concept of collaborating filtering has derived from the real world domain. Recommendations from other users are highly considered by users. They base their behavior by making choices through word of mouth, reference letters, news, media, public opinion, surveys, guide, etc. Recommender systems facilitate and alleviate people digging through a large amount of information to find the most appropriate and useful information. Recommender systems can provide personalized recommendations through different types of algorithms, including content based filtering, demographic filtering, collaborative filtering, and hybrid filtering methods (Bobadilla et al., 2012). In content-based filtering systems, recommendations are provided to users based on the information related to items that other users have rated previously (Bobadilla et al., 2012). This technique is quite useful for making recommendations for instance that of an adventure movie based on other users who have bought similar category of movies. Demographic filtering techniques on the other hand make recommendations to users taking the positive ratings made by other users who have shared demographic information such as age, gender, qualifications etc. This might help provide recommendations to users based on demographic information and could help provide culturally appropriate information. Collaborative filtering methods allow providing recommendations to users based on how other users have rated items. Hybrid filtering methods allow recommender systems to provide recommendations through techniques based on both collaborative and demographic filtering methods (Bobadilla et al., 2012).

One of the most prominent approaches of recommender systems is collaborative filtering (CF), which uses the known preferences of several users in order to build predictions or valuable recommendations for other novel users. Collaborative filtering consists of various people collaborating in order to help each other to conduct filtering by taking notes of their reactions. The reactions are dubbed as annotations. In other words, collaborative filtering facilitates individuals in the decision making process based on the other people point of view. This method focuses on learning predictive models of user preferences, interests, attitudes, or various behaviors from community data.

The first recommender system dubbed as Tapestry (Goldberg et al., 1992) has introduced for the first time the keyword “collaborative filtering (CF)”. This phrase has been vastly used from the research community even though recommenders may not specifically collaborate with other preferences. The basic assumption of collaborating filtering is that if the user X and Y rate “n” items in the same way or have related preferences, they will behave alike even for new activities or items. CF consists of a user preferences database to predict further products or items that a new user might like.

Collaborating filtering incorporates many challenges that the research community must address. CF approaches are required to have the capability of managing highly sparse data, to scale with the large amount of users and items, to provide the appropriate and valuable recommendation and prediction in a short period. In contrast, content-based algorithms do not make use of any information or details about the items; items are handled as abstract entities. Therefore, collaborating filtering cannot predict items that have not been previously rated. On the other hand, content-based systems do not experience the issue dubbed as the cold-start problem because they benefit from the intrinsic item attributes. While collaborative filtering can develop recommendation to any pair of items as given an amount of sufficient

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

training data, content-based algorithms treat items with similar attribute as similar. Collaborative Filtering techniques are categorized in three kinds of major techniques, namely, memory-based collaborative filtering, model-based collaborative filtering, and hybrid collaborative filtering (Su & Khoshgoftaar, 2009).

Memory-based collaborative filtering employs the whole item-database of user for generating a prediction. Every user is part of a group of people having alike behaviors. A prediction can be elaborated by the system when identifying the neighbors of individuals of a novel user. Memory-based algorithms foretell the rating of items obtained by a user through aggregating the rating of other items generated by the same user or aggregate the other users' rating for the same item. These techniques make use of a similarity metric to figure out several fixed number of items rated by people that are more comparable to the specific item explored by the recommendation system. The most popular similarity metrics aiming to rate vectors are the Pearson correlation coefficient and the cosine-based similarity metric (Adomavicius & Tuzhilin, 2005). Memory-based CF algorithms are easy to implement and perform well for highly dense datasets. Major challenges of this technique are the dependence on user-ratings and their performance deterioration on sparse data. Another memory-based shortcoming is being heuristics-based; they entail extensive hand-tuning upon novel datasets. Moreover, they demand more memory than model-based techniques because of their requirement of accessing the whole dataset when developing recommendations.

Model-based collaborative filtering has been investigated to deal with the drawbacks that memory-based incorporates. The development of machine learning models can make the system learn to distinguish complex patterns by training the system and being able to intelligently predict items for collaborating filtering purposes. Model-based methods develop recommendations for rating by using parametric or semi-parametric model and training on the rating data.

Hybrid collaborative filtering combines CF with other recommender systems approaches just like content-based filtering systems to overcome the shortcoming of pure systems. Therefore, they enhance the performance and the prediction process. They are based on the external content that generally is not available and has increased complexity.

Herein the chapter, authors have explored the most significant pure and hybrid techniques of collaborating filtering. Employing both collaborating filtering and content-based techniques can help in alleviating many recommender system problems, rather than employing only one of respective methods.

Some researchers have pointed out that the first recommender system was Grundy (Rich, 1979), which advocated the utilization of stereotypes to build user models relying on limited information of each indvidual. These user models were exploited for recommending related books to each user and the recommendation appeared to be effective. The definition, collaborative filtering was used for the first time by Goldberg et al. (1992). They have founded an experimental mail system dubbed Tapestry developed at the Xerox Palo Alto Research Center. They assumed that filtering of information can be more effective when human beings are included in the process of filtering. Tapestry was based in both content-based and collaborative methods. It aimed to deal with any incoming stream of electronic documents. It was a framework composed of an indexer, document store, annotation store, filterer, little box, remailer, appraiser, and reader/browser. The reason why researchers developed tapestry relies on the dramatic growth of electronic mail usage. Users were faced with large incoming documents. In order to exploit this bundle of documents, mailing lists were created. The mailing list orientates users to select only those lists, which might have been of their interest. This filtering approach could be further ameliorated when users specified a filter that scans all lists. The major principle that Tapestry constitutes is that filtering can be more effective when humans engage in the filtering process.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

GroupLens (Resnick et.al, 1994) and Video Recommender (Hill et al, 1995) were the first recommender systems, which have incorporated collaborative filtering algorithms intended to automate predictions. The method proposed in Aggarwal et al. (1999) is based on the graph-theoretic approach and can estimate the nearest neighbor of user. There are works that have used person correlation coefficient to measure the similarity (Resnick et.al., 1994), however there are other cosine-based approaches (Breese et al., 1998) where the two users are considered as two vectors in a multi-dimensional space. The similarity between users can be estimated by calculating the cosine angle of two vectors.

Billsus and Pazzani (1998) have introduced a collaborative filtering technique where several machine learning methods just like Artificial Neural Networks are combined with feature extraction techniques. This technique outperformed the memory-based in terms of prediction accuracy and relevance. In Sarwar et al. (2001), the authors pointed out that item-based techniques can have better computation performance compared to traditional user-based collaborative methods. In addition, they appear to perform in comparable or better quality than best of user-based algorithms.

Many several algorithms have tackled the new user problem. Mainly they are based in the hybrid recommender systems. An interesting approach is introduced by Rashid et al. (2002), where various techniques are investigated to decide which is the best and relevant item for a new user. In order to avoid the new user problem, these approaches have made use of different scenarios like item entropy or user personalization.

In regards to the sparsity problem that collaborative techniques face, many significant methods have been explored. In order to mitigate the sparsity rating problem the collaborating system can utilize the information of the user profile to calculate user similarity. Pazzani (1999) have included gender, age, area code, education and employment information to develop effective recommendation and to enhance the problem of sparsity. In this case, this methodology of using not only user-rating, but other demographic information is dubbed as demographic filtering. The sparsity problem have been tackled from Huang et.al. (2004), too. They have mitigated the sparsity issue by carrying out the associative retrieval framework and related algorithms to explore transitive associations through their past feedback.

A significant piece of work was presented in Pennock et al. (2000). They have proposed a novel method named as Personality Diagnosis (PD). In this research work the probability whether users have the same "personality type" when having the same user preference profile for some specific items was calculated. The recommender system can make further predictions, if users like the new arrived items. The PD approach comprises some traditional similarity-weighting techniques where novel data can be included easily. Subsequently, personality diagnosis approach has a significant probabilistic interpretation, which may be utilized in order to analyze and interpret resulting predictions. They have undertaken the experimental phase by providing empirical results on the EachMovie dataset of user-ratings. Data of user profile have been collected by CiteSeer digital library. The Collaborating technique called PD can be viewed as hybrid method of memory-based and model-based algorithms. It has inherited the advantages of memory-based approaches by being fairy straight forward and recording all the data. In addition, it does not demand a compilation step in order to include new data. On the other hand, like other model-based approaches, its predictions are precise and the recommendations have a consequential interpretation. Every user's preference is treated as manifestation of their personality type. The personality type is materialized in a vector of the user's ratings in the database. After the system has obtained the active user's known ratings, authors have calculated the probability of having the same personality with every other user. PD makes better predictions than four other algorithms respectively two memory-based and two model-based approaches.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

An interesting approach regarding hybrid collaborative filtering was proposed in (Melville et al., 2002). They have introduced an elegant and effective approach in which was combined component of content and collaborative techniques. It was incorporated a content-based predictor in order to improve the existing user data and to generate personalized predictions via collaborative filtering. Their content-boosted collaborative filtering approach have given evidence to outperform the pure content-based predictor, the pure collaborative filter and even a naive hybrid approach. In order to mitigate the drawbacks of collaborating filtering algorithms, it was exploited content information of the items already rated. The content-based prediction is employed in order to convert the sparse matrix of user-ratings into a complete matrix. The framework was carried out in a domain of movie recommendation. They have utilized the user-movie rating from the EachMovie dataset (available from the Compaq Systems Research Center). The dataset is composed of a user-ratings data with the rating profile of each user and each movie. Regarding the methodology framework carried out in this research paper, they have compared their hybrid approach with a pure-based predictor and a naive hybrid approach. In order to perform comparison, it was incorporated a subset of rating data from the EachMovie dataset. Recommendation were developed for the withdrawn movie using the different predictors. The quality of the different recommendation algorithms was evaluated throughout the comparison of the predicted values for the withheld rating to the current ratings. They have addressed the sparsity problem by using the pseudo rating matrix, which embody a full user-rating matrix containing ratings for all movies. Therefore, it was aroused the probability of finding similar users with similar preference profile. Regarding the first-rater problem, the recommendation was developed using a content-based predictor for the user. Another body of work in the area of collaborative filtering was proposed by Canny (2002). This article has presented a novel collaborative method, which is intended to protect the privacy of user data. It has supported the Probabilistic Factor-analysis Model (PFM). In this approach the protection of data is carried out by a peer to peer protocol. Moreover, it deals with missing data and does not demand their default value. The novel approach exhibits advantages in speed and storage when compared to the preceding works. There are some significant contributions that this research paper has provided to the body of knowledge. First, this paper has presented the privacy-preserving scheme and explained the protocol generally. The second contribution consists of the new collaborative filtering that appears to be the most accurate CF comparing to the previous one, to their best knowledge. And the third is conducting testing and experiments in order to evaluate the accuracy, speed and storage.

Marlin et al. (2011) discussed the key problem in rating-based collaborative filtering: the possibility of a fundamental incompatibility among the features of recommender data sets and the supposition of valid estimation of statistical model experiencing missing data. In this research paper, it is explained and explored in depth the missing data process in the recommender systems and their influence in the standard statistical model estimation. The paper described components of recommender datasets linking them to the statistical theory of the model estimation. Authors have proposed a framework together with a set of modified evaluation protocols, which handle non-random missing data. Rating prediction provided by this paper, give strong evidence that their developed model can perform better than the standard models.

Big data have emerged substantially in our technology-based area, therefore some gaps and issues have emerged in the recommender systems. Liu & Li (2015) have addressed these issues by designing parallel recommendation algorithms, which constitute an enhanced parallel item-based collaborative filtering founded on Hadoop. Studies according to parallel recommender systems have not been explored in depth so far. The novel parallel item-based collaborative filtering introduced here aims to enhance the

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

performance of the recommendation due to introducing soft punishment on active users. The user-rating model is considered as a prediction model. Authors provide strong evidence that this novel parallel approach outperforms the previous algorithms in terms of speed and scalability.

In closing of Collaborating Filtering section, we must emphasize that because artificial data are usually not reliable due to the properties of CF tasks, it is more suitable to undertake live experiments with real world dataset.

STATISTICAL RELATIONAL LEARNING FOR COLLABORATIVE FILTERING

In this section, this research chapter will disclose herein the most prominent approaches, which combine various statistical relational learning methods in the collaborating filtering domain.

Latent structure of data has been taken into account in order to predict links that have been cumbersome to be analyzed from existing approaches. Hence, machine-learning research communities have examined the latent class model in combination with Bayesian nonparametric methods. Blei et al. (2003) discussed Latent Dirichlet Allocation (LDA), which constitutes a generative probabilistic model for collections of discrete data such as text corpora. LDA belongs to the three-level hierarchical Bayesian model where each item is modeled as a finite mixture over a corresponding set of topics. They introduce an efficient approximate inference technique based on variation methods and EM algorithm for empirical Bayesian parameter estimation. This approach is undertaken in document modeling, text classification and collaborative filtering. This chapter is interested to review collaborative filtering with LDA. Throughout their experimental phase, authors have made use of EachMovie collaborative filtering data. Regarding the collaborative filtering perspective, the model is trained on fully observed set of users. Then for each unobserved user, all movies are displayed except of the user's preferred one. The various algorithms are assessed due to the possibility that they assign to held-out movie. Furthermore, with a 1600 movies vocabulary, authors have found predictive perplexities, however the LDA model provides the best perplexities.

Miller et al. (2009) discussed a latent variable approach making use of Bayesian nonparametric techniques to infer the numbers of features at the same time when learning the respective entities with its corresponding feature. Each entity of the Bayesian nonparametric model has binary valued latent features that affect the relations depending on a set of known covariates. The authors claimed that their approach enables to simultaneously infer the number of the latent features and at the same time to infer the way these features influence observations. Their framework joins inferred features with known covariates aiming to predict the link prediction task. Previous existing class-based frameworks infer latent structure; consequently, proposed model is deemed more expressive than traditional methods and can utilize the output of previous approaches for initialization of its work. This chapter provides insightful evidence that their novel approach can perform well regarding link prediction tasks upon various dataset even on dataset that used to argue for class-based models. This approach exhibits expressiveness, therefore this lead to an enhancement of its performance in the datasets. Moreover, its advantage relies on its richer representations, which enable to mine patterns of interaction compared to class-based models.

Traditional collaborative filtering methods are based on the postulation that real-world data are stationary. Motivated by the sale prediction problem, Xiong et al (2010) introduced to the research community a factor-based algorithm that has the capability to take into consideration the time component when developing predictions. This paper has treated tensor factorization problem with a specific time dimen-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

sion constraint. In addition, it conducted a fully baysian treatment in order to attain automatic model complexity control. Sampling procedure are performed in order for the system to learn the model. The novel approach dubbed as Bayesian Probabilistic Tensor Factorization (BPTF) is tested in real world data-ratings. Real relational data are in process of evolving and experience strong temporal patterns. This proposed framework is able to model time evolving relational data. This paper has tackled the overfitting problem by extending the approach with Bayesian algorithms. In regard to scalability function, authors have incorporated Markov Chain Monte Carlo (MCMC) process to perform the learning task.

Another domain where SRLs for collaborative filtering have been implemented is the Customer Relationship Management (CRM) field of interest. Customer modeling in CRM can be defined as a special case of a relational learning problem. Perlich and Hung (2010) have presented a connection between relational learning domain and CRM analysis particularly in customer classification and product recommendation. This work has been evaluated in seven-real world CRM datasets where empirical results have been gathered. This paper illustrated that relational learning techniques can be useful in addressing CRM modeling tasks. Authors have focused on predictive customer modeling underlying classification and probability estimation for cost-sensitive, decision-making and product recommendation. Relational learning can explore large-scale set of models, thus having the capability to capture predictable dependencies. This paper has modeled CRM with statistical relational learning techniques for product recommendation task. This tool was not meant to model all CRM, but it focused on performing exploration and capturing relevant information. The illustrated example in this paper was the bookstore dataset. They have utilized this dataset to examine the relational learning capability regarding the recommendation aspect. It is implemented the unified recommendation framework based on the extension of a relational learning, known as Probabilistic Relational Model (PRMs). In order to address the recommendation tasks, they have included a special existence attribute into the order table and obtained dependency frameworks relevant to order table attribute. The performance is evaluated with well-studied metrics including precision, recall F-measure and rank score.

Xu et al. (2010) have investigated complex social networks with Infinite Hidden Relational Models (IHRMs) for tasks of community detection, link prediction and product recommendation. A social network based IHRM employs edges associated with random variables. Probabilistic dependencies between these variables are determined based on the relational structure model. The hidden variables can transfer information in order to obtain non-local probabilistic dependencies. This model can serve for predicting entity attributes. In addition, it conducts a significant cluster analysis.

Statistical relational learning in the domain of collaborative filtering has been employed also to tackle the learning of trust. Rettinger et al. (2011) have pointed out that from a trustee perspective; trust is described as many relations happening among the agent to be trusted and the state of environment. They exemplify how to apply and learn context sensitive trust by employing SRL in the form of Dirichlet process model dubbed as Infinite Hidden Relational Trust Model (IHRTM). Several experiments are carried out in the user-ratings data from the eBay platform. Authors have demonstrated with empirical results that the trustee is enabled to specify the structure of the trust-situation by the inherent clustering process. Authors have highlighted that the employment of collaborative filtering combined with relational data can lead to an improvement of the trust assessment performance. Moreover, one other contribution of this framework is to enhance the cold start performance due to the learning and transferring knowledge faster. Dealing better with dynamic behavior in an open multi-agent system is another benefit that this approach provides, which is illustrated by records interaction generated from a scenario composed of two players.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

A method for modeling large-scale multi-relational dataset was discussed in Jenatton et al. (2012). Their framework represents a bilinear structure capturing several orders of data interaction. Furthermore, it shares latent factors through various relations. Authors have employed their approach on a standard tensor-factorization dataset. Their application across various experiments provides evidence that can efficiently learn semantically meaningful representations. The model for relational data is appropriate to be deployed in multi-relational graphs and in natural languages. Regarding the algorithmic approach, in order to speed up the optimization, they have utilized several scenarios. They have claimed that cannot treat the NLP application as a standard tensor factorization problem. The proposed framework is deemed probabilistic and can be a benefit for the accounting domain because of the uncertainty presence of data. The paper empirically has demonstrated that this approach outperforms the previous existing frameworks.

One of the most well-known SRL approaches, Probabilistic Relational Model (PRMs) have been employed for collaborative filtering in (Newton & Greiner, 2014). For experimental purposes, they have exploited the EachMovie dataset. First, they have learned standard PRM, then they have specified the hierarchical PRMs, which comprise an extension of standard PRM throughout refining classes into hierarchies. It is shown that this approach allows greater presence of context sensitivity compared to standard PRM, thus being more expressive. This piece of work provides strong evidence that PRMs can be effectively employed in the collaborative filtering area. They have discussed the performance of standard PRMs deployed in the EachMovie dataset and then compared it with the effectiveness of hierarchical PRMs (Getoor, 2002).

Chulyadyo and Leray (2014) proposed a novel method to establish a personalized recommendation model based on PRM incorporating user's preferences. The paper demonstrates that carrying out PRM for collaborative filtering can mitigate the cold start problem and thus making able to generate personalized recommendations. This work has emerged from the need of intelligent recommender system in professional community and in large-scale systems where the system does not have the user profile, which is composed of incomplete and inconsistent data. This work intends to serve as a help for users by showing their search queries listed by relevance. Previous techniques could not address the problem of real world applications that demanded personalized recommendation, therefore, lacking the user profile existence.

Yuan et al. (2015) have examined the importance of retrieving the contextual information of users. Contextual information stands for the time and the locations of various users. These key terms are applied in the micro-blogging services and location-based social networks domain. In this paper, authors deal with contextual recommendation tasks and contextual search. User mobility behavior is associated with four major questions: who visits, which place, what time and for what activity. Thus, addressing these topical issues demand for effective methods in examining the mobility behavior. The underlying approach introduces a nonparametric Bayesian model, which does not require any parameter tuning. When given one out of the four aspects, the mechanism itself can provide the remaining others. A variety of experiments demonstrates that the novel model is truly effective in unveiling user's spatial temporal topic. In this paper it is implemented a novel probabilistic generative model based on non-parametric Bayesian in order to represent the mobility behavior of users in terms of spatial, temporal and activity. The key contribution of this new model is to disclose spatial temporal matters for each user by making recommendations aware in terms of context and search.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

SUMMARY

This chapter provided a comprehensive overview of the most compelling research papers that combine statistical relational learning for collaborative filtering tasks. Authors have conceptualized this research work by exploring first statistical relational learning and all the techniques that it comprises. Moreover, the chapter has explored the recommendation systems, its categorizations in three major approaches collaborative-based, content-based, knowledge-based, and demographic-based. Regarding collaborative techniques, the chapter has provided a profound classification, which include memory-based, model-based, and hybrid approaches. Authors have claimed that model-based techniques outperform memory-based. However, it is demonstrated in various previous researches that the hybridization of the two techniques can alleviate some limitation that each method imposes if they are employed purely. SRL have been combined for collaborative filtering purposes in recommender systems particularly for the product recommendation tasks. The most popular approaches for relational data are the PRM (an extension of Bayesian traditional Networks) and Rational Markov Networks (RMN).

In this survey work are introduced three tables, which summarize and categorize the whole research work that has been conducted throughout research chapter. Respectively, Table 1 describes the categorization of the statistical relational learning frameworks, Table 2 exhibits the classification and the state-of-the-art in collaborative filtering techniques, and finally the Table 3 refers to an overview of the research examples, which have combined techniques of statistical relational learning with collaborative filtering.

CONCLUSION AND FURTHER WORK

Collaborating Filtering incorporates many challenges that the research community must address. CF approaches are required to have the capability to manage highly sparse data, to scale with the large amount of users and items, to provide the appropriate and valuable recommendation and prediction in a short period, etc. Throughout the survey of SRL for CF approaches, it is demonstrated that carrying out PRM for collaborative filtering can mitigate the cold start problem and thus making able of generating personalized recommendations. Previous research works have addressed the sparsity problem by using the pseudo rating matrix, which enbody a full user-rating matrix containing ratings for all movies.

Table 1. Categorization of statistical relational learning techniques

	Types of Statistical Relational Learning Techniques		
	PRM	**MLN**	**BLN**
Class level model	Directed GM	Logical clauses	Bipartite directed
Parameter estimation	Machine Learning to fill CPT	Machine learning to Weights	Machine Learning to fill CPT
Structure learning	Score Based learning	ILP methods	ILP methods
Inference Graph	Bayesian network	Markov model	Bayesian network
Autocorrelation	Self-loops in class level model	Additional variables needed	It is not discussed in the research this feature

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Classification of collaborative filtering techniques

	Heuristic based	Model Based	Hybrid-based combining content-based with collaborative filtering
Collaborative Filtering	Commonly used techniques -Nearest neighbor -Clustering -Graph theory -Representative Corresponding research works; -(Rich, 1979), - Goldberg et al. (1992) -Resnick et al. (1994) -Hill et al. (1995) -Shardanand & Maes 1995 -Aggarwal et al. 1999 -Pennock & Horwits (2000) -Sarwar et al. (2001)	Commonly used techniques -Bayesian -Clustering -Artificial Neural Networks -Linear regression -Probabilistic models Corresponding research works; -Goldberg at al. (1992) -Breese et al. (1998) -Pennock & Horwits (2000) Huang et.al. (2004), -Pazzani and Billsus (2007) -Yu et al. (2002) - Canny (2002). -Marlin et al. 2011 - Liu & Li (2015)	-Linear combination of predicted ratings -Various voting schemes -Incorporating one component as a part of the heuristic for the other Corresponding research works; -Pazzani (1999) -Billsus & Pazzani (2000) -Mellville et al. (2002) - Rashid et al. (2002)

Table 3. Overview of statistical relational learning for collaborative filtering

	PRM for CF	MLN for CF	BLN for CF
SRL for Collaborative Filtering	Representative research examples: - Getoor (2002) -Perlich and Hung (2010) - Jenatton et al. (2012) -Newton & Greiner (2014) -Chulyadyo and Leray (2014)	Representative research examples: - Xu et al. (2010) - Rettinger et al. (2011)	Representative research examples: - Blei et al. (2003) -Miller et al. (2009) - Xiong et al (2010) - Yuan et al. (2015)

Sparsity rating problem can be addressed even by utilizing the information of the user profile to calculate user similarity. It is included gender, age, area code, education and employment information in order to develop effective recommendations and to overcome the problem of sparsity. In this case, this methodology of using not only user-ratings, but other demographic information is dubbed as demographic filtering. Regarding the first-rater problem, the recommendation must be developed using a content-based predictor for the user.The recommendation process has been further enhanced by generating effective prediction through using content-based prediction even for the other users. Several papers have illustrated that relational learning techniques can be useful in addressing CRM modeling tasks, too. In regards to collaborating filtering, it is emphasized that because artificial data are usually not reliable, it is more suitable to undertake live experiments with real world dataset.

In concluding this chapter, one open issue that can be suggested for future work is regarding the limitation that IHRTM impose on the process of not handling the trustees with strategies. An adaptive learning strategy can be implemented in the future. This is a domain that cannot be addressed by propositional machine learning algorithms; hence, statistical relational model would be appropriate for this problem. Finally, the combination of SRL for CF opens up new doors to many further applications in the future research. Open gaps can be tackled and future refinement of existing methods can be provided in order to advance in this new field of interest.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Adomavicius, G., & Tuzhilin, A. (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Gediminas AdPossible Extensions. *IEEE Transactions on Knowledge and Data Engineering*, *17*(6), 734–749. doi:10.1109/TKDE.2005.99

Aggarwal, C. C., Wolf, J. L., Wu, K. L., & Yu, P. S. (1999). Horting hatches an agg: A new graph-theoretic approach to collaborative. *Proceedings of the Fifth ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining*, ACM Digital Library.

Assis Costa, G., & Oliveira, J. M. (2014). A Relational Learning Approach for Collective Entity Resolution in the Web of Data. *Proceedings of the Fifth International Workshop on Consuming Linked Data*.

Bilsus, D., & Pazzani, M. (1998). Learning Collaborative Information Filters. *Proceedings of the Int'l Conf. Machine Learning*.

Blei, D. M. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*, 993–1022.

Bobadilla, J., Hernando, A., Ortega, F., & Gutiérrez, A. (2012). Collaborative filtering based on significances. *Information Sciences*, *185*(1), 1–17. doi:10.1016/j.ins.2011.09.014

Breese, J., Heckerman, D., & Kadie, C. (1998). Empirical Analysis of predictive algorithms for collaborative filtering. *Proceedings of the 14 th Conference on Uncertainty in Artificial Intelligence* (pp. 1-10).

Burke, R. (2002). Hybrid Recommender Systems: Survey and Experiments. *UMUAI*, *12*, 331–270.

Burke, R. (2007). Hybrid Web Recommender Systems. In The Adaptive Web (pp. 377 – 408). Berlin Heidelberg: Springer Verlag.

Canny, J. (2002). Collaborative Filtering with Privacy via Factor Analysis. *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval SIGIR '02,* Tampere, Finland (pp. 238-245).

Chulyadyo, R., & Leray, P. (2013). *Probabilistic Relational Models for Customer Preference*. Modelling and Recommendation.

Chulyadyoa, R., & Leraya, P. (2014). A personalized recommender system from probabilistic relational. *Proceedings of the 18th International Conference on Knowledge-Based and Intelligent* (pp. 1063 – 1072). Elsevier B.V.

Della Pietra, S., Della Pietra, V., & Lafferty, J. (1997). Inducing features of random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*(4), 380–392. doi:10.1109/34.588021

Domingos, P., & Richardson, M. (2004). Markov Logic: A unifying framework for statistical relational learning. *Proceedings of the ICML-2004 Workshop on Statistical Relational Learning and its Connections to other Fields.*

Esposito, F., Ferilli, T., Basile, M.A., & Di Mauro, N. (2009). Social Networks and Statistical Relational. *Int. J. of Social Network Mining*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Friedman, N., Linial, M., Nachman, I. & Pe'er, D. (1999). Using Bayesian Networks to Analyze. *Journal of Computational Biolog*, 1999, 601–620.

Genesereth, M.R., & Nilsson, N.J. (1987). Logical Foundations of Artificial Intelligence.

Getoor, L., & Taskar, B. (2007). *Introduction to Statistical Relational Learning*. London: The MIT Press.

Getoor, L. (2002). *Learning statistical models from relational* [Doctoral dissertation]. Stanford University.

Getoor, L., & Mihalkova, L. (2011). *Learning Statistical Models*. SIGMOND.

Goldberg, D., Nichols, D. Oki, B. M, & Terry, D. (1992). Using collaborative filtering to weave an information Tapestry. *Communications of the ACM*.

Hill, W., Stead, L., Rosenstein, M., & Furnas, G. (1995). *Recommndeing and Evaluating Choices in a Virtual Community of Use*. Conf. Human Factors in Computing Systems.

Huang, Z., Chen, H., & Zeng, D. (2004). Applying Associative Retrieval Techniques to Alleviate the Sparsity Problem. *ACM Transactions on Information Systems*, *22*(1), 116–142. doi:10.1145/963770.963775

Jenatton, R., Le Roux, N., Bordes, A., & Obozinski, G. (2012). *A latent factor model for highly multi-relational data.* In Advances in Neural Information Processing Systems (p. 19). NIPS.

Kersting, K., & De Raedt, L. (2007). Bayesian Logic Programming: Theory and Tool. In *L. a. Getoor (Ed.), An Introduction to Statistical Relational Learning*. MIT Press.

Khoshneshin, M. (2012). *Latent feature networks for statistical relational learning [PHD]*. Iowa Research Online.

Khosravi, H., & Bina, B. (2010). A Survey on Statistical Relational Learning. In *Advances in Artificial Intelligence* (pp. 256–268).

Liu, Q., & Li, X. (2015). A New Parallel Item-Based Collaborative Filtering Algorithm Based on Hadoop. *Journal of Software*, *10*(4), 416–426. doi:10.17706/jsw.10.4.416-426

Marlin, M. B., Zemel, R., Roweis, S. T., & Slaney, M. (2011). Recommender Systems: Missing Data and Statistical Model Estimation. *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence* (pp. 2686-2691).

Melville, P., Mooney, R. J., & Nagarajan, R. (2002). Content-Boosted Collaborative Filtering for Improved Recommendations. *Proceedings of the Eighteenth National Conference on Artificial Intelligence(AAAI-2002)* (pp. 187-192). Edmonton, Canada: American Association for Artificial Intelligence.

Miller, K. T., Griffiths, T. L., & Jordan, M. I. (2009). Nonparametric Latent Feature Models for Link Prediction. *Proceedings of the 23rd Annual Conference on Neural Information Processing Systems 2009,* Vancouver, Canada.

Neville, J., Rattigan, M., & Jensen, D. (2003). Statistical Relational Learning: Four Claims and a Survey. *Proceedings of the Workshop on Learning Statistical Models from Relational Data, 18th International Joint Conference on Artificial Intelligence* (p. 5).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Newton, J., & Greiner, R. (2004). Hierarchical Probabilistic Relational Models for Collaborative Filtering. *Workshop on Statistical Relational Learning, 21st International Conference on Machine Learning* (pp. 1-6).

Nilashi, M., Bagherifard, K., Ibrahim, O., Alizadeh, H., Nojeem, L. A., & Roozegar, N. (2013). Collaborative filtering recommender systems. *Research Journal of Applied Sciences*, *Engineering and Technology*, *5*(16), 4168–4182.

Pazzani, M. (1999). A Framework for Collaborative, Content-Based, and Demographic Filtering. *Artificial Intelligence Review*, *13*(5/6), 393–408. doi:10.1023/A:1006544522159

Pazzani, M. J., & Billsus, D. (2007). Content-based Recommendation Systems. In The Adaptive Web (pp. pp. 325-341). Berlin Heidelberg: Springer. doi:10.1007/978-3-540-72079-9_10

Pennock, D. M., Horvitz, E., Lawrence, S., & Giles, C. L. (2000). *Collaborative Filtering by Personality Diagnosis: A Hybrid Memory- and* (pp. 473–480). Uncertainty in Artificial Intelligence.

Perlich, C., & Huang, Z. (2010). Relational Learning for Customer Relationship Management. *Information Systems*.

Rashid, A. M., Albert, I., Cosley, D., Lam, S. K., McNee, S. M., Konstan, J. A., & Riedl, J. (2002). Getting to Know You: Learning New User Preferences in Recommender Systems. *Proc. Int'l Conf. Intelligent.* 10.1145/502716.502737

Resnick, P., Iakovou, N., Sushak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: An Open Architecture for Collaborative Filtering of NetNews. *Computer Supported Cooperative Work*.

Rettinger, A. (2009). Statistical Relational Learning with Formal Ontologies. In Machine Learning and Knowledge Discovery in Databases (pp. 286-301). doi:10.1007/978-3-642-04174-7_19

Rettinger, A., Nickles, M., & Tresp, V. (2011). Statistical relational learning of trust. *Machine Learning*, *82*(2), 191–209. doi:10.100710994-010-5211-x

Rich, E. (1979). User Modeling via Stereotypes. *Cognitive Science*, *3*(4), 329–354. doi:10.120715516709cog0304_3

Siguenza-Guzman, L., Saquicela, V., Avila-Ordóñez, E., Vandewalle, J., & Cattrysse, D. (2015). Literature review of data mining applications in academic libraries. *Journal of Academic Librarianship*, *41*(4), 499–510. doi:10.1016/j.acalib.2015.06.007

Sarwar, B., Karypis, G., Konstan, J. & Riedl, J. (2001). Item-Based Collaborative Filtering Recommendation Algorithms. *Proceedings of the 10th Int'l WWW.*

Shardanand, U., & Maes, P. (1995). *Social Information Filtering, Algorithms for Automating 'Word of Mouth*. Conf. Human Factors in Computing Systems.

Singla, P., & Domingos, P. (2005). *Discriminative Training of Markov Logic Networks*. American Association for Artificial Intelligence.

Su, X., & Khoshgoftaar, T. M. (2009). *A Survey of Collaborative Filtering Techniques*. Advances in Artificial Intelligence.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Taskar, B., Abbeel, P., & Koller, D. (2002). Discriminative Probabilistic Models for Relational Data. *Proceedings of the Eighteenth conference on Uncertainty in artificial intelligence UAI'02* (pp. 485-492).

Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge, UK: Cambridge University Press. doi:10.1017/CBO9780511815478

Winkler, W. (1999). *The state of record linkage and current research problems*. U.S: Statistical Research Division, Census Bureau.

Wong, T. (2015). Learning Markov logic networks with limited number of labeled training examples. *International Journal of Knowledge-based and Intelligent Engineering Systems*, *18*(2), 91–98.

Xiong, L., Chen, X., Huang, T., Schneidery, J., & Carbonellz, J. G. (2010). Temporal Collaborative Filtering with Bayesian Probabilistic Tensor. In SIAM Data Mining.

Xu, Z., Tresp, V., Rettinger, A., & Kersting, K. (2010). Social Network Mining with Nonparametric relational Models. *Advance in Social Network Mining An Analysis*, 77-96.

Yuan, Q., Cong, G., Zhao, K., Ma, Z., & Sun, A. (2015). Who, where, when and what: A Nonparametric Bayesian approach to context-aware recommendation and search for twitter users. *ACM Transactions on Information Systems*, *33*(1), 1–33. doi:10.1145/2699667

Zhu, L., Lei, Q., Liu, G., & Liu, F. (2015). Processing Recommender Top-*N* Queries in Relational Databases. *Journal of Software*, *10*(2), 162–171. doi:10.17706/jsw.10.2.162-171

KEY TERMS AND DEFINITIONS

Collaborative Filtering: Collaborative filtering allows recommendations to users taking into account how other users have rated items. Collaborative filtering needs an extensive database containing exclusively the ratings made by users over the items (Bobadilla et al., 2012).

Content-Based Filtering: Content-based filtering systems allow recommender systems to provide recommendations to users based on the information related to items that other users have rated previously.

Data Mining: Data Mining is defined as the process of analyzing large information repositories for deriving and discovering useful patterns and identifying hidden relationships among the data (Siguenza-Guzman et al., 2015).

Hybrid Collaborative Filtering: Hybrid filtering methods allow recommender systems to provide recommendations through techniques based on both collaborative and demographic filtering methods (Bobadilla et al., 2012).

Recommender Systems: Recommender systems are software tools and techniques for suggesting items to users by considering their preferences in an automated fashion (Nilashi et al., 2013).

Social Network Mining (SNM): Its major purpose is to extract information about network objects, behaviors and activities that cannot be yielded if the entities were to be examined individually.

This research was previously published in Collaborative Filtering Using Data Mining and Analysis edited by Vishal Bhatnagar; pages 250-269; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 34
An Artificial Intelligence–Based Vehicular System Simulator

Marvin T. Chan
University of Regina, Canada

Christine Chan
University of Regina, Canada

Jonathan T. Chan
University of Regina, Canada

Craig Gelowitz
University of Regina, Canada

ABSTRACT

This paper presents a vehicular system simulator, which enables the human player to race a car against three system-controlled cars in a three-dimensional road system. The objective of the vehicular system simulator is not to support defeating the opponent in a car race, but to provide the player with a challenging and enjoyable racing experience. Therefore, it is important that the system simulates human driving behavior and adopts cognitive computing. The paper discusses development of the vehicular system simulator using the artificial intelligence (AI) techniques that are supported in the game engine of Unity. The design and implementation of the vehicular system simulator are presented. The discussion includes some possible extensions of the current version of the system so that it can be adapted to be a simulation system for education purposes

1. INTRODUCTION

A vehicular system simulator provides a popular game environment. In a traditional game like chess, a situation is presented where "players engage in an artificial conflict, defined by rules, that results in a quantifiable outcome" Salen & Zimmerman (2004). Such artificial conflicts are often represented as a puzzle or a challenge, and having the puzzle solved or the challenge resolved provides a real-world purpose to the game players (Liu et al., 2011). These games are sometimes referred to as "serious" games (Wood & Reiners, 2012), but they have been increasingly replaced by electronic games, especially for the "game generation" (Prensky, 2001) of "digital natives". This younger generation of game players typically grew up playing games and, in contrast to the "digital immigrants" (Prensky, 2001) of the older

DOI: 10.4018/978-1-7998-0951-7.ch034

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

generation, they are adapt in "dealing with large amounts of information quickly even at the early ages, using alternative ways to get information, and finding solutions to their own problems through new communication paths" (Akilli, 2007).

The Artificial Intelligence (AI) field has witnessed a similar transition from the "classical AI games" such as Samuel's Checker Player (Samuel, 1967) and Waterman's Poker Player (Waterman, 1970) to the contemporary AI techniques adopted in electronic games. The goal of the game is no longer a quantifiable outcome of winning over the opponent as in the case of a checker or poker game. Instead, a contemporary game contains changing environments, multiple objectives, and dynamic aspects of the game that are revealed to the player as the game unfolds. The objective is to offer to the human player an enjoyable experience through his or her interaction with the game, and this does not involve any specific quantifiable outcome.

The vehicular system simulator has been developed as a racing game, called Racer, with the goal of offering to the player a challenging and enjoyable experience in a car race against a system-controlled car. Although the player's goal within the game is to win the race against the system-controlled car, the AI techniques adopted in the game are primarily designed to give the player an enjoyable time racing his or her car. In other words, the objective of winning by either side is not given the highest priority.

This paper proceeds as follows. Section 2 provides some background literature relevant to the work. Section 3 presents some components of the vehicular system simulator. Section 4 presents implementation of the vehicular system; Section 5 discusses some AI techniques adopted in the vehicular system simulator, their effects on system performance, and some exploration on feasibility of extending the current system into an educational software system useful for teaching transportation engineering concepts. Section 6 includes the conclusion and some ideas on future directions of the work.

2. BACKGROUND LITERATURE

Some research studies on cognitive vehicular systems are briefly summarized as follows. Xie & Xu (2015) provided an overview of the algorithms and technologies useful for extracting static properties of vehicles from videos so as to support criminal and traffic violations. With the rapid growth of the video surveillance technology, a large database of video and images has been created. The database can become an important resource for police in performing traffic events analysis: crucial information such as speed and license plate of vehicles can be extracted and examined. Yoshizawa & Iwasaki (2015) examined a type of driving behavior, called "aimless driving", which is responsible for a large number of fatal traffic accidents in Japan. The study investigates this "aimless driving" behavior by giving subjects non-visual secondary tasks of four difficulty levels while they watch pedestrians and their responses to objects in the road on which they travel are measured. The results indicated that even non-visual tasks influence eye movement and the subjects were not able to react well to objects on the road. Xu et al. (2014) studied the use of semantic technologies for annotating videos so as to enhance comprehensibility of the videos. Their work proposed a video annotation platform, which supports users by providing a search interface of annotated video resources. Some preliminary results indicated that applying semantic technologies for annotation of videos can enhance reusability, scalability, and extensibility of the video resource and support increased adoption of this resource for traffic events analysis.

Much research work that focuses on developing AI-controlled components of game systems called "bots" can be found. These "bots" approximate or emulate human game playing styles. Development of the "bots" often has been motivated by the belief that a human player's enjoyment in the gaming

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

experience will be higher if he or she can be led to believe that the opponents in the game are other human players. An example of this type of game is the popular multiplayer first person shooter game of Counter-Strike (Valve Corporation, 2016), in which the objective is to eliminate all players on the enemy or opponent team. This type of "bots" also appears in the early video game of PacMan, in which the computer-controlled ghosts in the maze are able to move towards the player-controlled characters because the former incorporated some path finding algorithms (DeNero & Klein, 2010). Some other AI algorithms commonly implemented in games include finite state machines, fuzzy state machines, and decision trees. Decision trees are often used to represent the decision making process involved in games like Checkers or Chess. Since a decision tree can specify a number of possible game states given a current state, and support a search process for a goal state, this approach is adopted so that the human player is defeated provided the AI-controlled game can identify the path to the goal state more efficiently than the human player.

Some other AI techniques used for building bots so that they would mimic or emulate human behavior can also be found. Khoo and Zubek (2002) adopted a behavior-based action selection technique in the software development kit (SDK) of Flexbot, which was used for developing the AI-controlled bot of Groo. This provided the basis for development of 32 bots, which can run simultaneously on a single machine. The human players who played against the bots believed the bots effectively emulated human behavior (Khoo & Zubek, 2002). Tatai and Gudwin (2003) discussed adopting semiotics in their development of a bot to assist or attack the human player in the game of Counter-Strike; their aim was to design the AI system in such a way that the human player is entertained and challenged and not simply defeated. Hence, the bot's performance is gauged not by its efficiency in killing its opponents, but by its "capability to emulate as close as possible all the activities commonly performed by a human player" (Tatai & Gudwin, 2003). Due to this consideration, the bot cannot be perceived by the human to be waiting because human players expect bots to be continuously working. In other words, the bot's reaction time was one of the key criteria in measuring its performance (Tatai & Gudwin, 2003).

In a vehicular system simulator, the user's expectation is to control the car by making small adjustments to the car's direction while driving because that is how a car is controlled in real life. Therefore, a vehicular system simulator must be able to make continuous adjustments to the car's direction as opposed to simply driving in a straight line and turning only when a curve in the road approaches. To ensure these expectations on human driving behavior can be satisfied, some AI techniques used for designing and implementing the vehicular system simulator will be discussed in the following sections.

3. VEHICULAR SYSTEM COMPONENTS

The vehicular system simulator consists of components that involved AI concepts and techniques; these are presented as follows.

3.1. The Road System

Two considerations were involved in design of the road system in the simulator: (1) how to design the road itself, and (2) how the navigation on the road can be realized. The road system provides the 3D graphical environment on which the cars travel, and the road consists of obstacles, curves, and corners which require careful navigation of the car. To guide the car's navigation on the road, the Unity engine

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. A screenshot of waypoints at key positions on the track



provides the waypoint system, which clearly specifies identifiable points on the road that support a car's navigation on the road. The waypoint system includes a set of waypoints and each waypoint is a coordinate in 3D space that represents a key position on the racetrack. In the vehicular system simulator, the waypoints are stored as a sorted list of positions in 3D space using the List generic class from the System.Collections.Generic namespace (Microsoft, 2016). This list of waypoints is read as input, and then the system iterates through the list until all the waypoints have been passed by the system-controlled car. The system only iterates to the next waypoint in the list when the system-controlled car is detected to be within a specified distance from its current waypoint. Figure 1 shows a screenshot of the waypoints on the track.

To turn the car towards its current waypoint, the system executes a series of vector calculations so as to define the amount of steering the car requires. The calculations are done using an initial vector based on the position of the current waypoint and the current position of the car. The calculations then generate an output vector that forms the basis for setting the steering and braking output levels of the car. However, due to the nonlinear relationship between the input and output vectors, the car could not be controlled using the output vector generated only, and the additional mechanism of the conditional monitoring system was also adopted.

Depending on the values of the initial calculated output vector produced from the waypoint system's vector calculations, the steering and braking output levels applied to the car are adjusted by the conditional monitoring system. When both the waypoint system and the conditional monitoring system were employed for controlling the car, it was able to drive smoothly on the road.

Instead of adding the conditional monitoring system, it was possible to augment the waypoint system with a user-defined nonlinear mathematical function so as to better control the car. However, this option was discarded because it may require complex mathematical manipulations.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

3.2. Artificial Environment Perception

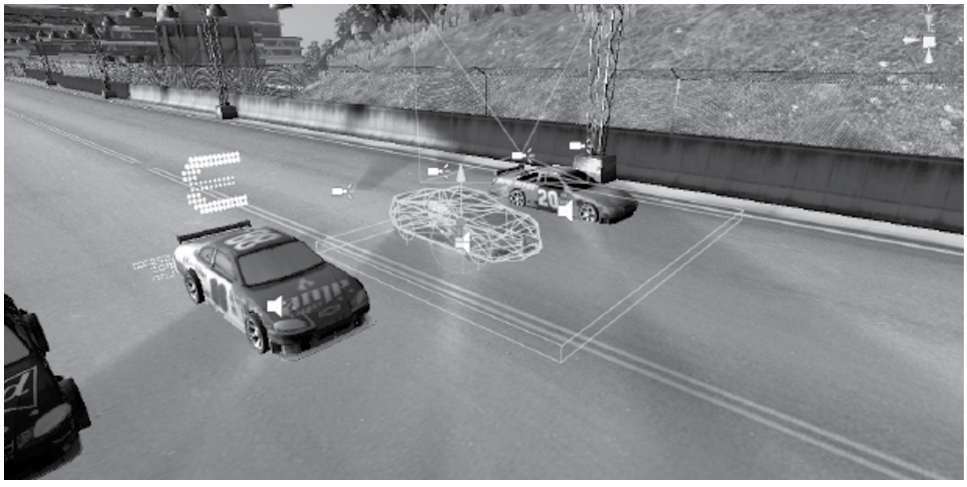
To adequately steer the car away from obstacles on the road requires that the car can adjust to changing road conditions. In other words, the car needs to exhibit environmental perception. The Unity engine provides the mechanism of trigger detection that supports artificial environment perception. To implement trigger detection, a trigger detection area has to be first defined for a car, and then, when some other object enters into this trigger detection area, it can be detected. For example, if a wall comes into the trigger detection area, the vehicular system becomes aware of where the wall is with respect to the car. The system then adjusts the steering and braking levels of the car accordingly so as to steer the car away from the wall. Hence the trigger detection component is important for realizing artificial environment perceiption.

An alternative technique considered for implementing artificial environment perception was the ray-casting technique. Ray-casting involves drawing a ray or line from an object towards a specified direction, which in the case of the vehicular system simulator, is the direction the car is heading. This drawn line or ray can be used for determining whether a car is about to collide with another object. However, the ray-casting mechanism can only detect collisions in one direction. By comparison, the trigger detection mechanism can detect collisions in all directions, and therefore, the latter is adopted in the system. Figure 2 shows a screenshot of the trigger detection area defined for the system-controlled car.

3.3. Car Definition

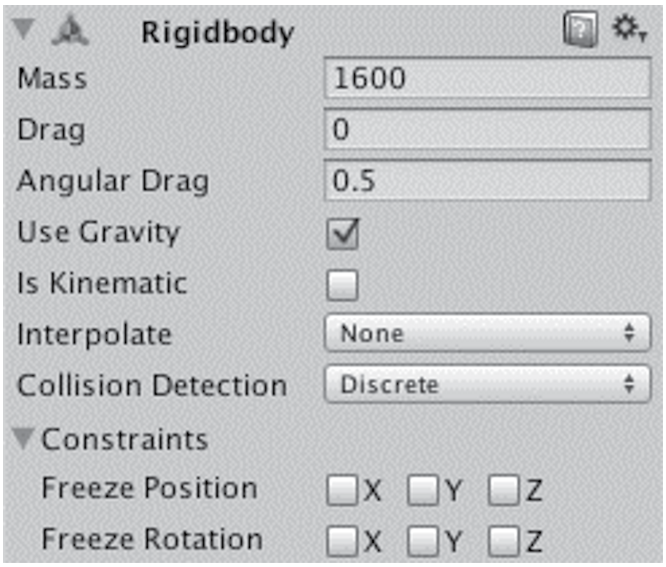
The Unity game engine (trademark of Unity Technologies) provides some features that enable the car to be defined according to user preferences, and two of these are: (1) the rigid-body component, and (2) the wheel-collider component. The rigid-body component enables the user to modify properties of the rigid body object and how physical forces are exerted upon it. Using the rigid body interface, the user can modify the properties of a car such as its mass, drag, and angular drag, as well as its kinetic properties of velocity, angular velocity, rotation, and angular rotation. Figure 3 shows the rigid body interface, which enables users to define the values of the properties of a car.

Figure 2. A screenshot of the trigger detection area defined for the system-controlled car



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. A screenshot of the interface of the rigid-body component



The component of the wheel collider in Unity supports the high-level abstraction implementation of wheel-type objects. This component enables users to specify the various properties of a car that directly affects the frictions between the tires and a road, and hence, a car's ability to navigate on a road system. The user can control the values of the properties of forward and sideways frictions, which are critical for determining how a car drives on a road and in simulating its braking distance. Figure 4 shows the interface of the wheel collider component in Unity, which enables the user to provide input to the various properties of the wheel-collider component of a car.

The coefficient of sideways friction of the wheel collider component plays an important role in situations when the car may slip, overturn, or when it is travelling in super-elevation or in curvilinear motion. On the other hand, when simulating braking distance in rectilinear motion, the coefficient of forward friction is proportional to the forward friction stiffness value in the wheel collider component, which is user-defined. When simulating curvilinear motion, the coefficient of sideways friction is determined by the stiffness value in the parameter of sideways friction of the wheel collider component. Figure 5 is a screenshot of the wheel-collider on the wheels of a car object.

4. IMPLEMENTATION OF THE VEHICULAR SYSTEM

4.1. Unity Game Engine

The vehicular system simulator was built using the Unity game engine developed by Unity Technologies (2016). Unity is a game development engine combined with an integrated development environment (IDE) that is capable of creating applications for Windows, Mac OS X, Linux, as well as other mobile platforms (Unity Technologies, 2016). The Unity game engine can perform graphical rendering and physics calculations, thereby enabling developers to focus on positioning, scaling, and adding realistic physics to objects in the vehicular system.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. A screenshot of the interface of the wheel-collider component

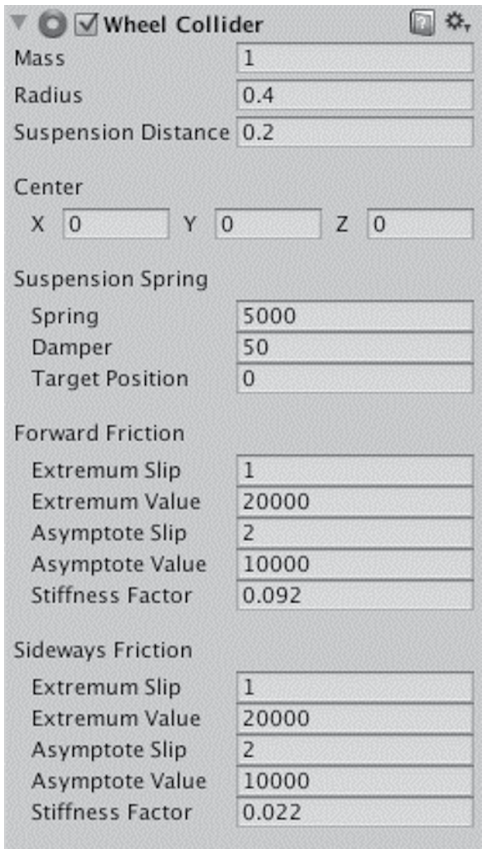
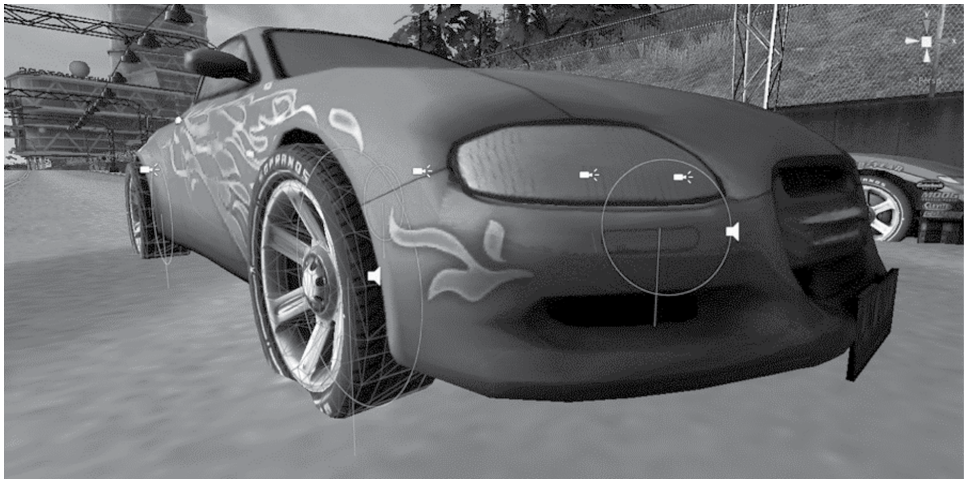


Figure 5. A screenshot of the wheel colliders (displayed as green rings) on the wheels of a car object



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.1.1. MonoDevelop

The Unity game engine works together with MonoDevelop for controlling the behavior of objects. MonoDevelop is an open source IDE developed by Xamarin and the Mono community (MonoDevelop Project, 2016), which is primarily used for development in the C# programming language. In the environment of MonoDevelop, developers can write scripts in C#, Javascript, or Boo to control the logic and behaviors of objects defined in Unity (2016). For example, to implement the braking system on the wheels of the player-controlled car, a script can be written that takes keyboard input from the user and applies the input value to the attribute of the velocity of the car's wheel defined in Unity. Hence, the combined tools of Unity and MonoDevelop enable developers to focus on the development of the system components and they can leave issues related to 3D graphics rendering and physics calculations to the combined platform.

4.2. Programming Languages and Data Structures

The primary programming language used for writing the scripts to control the behaviors of objects was C#, which has advantages because firstly, C# is object-oriented and class-based. Secondly, the language C# has a variety of libraries and frameworks built specifically for development with Unity. The libraries consists of many generic classes and interfaces such as the List, Stack, and Dictionary (Unity Technology, 2016). Thirdly, online documentation is available to support development.

4.3. Vehicular System Modules

The system consists of three main modules, which include: (1) the player-controlled car module, (2) the system-controlled car module, and (3) the environment module. They are discussed as follows.

The player-controlled car module consists of four sub modules: (1) the body, (2) the wheels, (3) the heads-up-display (HUD), and (4) the player controller script. The body of the car is the 3D model that the player sees in the system; this sub module consists of the colliders, which enable the car to collide with objects on the road. The wheels of the car contain wheel colliders, which enable the wheels to make contact with the road and drive the car forward. The wheels sub module contains a script, which is responsible for animating rotation of the wheels as the car travels on the road. The HUD component of the car is controlled by the HUD.cs script; this is responsible for displaying the current speed, position, and lap of the car. Finally, the sub module of the player-controller script, called PlayerCarController.cs, takes the user input from the keyboard and applies the appropriate steering, accelerating, and braking output levels to the player-controlled car.

The system-controlled car consists of four sub modules: (1) the body, (2) the wheels, (3) the trigger area, and (4) the controller script. The body and wheels of the system-controlled car are identical to the body and wheel components of the player-controlled car. The trigger area component of the car is controlled by the CarSensor.cs script; it detects when an obstacle such as a wall enters into the trigger area. When this happens, the system controller is informed and responds by adjusting its control of the car. The system controller script, called AICarController.cs, contains all the implemented artificial intelligence techniques and algorithms responsible for ensuring the system-controlled car can effectively race against the player-controlled car.

The environment module consists of two sub modules, which are the road system and the waypoints system. Both sub-modules have been presented earlier.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

5. DISCUSSION

5.1. Development Tools

Development of the vehicular system simulator was supported by the implementation tools. The Unity game engine provides high-level abstraction programming tools and an intuitive user interface. The engine can support developers in implementation of AI concepts and enable them to ignore lower level development details such as graphics rendering and physics calculations. The combined tools of MonoDevelop and Unity supports the system implementation process because MonoDevelop consists of auto-correction features for many libraries and SDKs used in Unity. For example, the combined tools support the developer in implementation of the waypoint system because the Unity engine supports positioning waypoints in 3D space. Hence, the developer can use this technique of a waypoint system instead of more traditional AI techniques like the A* search method, which is also an effective path finding technique. For controlling the car, the system of conditional monitoring was adopted instead of developing the nonlinear mathematical function, mainly for the sake of simplicity. In terms of collision prevention, the trigger detection technique was used instead of ray-casting because the former is more effective than the latter in that the technique of ray-casting only supports artificial environment perception in one single direction per casted ray. By comparison, the trigger detection technique supports simultaneous perception in all directions surrounding the car. Moreover, the Unity engine has built-in high-level abstractions for trigger detection, which facilitates implementation.

A possible direction for future work is to extend the vehicular system into a simulation system for education purposes, and more specifically, for teaching transportation engineering concepts. The components developed in the system can be used for illustrating graphically many concepts in transportation engineering. For example, one such concept is “braking patterns”, which can include the concepts of: (a) normal braking, (b) accelerated braking, and (c) stonewall braking. Normal braking is the act of slowing down gradually when there is no immediate danger of collision while accelerated braking is the act of decelerating quickly in anticipation of an impending collision. Stonewall braking is the actual occurrence of a collision which forces the car to stop abruptly. The vehicular system simulator can illustrate visually these concepts by defining one or more than one car and their properties. In scenario (a), one car has to be defined, and its braking pattern can be implemented by defining the side friction parameter. Since in scenario (b), two cars are involved and the braking pattern of the first car affects that of the second one, two cars need to be defined. In scenario (c), a stationery object, such as a wall, and a car are involved.

All three of these braking patterns can be simulated using the implementation mechanisms of the car’s artificial environment perception system through the use of a trigger area. The trigger area detects how far away an obstacle is and hence, the distance of a potential collision. Based on the distance, the system can determine whether normal or accelerated braking is required. In addition, the trigger area detects the direction of the potential collision, so that the car can adjust its turn to steer away from the object. If the potential colliding object is a safe distance away, the driving system will apply normal braking to gradually decrease the car’s speed. However, if the colliding object is close to the car, accelerated braking will be applied to slow the car down more suddenly. The value of the side friction parameter can be adjusted to support braking while turning. If the car does come into collision with an object, the magnitude and direction of the force of the impact on the car can be detected and the information displayed to the user for further analysis. These sample features of the vehicular simulation system can all be useful for visual and real time presentation of concepts in transportation engineering.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

5.2. Extension Into an Educational Simulation System

A possible direction for future work is to extend the vehicular system into a simulation system for education purposes, and more specifically, for teaching transportation engineering concepts. The incorporation of game or simulation systems for teaching transportation engineering has been the focus of research, and some sample works are presented as follows.

Chen & Levinson (2006) investigated the efficacy of using simulation for teaching the topic of transportation network growth by conducting an experiment at the Civil Engineering Department of the University of Minnesota. A network growth simulator program (SONG 1.0) was incorporated into a senior/graduate class in transportation system analysis. Instead of aiming to generate a perfect reproduction of the real world SONG 1.0 aims to help students explore the micro-world of transportation network systems so that they are stimulated and encouraged to come up with innovative ways of conceptualizing the network growth and planning process. The experimental results demonstrate that the use of SONG 1.0 effectively enhanced students' in-depth understanding about development of network patterns, which in turn helped to develop their skills in problem-solving and decision-making. Zhu *et al.* (2011) reported on a transportation planning software package called Agent-Based Demand and Assignment Model (ADAM). The agent-based modeling approach assumes that the aggregate urban travel demand arises from a multidimensional selection process, which includes choices related to residential and business locations, trip origins, trip destinations, routes, etc. All agents in the model are assumed to have individual characteristics, goals, and patterns of travel behaviors. The agents exchange information on their experiences and adjust their travel choices according to the available information. ADAM was incorporated as part of a teaching strategy, and evaluations showed that ADAM helped students in forming opinions, evaluating projects, and making judgments. Students who used ADAM reported an enhanced understanding of the concepts taught in the class, and students who prefer visual and active learning were found to benefit more than others. Kasaraneni (2009) presented a web-based game system that demonstrates driver-behaviors at signalized intersections. It was used as a supplementary tool in an introductory course on Transportation Engineering, and helped students understand decisions of drivers at the onset of the yellow signal based on existing traffic conditions. The game was implemented on the .NET Framework using Microsoft Visual C#, ASP, and Microsoft Access (all trademarks of Microsoft, U.S.A.). Based on collected data on student engagement in the course, it was found the game system helped students to better focus their attention. The creative instructional activities also provided students with hands-on experience of simulated real world scenarios. The game system discussed in Kasaraneni (2009) was further developed in Abbas & McNair (2014) with the objective of investigating game-aided pedagogy for improving engagement and learning outcomes of students in transportation engineering. The enhanced game system consists of three layers: (1) a web-based user interface to display simulated traffic flows, (2) a game control layer to implement logic controls based on the game context and user commands, and (3) a data source layer, which consists of the two databases of traffic snapshots and traffic scenarios. Assessment of student engagement and learning outcomes was conducted by means of questionnaires and statistical analysis on the collected data indicated the game enhanced learning outcomes.

The components developed in the vehicular system simulator presented in this paper can be ex-tended so that they support student interaction. For example, the consideration of a vehicle's braking distance is important for the design of a road system. The concepts involved can be categorized into linear or curvilinear motion of a vehicle, along a straight or curved road. If it is a curved road, the motion can be further divided into (1) horizontal curvilinear (i.e. on a flat curved road), or (2) vertical curvilinear (i.e.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

on a curved road that travels up and down hilly terrain). Typically, in a transportation engineering class, the concept of braking distance is explained by means of an equation, which expresses the relationships among the variables of initial velocity or speed, final speed, gravity, coefficient of friction between the road and the tires, and the grade of the road. For the last variable of the grade of a road, an uphill road has a positive grade value, and a downhill road has a negative grade value. The equation for braking distance is often used in road design, such that if a velocity of v_o is allowed on the road, then the road must be at least Db meters long so that cars can safely come to a stop. The equation for defining braking distance is given as follows (Papacostas & Prevedouros, 2001),

$$D_b = \frac{v_0^2 - v^2}{2g\left(f \pm G\right)}$$

where

D_b = braking distance,
$v0$ = initial velocity;
v = final velocity;
g = gravity;
f = coefficient of friction;
G = gradient (+ve = uphill, -ve = downhill).

This can be modeled in a simulation software system so that the student or user I supported in providing input values for the velocities, and the equation embedded in the system will determine the length of the road to ensure safe deceleration of a vehicle.

Some features in the current version of the vehicular system simulator can also be used for illustrating graphically concepts in transportation engineering. For example, one such concept is "braking patterns", which can include the concepts of: (a) normal braking, (b) accelerated braking, and (c) stone-wall braking. Normal braking is the act of slowing down gradually when there is no immediate danger of collision while accelerated braking is the act of decelerating quickly in anticipation of an impending collision. Stonewall braking is the actual occurrence of a collision which forces the car to stop abruptly. The vehicular system simulator can illustrate visually these concepts by defining one or more than one car and their properties. In scenario (a), one car has to be defined, and its braking pattern can be implemented by defining the side friction parameter. Since in scenario (b), two cars are involved and the braking pattern of the first car affects that of the second one, two cars need to be defined. In scenario (c), a stationery object, such as a wall, and a car are involved.

All three of these braking patterns can be simulated using the implementation mechanisms of the car's artificial environment perception system through the use of a trigger area. The trigger area detects how far away an obstacle is and hence, the distance of a potential collision. Based on the distance, the system can determine whether normal or accelerated braking is required. In addition, the trigger area detects the direction of the potential collision, so that the car can adjust its turn to steer away from the object. If the potential colliding object is a safe distance away, the driving system will apply normal braking to gradually decrease the car's speed. However, if the colliding object is close to the car, accelerated braking will be applied to slow the car down more suddenly. The value of the side friction parameter can be adjusted

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

to support braking while turning. If the car does come into collision with an object, the magnitude and direction of the force of the impact on the car can be detected and the information displayed to the user for further analysis. These sample features of the vehicular simulation system can all be useful for visual and real time presentation of concepts in transportation engineering.

6. CONCLUSION

The vehicular system simulator was tested by over twenty users, who all reported that they thoroughly enjoyed the experience of racing provided by the system. It was observed that there was approximately equal number of players who were able to win the race against the system-controlled cars. This suggests that not only did the system provide an entertaining gaming experience; it also provided a reasonably engaging and challenging gameplay.

Future work would involve incorporating other path finding techniques into the vehicular system simulator so that it can be more dynamic and exciting. While the waypoint system in the current version of the system is satisfactory, path finding methods such as the A* search method would enable the system-controlled car's path to be dynamically determined based on its current position. This would make it more difficult for the human players to predict the behavior of the system-controlled cars, and the unpredictable behavior would mean a more challenging racing experience for the user.

Furthermore, while this vehicular system simulator was designed for entertainment, the system can be extended to become a simulator system for education. By modifying some of the parameter values of the car, such as the braking coefficients and torque values, and by modifying the road system environment to resemble that of a municipal road system, the current version of the vehicular system can be adapted to become a simulation software for teaching transportation engineering concepts, for training learner-level drivers, or for familiarizing drivers with the roadway layouts and traffic laws of foreign or unknown cities.

ACKNOWLEDGMENT

The first author wishes to acknowledge the generous support of the Graduate Studies Scholarship of the Faculty of Graduate Studies and Research, and the second author wishes to acknowledge Derril McLeod Environmental Scholarship and the Academic Silver Scholarship of University of Regina.

REFERENCES

Abbas, M., & McNair, L. D. (2014). Game-Aided Pedagogy to Improve Students' Learning Outcomes and Engagement in Transportation Engineering. *Proceedings of the 121st ASEE Annual Conference & Exposition*. American Society for Engineering Education, Indianapolis, IN.

ACM. (2016). ACM Policy and Procedures on Plagiarism. Retrieved from http://www.acm.org/publications/policies/plagiarism_policy

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Akilli, G. K. (2007). Games and Simulations: A New Approach in Education? In D. Gibson, C. Aldrich, & M. Prensky (Eds.), *Games and Simulations in Online Learning: Research and Development Frameworks* (pp. 1–3). Hershey, PA, USA: IGI Global. doi:10.4018/978-1-59904-304-3.ch001

Chen, W., & Levinson, D. (2006). Effectiveness of Learning Transportation Network Growth through Simulation. *Journal of Professional Issues in Engineering Education and Practice*, *132*(1), 29–41. doi:10.1061/(ASCE)1052-3928(2006)132:1(29)

Czajkowski, K., Fitzgerald, S., Foster, I., & Kesselman, C. (2001). Grid Information Services for Distributed Resource Sharing. *Proceedings of the 10th IEEE International Symposium on High Performance Distributed Computing* (pp. 181–184). 10.1109/HPDC.2001.945188

DeNero, J., & Klein, D. 2010. Teaching introductory artificial intelligence with Pac-Man. *Proceedings of the Symposium on Educational Advances in Artificial Intelligence*

Foster, I., & Kesselman, C. (1999). *The Grid: Blueprint for a New Computing Infrastructure*. San Francisco: Morgan Kaufmann.

Foster, I., Kesselman, C., Nick, J. M., & Tuecke, S. (2003). The physiology of the grid. In *Grid computing: making the global infrastructure a reality* (pp. 217-249).

Kasaraneni, Y. (2009). Improving Dilemma Zone Protection Control Issue at Signalized Intersection Using a Web-Game. Retrieved from http://scholar.lib.vt.edu/theses/available/etd-09182009-143008/unrestricted/Kasaraneni_Y_T_2009.pdf

Khoo, A., & Zubek, R. (2002). Applying Inexpensive AI Techniques to Computer Games. *Interactive Entertainment*, *17*(4), 48–53.

Liu, Y., Alexandrova, T., & Nakajima, T. (2011). Gamifying intelligent environments. *Proceedings of the 2011 ACM Multimedia Conference and Co-Located Workshops Ubi-MUI '11 Workshop* (pp. 7-12). 10.1145/2072652.2072655

May, P., Ehrlich, H. C., & Steinke, T. (2006). ZIB Structure Prediction Pipeline: Composing a Complex Biological Workflow through Web Services. In W. E. Nagel, W. V. Walter, & W. Lehner (Eds.), *Euro-Par '06, LNCS* (Vol. 4128, pp. 1148–1158). Heidelberg: Springer. doi:10.1007/11823285_121

Microsoft. (2016). Microsoft Developer Network. Retrieved from http://msdn.microsoft.com/en-US/

MonoDevelop Project. (2016). MonDevelop. Retrieved from http://monodevelop.com

National Center for Biotechnology Information. (2016). Retrieved from http://www.ncbi.nlm.nih.gov

Papacostas, C. S., & Prevedouros, P. D. (2001). *Transportation Engineering and Planning* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Prensky, M. (2001). Digital Natives, Digital Immigrants. *On the Horizon, 9(5),* 1-6.

Salen, K., & Zimmerman, E. (2004). *Rules of Play: Game Design Fundamentals*. MIT Press.

Samuel, A. L. (1967). Some Studies in Machine Learning Using the Game of Checkers. II-Recent Progress. *IBM Journal of Research and Development*, *11*(6), 601–617. doi:10.1147/rd.116.0601

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Smith, T. F., & Waterman, M. S. (1981). Identification of Common Molecular Subsequences. *Journal of Molecular Biology*, *147*(1), 195–197. doi:10.1016/0022-2836(81)90087-5 PMID:7265238

Tatai, V. K., & Gudwin, R. R. (2003). Using a Semiotics-Inspired Tool for the Control of Intelligent Opponents in Computer Games. *Proceedings of the International Conference on Integration of Knowledge Intensive Multi-Agent Systems* (pp. 647-652). 10.1109/KIMAS.2003.1245115

Unity Technologies. (2016). Unity. Retrieved from http://unity3d.com

Valve Corporation. (2016). Counter-Strike. Retrieved from http://www.valvesoftware.com/games/css.html

Waterman, D. A. (1970). Generalization Learning Techniques for Automating the Learning of Heuristics. *Artificial Intelligence*, *1*(2), 121–170. doi:10.1016/0004-3702(70)90004-4

Wood, L. C., & Reiners, T. (2012). Gamification in Logistics and Supply Chain Education: Extending Active Learning. In P. Kommers, T. Issa, & P. Isaias (Eds.), *IADIS International Conference on Internet Technologies & Society 2012* (pp. 101-108). IADIS Press, Perth, Australia.

Xie, F., & Xu, Z. (2015). Semantic Based Annotation for Surveillance Big Data Using Domain Knowledge. *International Journal of Cognitive Informatics and Natural Intelligence*, *9*(1), 16–29. doi:10.4018/IJCINI.2015010102

Xu, Z., Zhi, F., Liang, C., Mei, L., & Luo, X. (2014). Generating Semantic Annotation of Video for Organizing and Searching Traffic Resources. *International Journal of Cognitive Informatics and Natural Intelligence*, *8*(1), 51–66. doi:10.4018/ijcini.2014010104

Yoshizawa, A., & Iwasaki, H. (2015). Influence of Nonvisual Secondary Tasks on Drivers Pedestrian Detection. *International Journal of Cognitive Informatics and Natural Intelligence*, *9*(4), 21–32. doi:10.4018/IJCINI.2015100102

Zhu, S., Xie, F., & Levinson, D. (2011). Enhancing Transportation Education through Online Simulation Using an Agent-Based Demand and Assignment Model. *Journal of Professional Issues in Engineering Education and Practice*, *137*(1), 38–45. doi:10.1061/(ASCE)EI.1943-5541.0000038

This research was previously published in the International Journal of Software Science and Computational Intelligence (IJSSCI), 9(1); edited by Yingxu Wang; pages 55-68; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 35
A Novel Architecture for Learner-Centric Curriculum Sequencing in Adaptive Intelligent Tutoring System

Ninni Singh
University of Petroleum and Energy Studies, India

Neelu Jyothi Ahuja
University of Petroleum and Energy Studies, India

Amit Kumar
University of Petroleum and Energy Studies, India

ABSTRACT

An ideal face-to-face tutor learner interaction aims to offer learning to the learner in a manner that best suits an individual learner's learning level and learning style. This ability of differentiated instruction has been built in Seis-Tutor Intelligent Tutoring system, developed to offer subject matter knowledge of 'Seismic Data Interpretation,' a field of geo-physics. The detailed architecture of learner-centric curriculum sequencing module, built to this effect, with its components, sub-components, their interconnected functioning, to generate exclusive learning path, have been described. An algorithm for learner-centric curriculum sequencing, a mathematical model and proposed implementation using a case study has been elaborated.

INTRODUCTION

Artificial intelligence is an advanced field of research. Its use in the field of education particularly in making teaching and learning more effective, has been significantly exploited in recent past. This has resulted in development of variety of educational computer artifacts. Further developments have caused origin of the field of intelligent tutoring. This field developed at the intersection of disciplines

DOI: 10.4018/978-1-7998-0951-7.ch035

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

of computer science, cognitive psychology and educational research has gained immense popularity in current times, particularly due to the development of Intelligent Tutoring Systems (ITS) (Nwana et al., 1990). These systems termed as cognitive tutors offer tutoring in a manner that best suits the learning preference of the learner. It is a program that not only behaves like a human tutor but also follows the rules and instructions based on the progress and behavior of the learner.

ITS is an intelligent computer system which solves the learner's problems by bestowing feedback and hints. It traces the learner's activity, assesses the learner's mastery on subject matter and seeks to understand the learner's psychological mind. Cognizing psychological mind of learner makes a tutoring system an "intelligent" tutoring system because it solves the learner issues and offers the tutoring content in a manner that learner can grasp easily and effectively. An ITS is distinctly different from a typical e-learning system, such as a web-based learning system, that facilitates a learner to explore a specific domain or course contents via internet. Like ITS, these systems do not adapt to the learner's learning needs and also do not offer learner-specific feedback and hints. ITS being adaptive in nature, has gained immense popularity in current times.

Now-a-days, adaptation and personalization are two important aspects of computer aided instruction systems. On web-based education systems, study material resources are available, but with their overloading and accessibility issues. Hence, it is difficult for a learner to achieve his/her learning goals in cyberspace. In an education system, categorization of learning material and related activities, to present to learner, in learner exclusive manner, targeted to achieve learning objectives, is termed as learning path. This learning path may vary from person to person depending of his/her profile. This is aimed to offer individualistic learning experience to each learner. This is because each individual imbibes knowledge at different speed. The knowledge of a particular learner may grow quite fast and for others the time taken may vary greatly. The response of an individual to learning material too, may be different. A given page may be difficult to understand for a beginner learner and be unimportant and dull for an advanced leaner (Gutiérrez, & Pardo, 2007).

BACKGROUND AND PRELIMINARIES

Very limited work is reported in learner path sequencing in intelligent tutoring systems. However, a reasonably sizable work on curriculum sequencing is reported in e-Learning systems, web-based hypermedia systems and Learning Management Systems. But our systematic review reveals that the ITS systems have remained unexplored in this context so far. Till now, the organization of learning material in intelligent tutoring system, is commonly based on heuristic rules, i.e. during construction of domain, programmer defines rules for all possible situations. This means that initially all learners are provided with same predefined curriculum, but as learners progress, the remedial actions on each activity of the learner are recommended, much the same way like the human tutors do. For example, ITS helps the learner to solve a particular question by providing necessary hints for the question (Brusilovsky, Schwarz, & Weber, 1996, Skinner et al., 1958, Wu et al., 1991).

Some work on curriculum sequencing in ITS is reported below: A depth first traversal algorithm was used in Knowledge-based systems, Hyperbook systems as a sequencing technique (Baldoni, Baroglio, & Patti, 2001).

Chen, (2011) proposed personalized and remedial e-Learning system (PDRLS), providing learning path for specific learners on the basis of their previous knowledge about the course. Firstly, experts describe the relationship among the concepts. Learner pretest plays a vital role in the identification of

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

learning path. For identifying learner previous knowledge structure, they used pathfinder network algorithm, which takes outcomes of pretest as input. At last they compare knowledge structures and generate the appropriate learning path.

Concept map is a way of organizing knowledge in form of courses, topics and subtopics, represented as squares, circles and connection between them indicating the degree of relativeness or relation. The idea of concept map is proposed by (Novak, & Cañas, 2008) on meaningful theory of Ausubel.

Xie, Zou, Wang, Wong, Rao, & Wang, (2017) proposed a technique, to resolve issue related to finding suitable learning path for a group of learners. Profile based framework was proposed for determining their learning path by considering various parameters (temporal boundaries and prototype based aggregation) related to learning.

Further, new approach using data mining techniques has also been explored. With the objective to discover new and meaningful knowledge data mining techniques such as prediction, association, classification have been used. Such systems monitor learner web usage as they navigate through the web pages. Hsieh, & Wang, (2010) proposed web-based learning system that works in two important steps. First is learning path recommendation and second, learning path construction. First, the system generates an initial learner course using data mining technique based on an apriori algorithm. Later learning path is constructed using formal concept analysis that determines the relationship between the course concepts and generates appropriate learning path.

Techniques like fuzzy rule matching have been used to discover association between learner's requirements and list of course materials based on web navigation pattern (Li, & Huang, 2006).

Another Karampiperis, & Sampson, (2004) experimentation has been use of the concept of ontologies and learning object metadata to address the adaptive learning object sequencing problem in intelligent learning management system. They also presented the decision framework for extracting the appropriate learning path according to steps of learner's navigation.

In recent past, nature-inspired and evolutionary algorithms have been used to solve adaptive curriculum sequencing problem. Classification problem of artificial neural network (ANN) Dron et al., (2002) has been used for selection of appropriate learning object.

Chen (2008) proposed a genetic based learning path sequencing method for generating learner tailored learning path. They used level of difficulty as well as concept relation degree parameters for the fitness functions.

Another approach was use of genetic algorithms in e-learning systems. Their proposed methodology is based on a pre-test, to collect incorrect learning concept of learners, then GA is employed to construct a near optimal learning path according to incorrect response patterns of pre-test.

Ant Colony Optimization (ACO), a bio-inspired artificial intelligence technique has also gained attention. The first ant system was proposed by Seridi, Sari, Khadir, & Sellami, (2006) and was successfully applied in solving traveling salesman problem. The optimization of learning path in educational hypermedia system is done in such a way that the structure of the learning material is represented by a graph with valued arcs whose weights are optimized by virtual ants (learners) that release virtual pheromones along their path. In Hong, Chen, Chang, & Chen, (2007) the adaptation of the learning path is done towards student's learning style and optimized using Ant's Algorithm. Bonabeau, Dorigo, & Theraulaz, (1999) introduced an approach to recommending the sequence of e-learning modules for distance learners based on self-organization theory. The architecture they named as 'educational wayfinding support' that supports the recording, processing and presentation of collective behavior designed to create a feedback loop, guiding learners on successful path towards learning goals based on ant-algorithm. Similar techniques

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

can be applied using a probabilistic graph where nodes are pedagogical items, arcs are hypertext links with their associated probabilities and students play the role of ants that traverse the graph (Dron et al., 2002, Weber, & Specht, 1997, Brusilovsky, Ritter, & Schwarz, 1997). Paraschool (Dahbi, Elkamoun, & Berraissoul, (2006), Van den Berg, Van Es, Tattersall, Janssen, Manderveld, Brouns, & Koper, (2005) applied the ACO heuristics to pedagogic material navigation problem, and Valigiani, Biojout, Jamont, Lutton, & Collet, (2005) experimented with an "ant-hill" method which laid the pheromone depending on how students validated an item (success or failure), so as to optimize learning path with different students that have different views. Style-based ant colony system (SACS) by Semet, Lutton, & Collet, (2003), Wang, Wang, & Huang, (2008) is based on group learning pattern and employed an extended ant-colony system approach. It is proposed to construct a user model for finding suitable learning path from graph-based path structure for solving most global optimal problem.

Sengupta, Sahu, & Dasgupta, (2012) also proposed an ant colony optimization approach to achieve adaptivity in sequencing of learning concept. In order to determine the relation among the concept they utilize the frequent graph pattern.

Kardan, Ebrahim, & Imani, (2014) proposed two stage robust learning path algorithm termed as ACO Map. In first stage, they determine the knowledge level of the learner based on the outcomes of pretest, in the second stage, they utilize the metaheuristic method, i.e. ant colony optimization, to generate a learning path which best suits the learner profile.

In ELM-ART ITS (Brusilovsky, Schwarz, & Weber, 1996) provides adaptive hypermedia and path sequencing which help learner to navigate through the learning material. The course content of this ITS is hosted on internet in hypertext form. It comprises of two main components: reference manual and textbook. Textbook has tree like hierarchical structure, which contains topics, subtopics, and related contents, while manual is available for reference to course content. ELM-ART provides navigation support through two hypermedia techniques, adaptive annotation and adaptive sorting. In adaptive annotation, visual artifacts like colors, icons of font are used to represent links, and in adaptive sorting, similarity between links is computed and most relevant links are presented first. The negative side of their proposed technique is that learner gets lost in this complex hyperspace.

In view of positive influence of intelligent and adaptive techniques in education systems, research teams have experimented with intelligent features in intelligent tutoring systems. One of the initial features explored in these systems is use of sequencing technologies (question sequencing, page sequencing, and curriculum sequencing, etc.). Subsequent features to follow are adaptive problem-solving support, navigation support, etc. Literature study reveals, intelligent technologies have been in experimentation and research since long. Presented below is a systematic cataloging of prevalent systems. Table 1 shows the integration of features of adaptive hypermedia and intelligent tutoring in web based tutoring systems. Table 2 elaborates web-based hypermedia systems that incorporate intelligent techniques. Table 3 illustrates web-based intelligent tutoring systems.

PROPOSED ARCHITECTURE AND METHODOLOGY

This section elaborates the work flow of proposed curriculum sequencing module built in Seis-Tutor intelligent tutoring system. Curriculum planning and its sequencing as per learner's learning style and level, are an easy and natural process exhibited by a human tutor. Planning and sequencing course content dynamically is one of the cognitive intelligences, exhibited by a human. To mimic this cognitive intel-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. A Web Based Tutoring Systems that incorporate functionality of Intelligent Tutoring and adaptive Hypermedia Systems

Sno.	System	Curriculum Sequencing	Navigation Support	Reference
1.	ELM-ART	Links/ Pages	Explanatory note	(Brusilovsky, Schwarz, & Weber, 1996).
2.	ELM-ART-II	Tests, Course	Explanatory note	(Weber, Specht, 1997).
3.	VC Prolog Tutor	Page		(Peylo, Teiken, Rollinger, & Gust, 1999)
4.	PAT-InterBook	Task	Explanatory note	(Brusilovsky, Ritter, & Schwarz, 1997).

Table 2. Intelligent Technologies incorporated in Adaptive Web- based Hypermedia Systems

Sno.	System	Curriculum Sequencing	Navigation Support	References
1.	AST	Course	Explanatory note	Specht, Weber, Heitmeyer, & Schöch, 1997)
2.	ADI	Course (Interest + Knowledge)	Explanatory note	(Schoch, Specht, & Weber, 1998)
3.	ACE	Course (Interest + Knowledge)	Explanatory note, Hiding	Specht, & Oppermann, 1998)
4.	InterBook	Page	Explanatory note	Bursilovsky, Eklund,
5.	PT		Hiding	(Kay, & Kummerfeld, 1997)
6.	AHA		Explanatory note, Hiding	(Bra, & Calvi, 1998)
7.	KBS HyperBook	Course	Explanatory note	(Henze, Naceur, Nejdl, & Wolpers, 1999)

Table 3. Web- Based Intelligent Tutoring Systems

Sno.	System	Curriculum Sequencing	Problem solving Assistance	References
1.	Medtec	Task	-	(Eliot III, Neiman, & LaMar, 1997)
2.	DCG	Course	-	(Vassileva et al.,1997)
3.	PAT-Online	-	Partial	(Ritter et al.,1997)
4.	ILESA	Problems, Lesson	Server	(López, Millán, Pérez-De-La-Cruz, & Triguero, 1998)
5.	PAT-JAVA	-	Server	(Ritter, Brusilovsky, & Medvedeva, 1998)
6.	ADELE	-	Server	(Shaw, Johnson, & Ganeshan, 1999)
7.	ADIS	-	Server	(Warendorf, & Tan, 1997)
8.	MANIC	Topic		(Stern, & Woolf,1998)
9.	TEMAI	-	Partial	(Paiva, & Machado, 1998)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

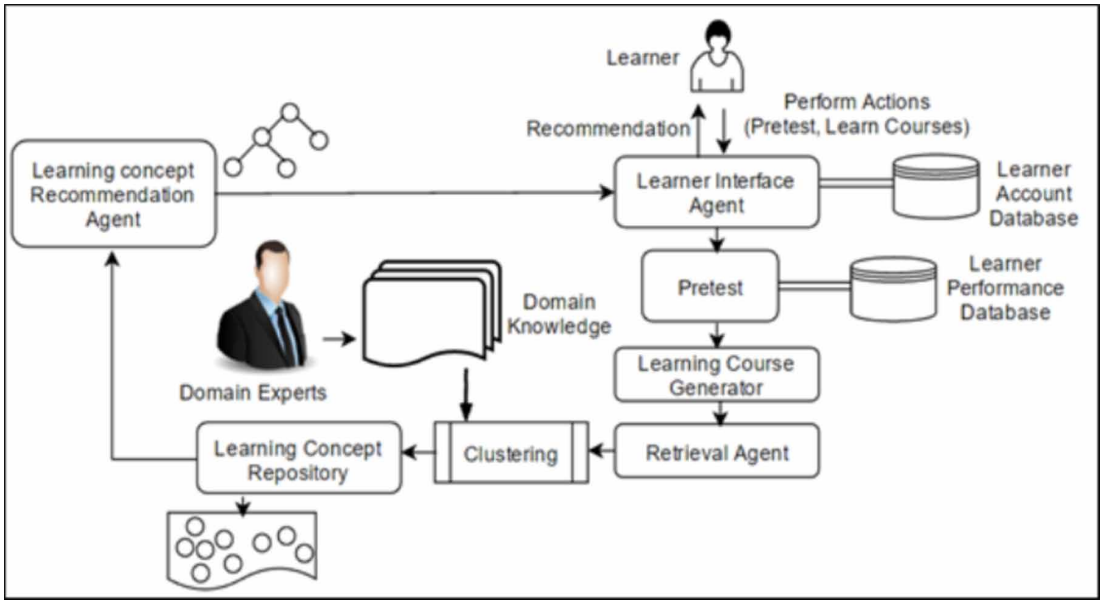
ligence of human by machine, is one of key research aspects discussed here. The proposed architecture of curriculum planning and sequencing module of SeisTutor in intelligent tutoring system has been described below.

Curriculum sequencing using evolutionary techniques has been proposed to make the tutoring adaptive to the learner. To better detail the functionality of Seis-Tutor, this part of the system is termed as 'Tutoring Engine'. The architecture of tutoring engine showing three agents and three databases is presented in Figure 1. The three agents are 'learner interface agent,' 'retrieval agent' and 'learning concept recommendation agent'. The three databases are 'learner account database,' 'learner performance database' and 'domain knowledge database'. The sub-components that carry out the actions within tutoring engine are the 'learner interface agent,' 'learning course generator,' 'retrieval of learning concept' and 'learning concept recommendation agent.'

Learner Interface Agent

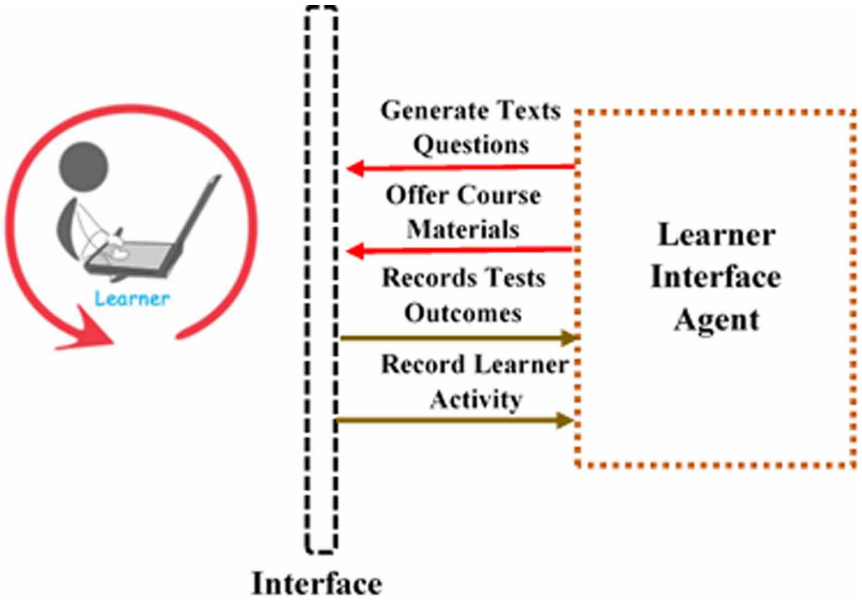
The objective of learner interface agent is to offer an adaptable learning interface to the learner to interact with the intelligent tutoring system. The aim of pretest module is to generate the test items that help the tutoring system to adjudge pre-knowledge level and learning preferences of the individual learner. It then allots a level and learning style to the learner. These are provided to learning course generator, which generates an exclusive tutoring strategy for the learner that is expressed or represented as exclusive learning path implemented through one of the soft computing techniques (see Figure 2).

Figure 1. Curriculum Sequencing architecture in Seis-Tutor



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Learner Interface Agent work flow in Seis-Tutor

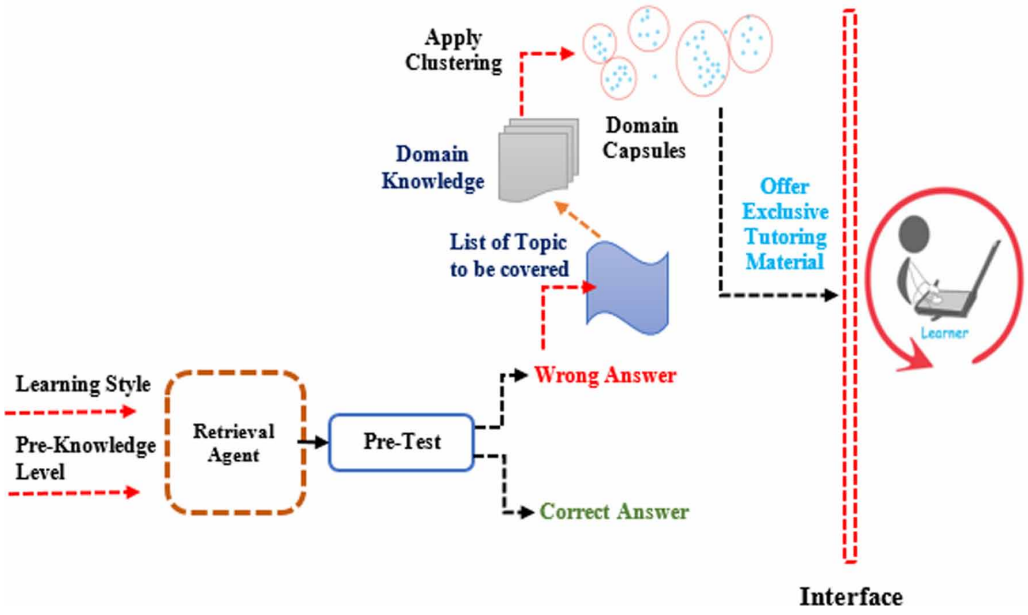


Retrieval Agent

After the identification of exclusive tutoring strategy, retrieval agent triggers its actions, i.e. it applies clustering technique on domain knowledge. The purpose of clustering technique is classification of the pool of course contents into clusters, each represented by a combination of pre-knowledge level and learning style, the repository so formed are termed as learner concept repository. The possible media formats for course content are presentation, document and video. Clustering these media has been implemented through K-Mean, hashing and unsupervised clustering technique. Based on the incorrect responses in the pretest, an appropriate curriculum sequence is determined. Curriculum sequence is represented in the form of a learning path. Retrieval agent examines responses of the learner and based on analyzed learner problems, through the ongoing learning process, Seis-Tutor determines appropriate learning path. Assuming the test comprises of one question each from each subtopic, and the response of the learner indicates that he/she has got some questions wrong, then his/her exclusive learning path will essentially comprise of those subtopics and in required detail, rather than those sub-topics, whose questions are answered correctly. An implementation has been built using, rule-based mechanism with production rules. Using these, exclusive learning path for a learner is chosen from a pool of available learning paths. Topics, sub-topics and related course content along with their association as defined by domain experts is maintained as part of the domain module of Seis-Tutor. Retrieval agent selects topics, subtopics, and course content from the learning concept repository and populates the learning path. Then the learning path is displayed through learner interface (see Figure 3).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Retrieval Agent work flow in Seis-Tutor



Learning Concept Recommendation Agent

The learning concept recommendation agent guides the learning process of the learner. It records the learning progress and performance data into the Learner Account Database (see Figure 4).

Flow of Control of Tutoring Engine

The section below presents the working of the tutoring engine. Firstly, the learner logs into the Seis-Tutor through learner interface, assisted by the learner interface agent. The learner interface agent authorizes, by checking for existence of learner profile in learner account database. If the learner is already registered, the agent loads the profile from learner account database and assists the learner to proceed with unfinished tutoring session, else considers learner as new to the system and administers pretest. The questions for pretest are taken from the pool maintained in learner performance database.

The pretest is made available to the learner through learner interface. The learner interface agent hands over the outcomes of pretest to the pretest module. The pretest module examines the pretest results, determines the pre-knowledge level and learning style of the individual learner and conveys to learner course generator. The learner course generator generates course

sequence based on adjudged pre-knowledge level and learning style of the learner. The exclusive learning strategy manifested as course (curriculum) sequence has been implemented using fuzzy soft computing techniques. This then is handed over for execution to the retrieval agent. Thus, an exclusive learning strategy is progressed for the learner and performance during ongoing tutoring is recorded. Accordingly, the learning path may undergo modification in real time, with an aim to provide best and exclusive learning experience (see Figure 5).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Work flow of Learner Concept Recommendation Agent in Seis-Tutor

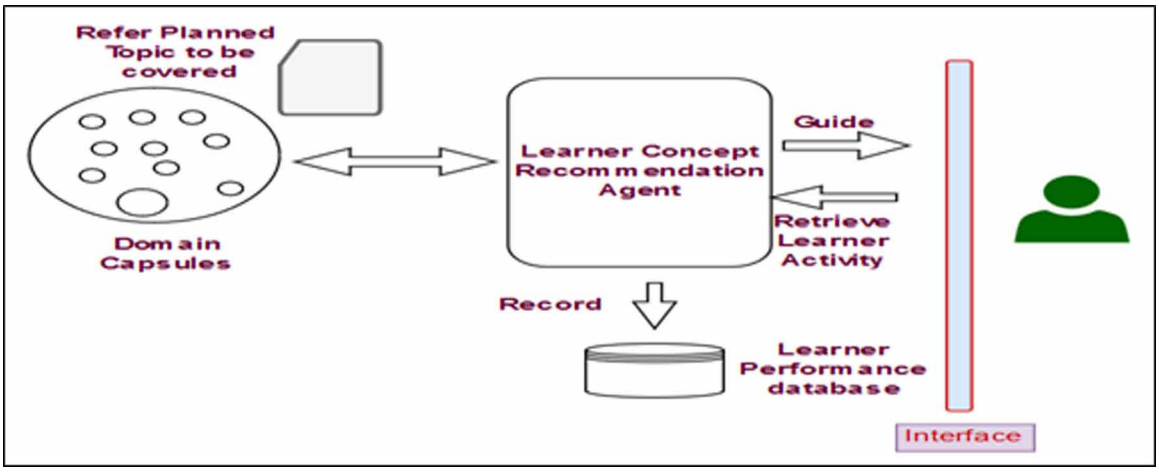
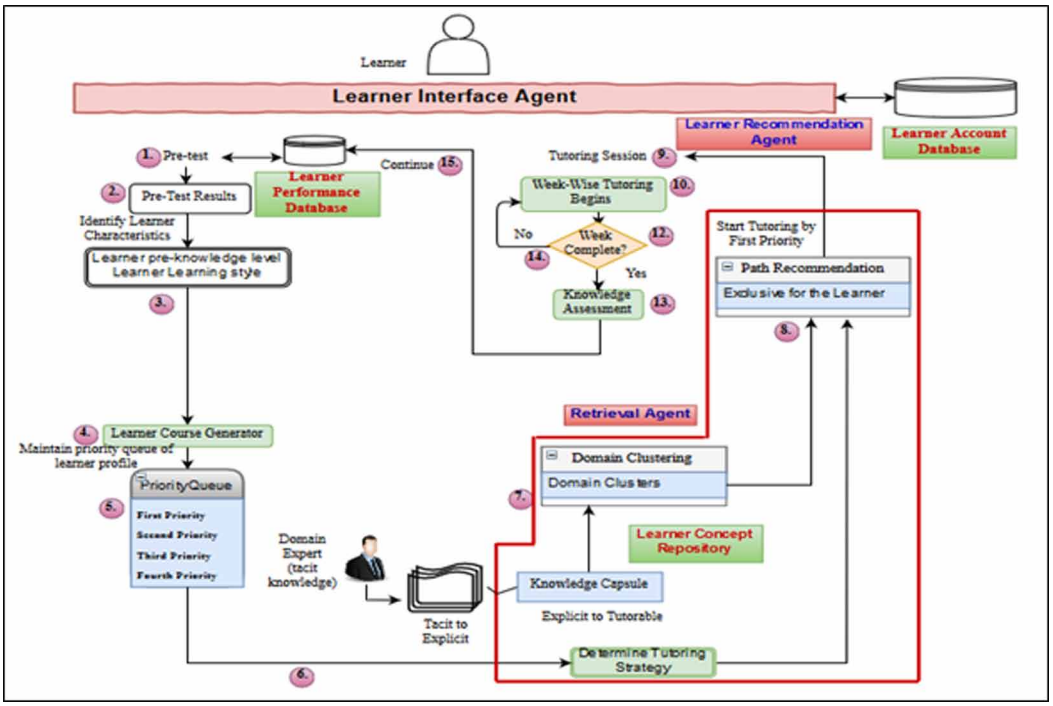


Figure 5. Step-Wise Execution of SeisTutor



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Nomenclature

Symbol	Definition
T	Set of topics to be covered
Q	Set of questions asked in Pretest.
$Q_{Correct}$	Correct response
Q_{wrong}	Incorrect response
Topic	Topic to be covered
t_1	Introduction
t_2	Seismic Waves and their Types
t_3	Reflection Seismology
t_4	Source and Receivers
t_5	Analog and Digital recording
t_6	Seismic Interpretation Methods
t_7	Faults in Petroleum Provinces
t_8	Stress & Strain
t_9	Seismic Expression of Faults
t_{10}	Interpretation of Fault data and 3-D data
t_{11}	Velocity Interpretation
t_{12}	Hydrocarbon Indicators and Modeling

Mathematical Justification of Proposed Architecture for Curriculum Sequencing

See Table 4 for Nomenclature.

Let us consider T as the set of topics to be covered in the tutoring session.

$$T = \{t_1, t_2, t_3, \ldots\ldots, t_{12}\} \tag{1}$$

Let us consider Q is the set of questions asked in Pre-Knowledge level test (Pretest).

$$Q = \{q_1, q_2, q_3, \ldots\ldots, q_{12}\} \tag{2}$$

Then,

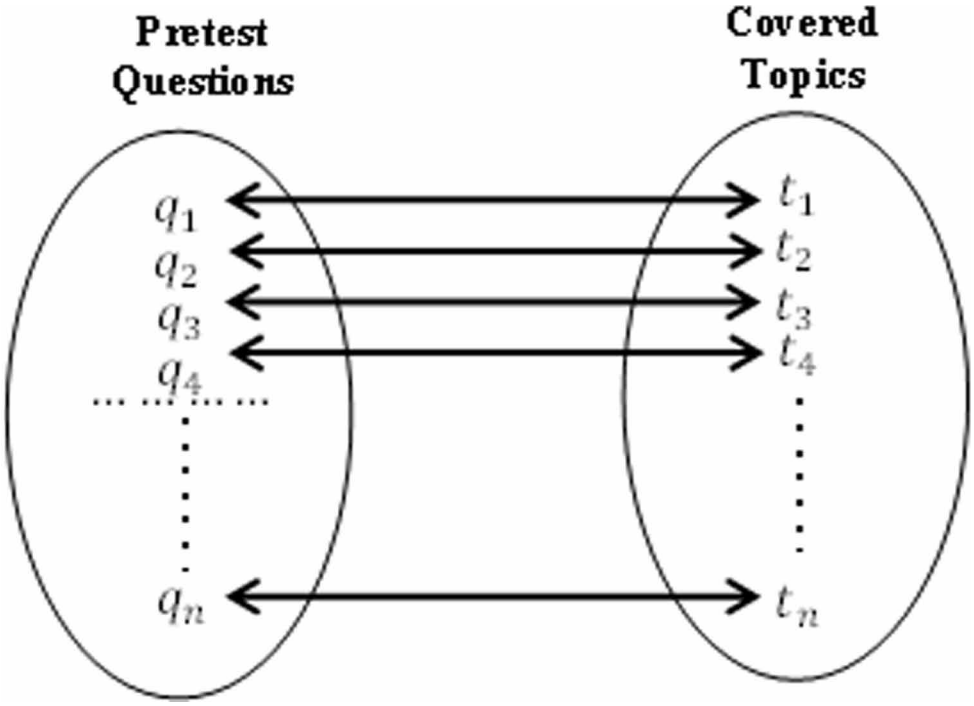
$$Q \in T \tag{3}$$

Equation (3) indicates that each question that was presented by the SeisTutor corresponds to the list of topics that were listed and detailed to be covered.

Note: The relation between questions and topics can be one-to-one and many-to-one, but just for the ease of representation one-to-one mapping has been presented in Figure 6.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. One-to-one mapping state



Let us consider $Q_{correct}$ as the set of questions answered correctly by the learner and Q_{wrong}, is the set of questions answered incorrectly.

$$Q_{correct} = \{c_1, c_2, c_3, \ldots\ldots, c_{12}\} \tag{4}$$

$$Q_{wrong} = \{w_1, w_2, w_3, \ldots\ldots, w_{12}\} \tag{5}$$

Let us consider learner gives correct responses to (Q) 20% of questions.

$$\text{As } Q_{correct} \in Q \text{ and } Q_{correct} \subseteq Q \tag{6}$$

$$Q_{correct} = \{c_3, c_6, c_{10}\} \tag{7}$$

Then 80% of questions Q belong to set of incorrect answered questions.

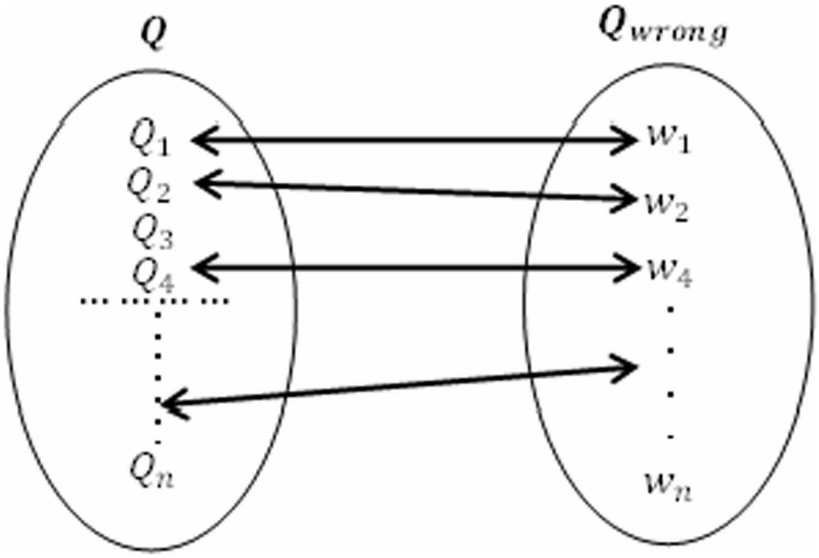
$$Q_{wrong} = (Q - Q_{correct}) \tag{8}$$

i.e.

$$Q_{wrong} = \{w_1, w_2, w_4, w_5, w_7, w_8, w_9, w_{11}, w_{12}\} \tag{9}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. One-to-one mapping between Pretest questions and incorrectly answered questions



Similarly, $Q_{wrong} \in Q$ and $Q_{wrong} \subseteq Q$ (10)

As Q_{wrong} Set consists of questions which are incorrectly answered from the question set Q (see Figure 7). Afore mentioned in equation (3) that $Q \in T$.

Thus, by following transitivity law of algebra, i.e. if Set $A \in B$, and Set $B \in C$, then $A \in C$ or $A \in B \in C$.

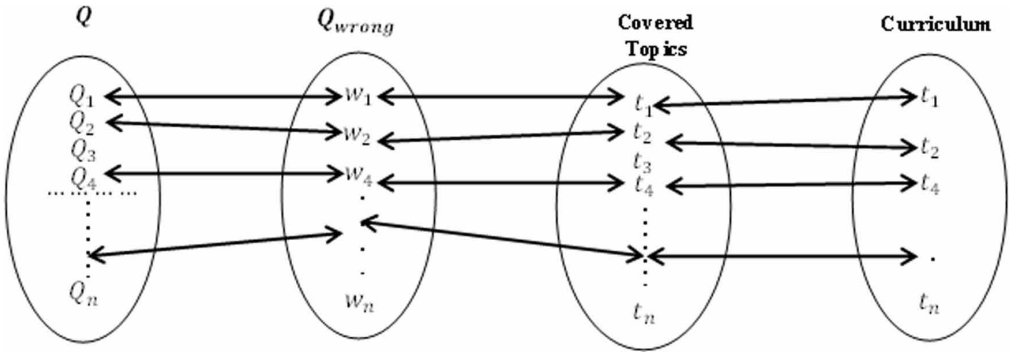
Thus, $Q_{wrong} \in Q \in T$ (11)

$$Topic = \left\{t_1, t_2, t_4, t_5, t_7, t_8, t_9, t_{11}, t_{12}\right\} \tag{12}$$

From equation (11), 12 exclusive curriculum design as developed in SeisTutor is mathematically proved. Same is illustrated by Figure 8.

Presented in Algorithm 1 illustrating step-wise implementation of learner-centric curriculum design based on outcomes of the pre-knowledge level test.

Figure 8. Curriculum designing



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Algorithm 1

Input: Outcomes of Pre-Knowledge level Test
Output: Exclusive curriculum design for the learner

Begin
1. Retrieve the answers opted by the leaner for the given set of questions Q in a variable $RT=\{R_1,R_2,R_3,\ldots\ldots,R_n\}$. Where RT is a set which enclose, learner's responses, Q is the set of asked questions, and $R_1,R_2,R_3,\ldots\ldots,R_n$ are respective learner responses.
2. Matching operation is performed between received results and the master copy of answers, which is stored the database. $MA=\{A_1,A_2,A_3,\ldots\ldots,A_n\}$. Where MA is a set which encloses correct responses and $A_1,A_2,A_3,\ldots\ldots,A_n$ are the respective correct answers.
3. Separate sets of correct responses $Q_{correct}$ and incorrect responses Q_{wrong} were created.
3.1. do
3.2. {
3.3. if($R_i==A_i$)
3.4. {
3.5. $Q_{correct}=Q_i$;
3.6. }
3.7. else
3.8. {
3.9. $Q_{wrong}=Q_i$;
3.10. }
3.11. $i++$;
3.12. } while ($i\leq n$);// where n is the number of questions asked in Pre-Knowledge pretest
4. Step 2 and 3 will repeat for all the response until while condition met.
5. Q_{wrong} Sets are taken into the consideration and perform a mapping operation between the topics covered and Q_{wrong} sets.
5.1. for ($k=1$; $k\leq n$; $k++$)
5.2. {
5.3. for ($l=1$; $l\leq n$; $l++$)
5.4. {
6. if ($Q_{wrong}==W_k==T_l$) // where W is a set which enclose the labels (that directly links to the topics) and $W = \{W_1, W_2, W_3,\ldots\ldots,W_n\}$ are the respective labels associated with the questions.
6.1.
6.2. {
6.3. Topic=T_l;
6.4. }
6.5. else
6.6. {
6.7. continue;
6.8. }
6.9. }
6.10. }.
7. Topic Set consists of the course collection which is exclusively designed for the learner by the SeisTutor.
End

EXPERIMENTAL ANALYSIS AND DISCUSSIONS

Seismic Data Interpretation

Subject matter identified for tutoring in the present work is 'Seismic Data Interpretation,' a sub-field of 'Geophysics'. The primary interest behind petroleum exploration is to uncover the geological substructure opportune for hydrocarbon aggregation. Geologists generally use seismic snaps to portray the earth substructure. Seismic snaps are manually interpreted by the seismologists by utilizing their own skill and expertise. This skill and branch of study is called 'Seismic Data Interpretation'. Due to the absence of formal rules of interpretation, geologists utilize their own experience (established over past years) and skills for interpretation and as a result same seismic image may be interpreted differently by different seismologists (Ahuja, & Diwan, 2012).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

To provide a fair idea of implementation, case studies has been depicted in Table 5, Table 6 and Table 7.

Case Study 1: Mathematical Justification of Differentiation in Curriculum Design Based on Pre-Knowledge Level

Let us consider three learners identified to gain knowledge learn 'Seismic Data Interpretation' through SeisTutor. After formal registration, learner applicants enter Initial assessment phase. The learner is presented with a brief test administered under two categories, "*Pre-Knowledge test"* which adjudges the learner's profile under one of the three different levels (profiles) "Beginner", "Intermediate" and "Expert", and "*Learning style test*" encompassing specific questions to adjudge the learning style of the learner. It was observed that one applicant was allotted 'Beginner profile,' one of them the 'Intermediate profile' and rest of them the 'Expert profile'. Now as per proposed algorithm for curriculum design these three applicants received exclusive curriculum plans, different from each other.

Figure 9 illustrates the learner profile base differentiation. Figure 9 indicates the discrimination of learner profile based on PreTest score percentage. The learner whose Pretest score is less than or equal to 20% categorized as 'Beginner profile' in a similar manner whose Pretest score is between 20% and 40% categorized as 'Intermediate Profile' and whose Pretest score is between 20% and 60% above (categorized as 'Expert Profile.'

Figure 10 describes the closest relation between correct response and the designed curriculum. As beginner provided correct response for only three question thus the designed curriculum comprises additional topics whereas an expert applicant provide correct response to seven questions, thus the designed curriculum comprises less number of topics.

Case Study 2: Mathematical Justification of Differentiation in Curriculum Design Within Similar Pre-Knowledge Level

Let us consider three learners identified to gain knowledge on 'Seismic Data Interpretation' domain via SeisTutor. After formal registration, learner applicants enter Initial assessment phase. The learner is presented with a brief test administered under two categories, "Pre-Knowledge test" which adjudges the

Table 5. Case Study 1

Beginner	Intermediate	Expert
--) Assuming Pretest score of learner to be 20% (categorized as 'Beginner') by Equation (3) --) $\boldsymbol{Q_{correct}}$=$\{c_2, c_6, c_9\}$ by Equation (4) --)$\boldsymbol{Q_{wrong}}$=$\{w_1, w_3, w_4, w_5, w_7, w_8, w_{10}, w_{11}, w_{12}\}$ by Equation (12) --) ***Topic***= $\{t_1, t_3, t_4, t_5, t_7, t_8, t_{10}, t_{11}, t_{12}\}$	--) Assuming Pretest score of learner to be between 20% and 40% (categorized as 'Intermediate') **by Equation (3)** --) $Q_{correct}$=$\{c_6, c_5, c_9, c_{12}, c_7\}$ **by Equation (4)** --) Q_{wrong}=$\{w_1, w_2, w_4, w_6, w_{10}, w_8, w_{11}\}$ **by Equation (12)** --) *Topic*= $\{t_1, t_2, t_4, t_6, t_{10}, t_8, t_{11}\}$	--) Assuming Pretest score of learner to be between 20% and 60% (categorized as 'Expert') **by Equation (3)** --) $Q_{correct}$=$\{c_1, c_2, c_3, c_9, c_{12}, c_7, c_{10}\}$ **by Equation (4)** --) Q_{wrong}=$\{w_3, w_5, w_6, w_8, w_{11}\}$ **by Equation (12)** --) *Topic*= $\{t_3, t_5, t_6, t_8, t_{11}\}$ It can be observed that the curriculum plan for the 'Beginner' is $\{t_1, t_3, t_4, t_5, t_7, t_8, t_{10}, t_{11}, t_{12}\}$, whereas that by 'Expert' is $\{t_3, t_5, t_6, t_8, t_{11}\}$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Learner Profile Discrimination

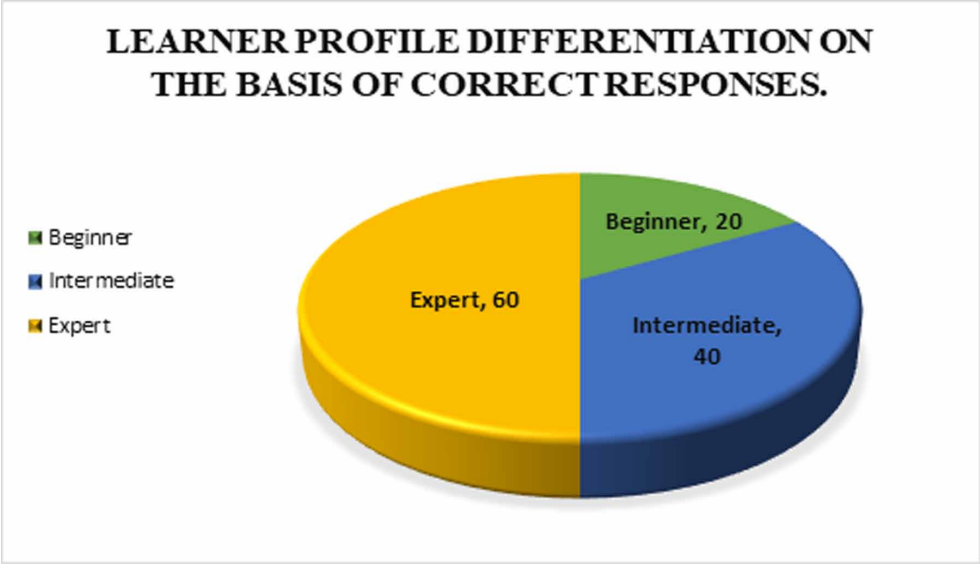
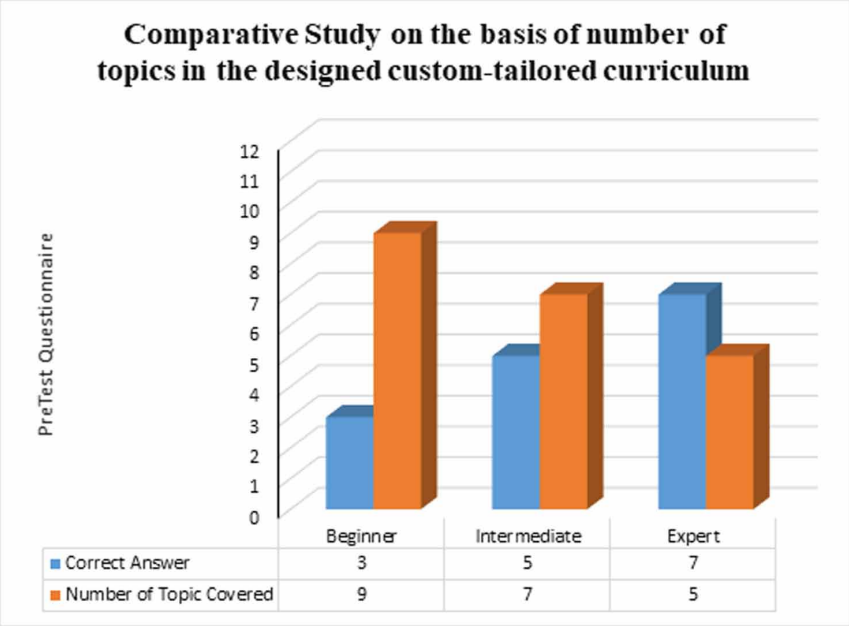


Figure 10. Compatitive study on designed Curriculum



learner's profile as one of the three different levels (profiles) "Beginner", "Intermediate" and "Expert", and "Learning style test" encompassing specific questions to adjudge the learning style of the learner. Considering seven such applicants, 3 were allotted '*Beginner'* profile, two were allotted '*Intermediate*' and one was allotted '*Expert'* profile. Presenting the differentiation amongst different applicants of same knowledge level or profile, here, the case of three '*Beginners*' has been taken. It can be noted below

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. Case Study 2

Beginner Profile 1	Beginner Profile 2	Beginner Profile 3
--) Assuming Pretest score of learner to be 5% (categorized as 'Beginner') by equation (3) --) $\boldsymbol{Q_{correct}}=\{c_2\}$. by equation (4) --) $\boldsymbol{Q_{wrong}}=\{w_1, w_3, w_4, w_5, w_6, w_7, w_8, w_9, w_{10}, w_{11}, w_{12}\}$. by equation (12) --) ***Topic***$= \{t_1, t_3, t_4, t_5, t_6, t_7, t_8, t_9, t_{10}, t_{11}, t_{12}\}$.	--) Assuming Pretest score of learner to be between 1% and 12% (categorized as 'Beginner') **by equation (3)** --) $Q_{correct}=\{c_3, c_5\}$. **by equation (4)** --) $Q_{wrong}=\{w_1, w_2, w_4, w_6, w_7, w_8, w_9, w_{10}, w_{11}, w_{12}\}$. **by equation (12)** **--)** $Topic= \{t_1, t_2, t_4, t_6, t_7, t_8, t_9, t_{10}, t_{11}, t_{12}\}$.	--) Assuming Pretest score of learner to be between 1% and 18% (categorized as 'Beginner') **by equation (3)** --) $Q_{correct}=\{c_4, c_9, c_{12}\}$. **by equation (4)** --) $Q_{wrong}=\{w_1, w_2, w_3, w_5, w_6, w_7, w_8, w_{10}, w_{11}\}$. **by equation (12)** --) $Topic= \{t_1, t_2, t_3, t_5, t_6, t_7, t_8, t_{10}, t_{11}\}$

in Table 6, that each '*Beginner*' is offered exclusive curriculum different from each other. Now as per proposed algorithm for curriculum design these three applicants received exclusive curriculum plans, which are different from each other.

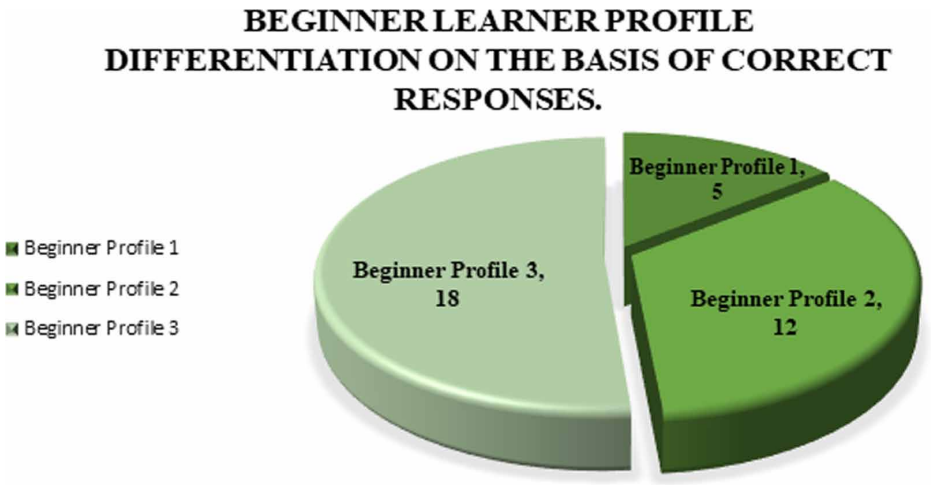
It can be observed that the applicants who belong to similar profiles have been offered distinct curriculum plan. Afore mentioned case study 2 illustrates the learner-centric curriculum plans for different learners of 'Beginner' profile.

Figure 11 illustrates the Beginner learner profile base differentiation. Figure 11 indicates the differentiation of Beginner learner profile based on PreTest score percentage. Here all the applicants belong to beginner profile, hence categorization is performed based on Pretest score. Beginner profile 1 learner Pretest score is 5%, Beginner profile 2 learner Pretest score is 12% and Beginner profile 3 learner Pretest score is 18%.

Figure 12 illustrates designed curriculum is custom-tailored within same level of learner.

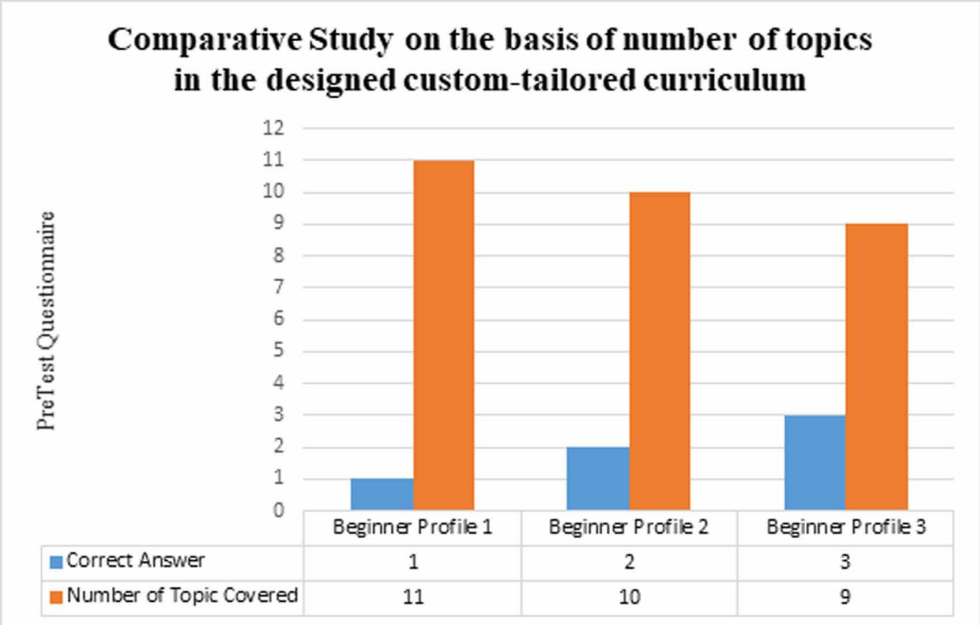
As beginner profile 1 answered only one question thus the designed curriculum comprises more number of topics in a similar manner beginner profile 3 applicant answered three questions thus the designed curriculum comprises less number of topics compared to beginner profile 1.

Figure 11. Beginner learner profile discrimination



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 12. Compatitive study on designed Curriculum



Case Study 3: Mathematical Justification of Curriculum Design for Learner-Cold Start

Let us consider learners identified to gain knowledge on 'Seismic Data Interpretation' domain using SeisTutor. After formal registration, learner applicants enter Initial assessment phase. The learner is presented with a brief test administered under two categories, "Pre-Knowledge test" which adjudges the learner's profile as one of the three different levels (profiles) "Beginner", "Intermediate" and "Expert", and "Learning style test" encompassing specific questions to adjudge the learning style of the learner.

Considering that the applicant has no prior knowledge related to seismic data interpretation, therefore as a result in "Pre-Knowledge test," he or she may not be able to provide a single correct response and be "Beginner" profile. Now, as per the proposed algorithm for curriculum design each applicant receives exclusive curriculum plan, which is different from each other. But in this case, it is very difficult to recommend custom-tailored curriculum. This condition is widely named as "Cold Start Problem."

It can be observed that when the applicants are cold start learners then SeisTutor offers a default curriculum plan.

Research shows that learners follow a variety of mechanisms for gathering and delineating information. Further they have different preferences to understand and process information. These preferences are termed as 'learning styles'. Systematic review of literature reveals that, if learner preferences are pre-identified, then teaching and learning can be provided in a manner that learner learns more effectively. The effectiveness of an ITS depends upon generation of tutoring strategy (A combination of a learner level and a learning style represents a tutoring strategy) by the pedagogy model.

Curriculum sequencing is very vital in ITS, as well as in the classroom, due to its systematic approach, which is congenial to learning. Moreover curriculum sequencing, maps appropriate to a course

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 7. Case Study 3

Beginner Profile
Now, as per the proposed algorithm for curriculum design, these applicants receive default curriculum, because of poor performance in *Pre-Knowledge test"* . The curriculum recommendation system works on bugs found in the PreTest. As aforementioned in Equation (3) that each question that was presented by the SeisTutor corresponds to the list of topics that were listed and detailed to be covered. --) Assuming Pretest score of learner to be 0 (categorized as 'Beginner') by equation (3) --) $\boldsymbol{Q}_{correct}=\{NILL\}$ by equation (4) --) $\boldsymbol{Q}_{wrong}=\{w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8, w_9, w_{10}, w_{11}, w_{12}\}$ by equation (12) --) $\boldsymbol{Topic}= \{t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8, t_9, t_{10}, t_{11}, t_{12}\}$

outline, to be covered. As in classroom teaching, a tutor is expected to provide learning for all learners irrespective of their individual knowledge levels, or learning preferences, hence, a generalized course delivery plan is developed and put into practice. The major drawback of this mechanism is that the teacher is not familiar with the pre-knowledge level, thus he or she is able to bring only very limited variation a teaching style. Whereas, in ITS the aim is to overcome this drawback, to a great extent. Till date, the focus on ITS research, has been majorly on tutoring a learner, as per his/her learning preference, but very limited attention, has been paid on curriculum sequencing, which is the design of an exclusive learning path for that particular learner. Every learner before being admitted to a course holds some knowledge about the domain. Thus, there is a need to predetermine the knowledge level of the learner and based on that, sequence the learning material for the learner, by choosing items from the pool of courseware. The research shows that if the courses are tutored as per the learner pre-knowledge level and preferences, the overall learning gain and degree of engagement of the learner with ITS, are enhanced.

CONCLUSION

Architecture of curriculum sequencing module of adaptive intelligent tutoring system Seis-Tutor, developed for this purpose has been detailed. A human tutor very naturally uses cognitive intelligence to provide learner with appropriate course content, offered through a customized learning path. This ability has been built into Seis-Tutor and it generates an exclusive learning path for the learner. The aim of the exercise is to provide best learning experience and significant learning gain for the learner, comparable to ideal 'face-to-face human tutor learner' classroom interaction.

ACKNOWLEDGMENT

This work is being carried out at University of Petroleum and Energy Studies with the reference number SR/CSI/140/2013. The authors thankfully acknowledge the funding support received from Cognitive Science Research Initiative, Department of Science and Technology for the project. The Authors thank the management of University of Petroleum and Energy Studies, for supporting the work and granting permission to publish it.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Ahuja, N. J., & Diwan, P. (2012). An expert system for Seismic data interpretation using visual and analytical tools. *International Journal of Scientific & Engineering Research*, *3*(4).

Baldoni, M., Baroglio, C., & Patti, V. (2001). *Structureless, intention-guided web sites: planning based adaptation. In* HCI (pp. 237–241). .

Bonabeau, E., Dorigo, M., & Theraulaz, G. (1999). *Swarm intelligence: from natural to artificial systems (No. 1)*. Oxford university press.

Bra, P. D., & Calvi, L. (1998). AHA! An open adaptive hypermedia architecture. *New Review of Hypermedia and Multimedia*, *4*(1), 115–139. doi:10.1080/13614569808914698

Brusilovsky, P., Ritter, S., & Schwarz, E. (1997). Distributed intelligent tutoring on the Web. *Artificial Intelligence in Education: Knowledge and Media in Learning Systems*, *482*, 489.

Brusilovsky, P., Schwarz, E., & Weber, G. (1996). ELM-ART: An intelligent tutoring system on World Wide Web. In Intelligent tutoring systems (pp. 261-269). Springer Berlin/Heidelberg.

Brusilovsky, P., Eklund, J., & Schwarz, E. (1998). Web-based education for All: A Tool for Developing Adaptive Courseware. In *Proceedings of the 7th International World Wide Web Conference*.

Chen, C. M. (2008). Intelligent web-based learning system with personalized learning path guidance. *Computers & Education*, *51*(2), 787–814. doi:10.1016/j.compedu.2007.08.004

Chen, L. H. (2011). Enhancement of student learning performance using personalized diagnosis and remedial learning system. *Computers & Education*, *56*(1), 289–299. doi:10.1016/j.compedu.2010.07.015

Dahbi, A., Elkamoun, N., & Berraissoul, A. (2006). Adaptation and optimisation of pedagogical paths by ants's algorithm. In 2nd Information and Communication Technologies ICTTA'06 (Vol. 1, pp. 546-551). doi:10.1109/ICTTA.2006.1684429

Dron, J. N. (2002). *Achieving self-organisation in network-based learning environments* [Doctoral dissertation] University of Brighton.

Eliot, C. R., III, Neiman, D. E., & LaMar, M. M. (1997, November). Medtec: A Web-Based Intelligent Tutor for Basic Anatomy. In WebNet (Vol. 97, pp. 1-5).

Gutiérrez, S., & Pardo, B. (2007). Sequencing in web-based education: approaches, standards and future trends. In *Evolution of Teaching and Learning Paradigms in Intelligent Environment* (pp. 83-117).

Henze, N., Naceur, K., Nejdl, W., & Wolpers, M. (1999). Adaptive hyperbooks for constructivist teaching. *KI, 13*(4), 26-31.

Hong, C. M., Chen, C. M., Chang, M. H., & Chen, S. C. (2007, July). Intelligent web-based tutoring system with personalized learning path guidance. In *Seventh IEEE International Conference on Advanced Learning Technologies ICALT 2007* (pp. 512-516). 10.1109/ICALT.2007.167

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hsieh, T. C., & Wang, T. I. (2010). A mining-based approach on discovering courses pattern for constructing suitable learning path. *Expert Systems with Applications*, *37*(6), 4156–4167. doi:10.1016/j.eswa.2009.11.007

Karampiperis, P., & Sampson, D. (2004, August). Adaptive instructional planning using ontologies. In *Proceedings. IEEE International Conference on Advanced Learning Technologies, 2004* (pp. 126-130). 10.1109/ICALT.2004.1357388

Kardan, A. A., Ebrahim, M. A., & Imani, M. B. (2014). A new personalized learning path generation method: ACO-Map. *Indian Journal of Social Research*, *5*(1), 17.

Kay, J., & Kummerfeld, B. (1997). User models for customized hypertext. In *Intelligent Hypertext* (pp. 47–69). Springer Berlin Heidelberg. doi:10.1007/BFb0023959

Li, Y., & Huang, R. (2006). Dynamic composition of curriculum for personalized e-learning. *Frontiers in Artificial Intelligence and Applications*, *151*, 569.

López, J. M., Millán, E., Pérez-De-La-Cruz, J. L., & Triguero, F. (1998). ILESA: a Web-based intelligent learning environment for the simplex algorithm. In *Proc. of CALISCE* (Vol. 98, pp. 399-406).

Novak, J. D., & Cañas, A. J. (2008). The theory underlying concept maps and how to construct and use them.

Nwana, H. S. (1990). Intelligent tutoring systems: An overview. *Artificial Intelligence Review*, *4*(4), 251–277. doi:10.1007/BF00168958

Paiva, A., & Machado, I. (1998). Vincent, an autonomous pedagogical agent for on-the-job training. In Intelligent tutoring systems (pp. 584-593). Springer Berlin/Heidelberg. doi:10.1007/3-540-68716-5_64

Peylo, C., Teiken, W., Rollinger, C. R., & Gust, H. (1999). Der VC-Prolog-Tutor: Eine Internet-basierte Lernumgebung. *KI, 13*(4), 32-35.

Ritter, S. (1997, August). PAT Online: A Model-tracing tutor on the World-wide Web. In *Proceedings of Workshop" Intelligent Educational Systems on the World Wide Web" at AI-ED* (Vol. 97, pp. 11-17).

Ritter, S., Brusilovsky, P., & Medvedeva, O. (1998). Creating more versatile intelligent learning environments with a component-based architecture. In Intelligent Tutoring Systems (pp. 554-563). Springer Berlin/Heidelberg. doi:10.1007/3-540-68716-5_61

Schoch, V., Specht, M., & Weber, G. (1998, June). ADI–An empirical evaluation of a tutorial agent. In *Proceedings of ED-MEDIA* (pp. 98-10).

Semet, Y., Lutton, E., & Collet, P. (2003, April). Ant colony optimisation for e-learning: Observing the emergence of pedagogic suggestions. In *Proceedings of the 2003 IEEE Swarm Intelligence Symposium SIS'03* (pp. 46-52). 10.1109/SIS.2003.1202246

Sengupta, S., Sahu, S., & Dasgupta, R. (2012). Construction of learning path using ant colony optimization from a frequent pattern graph. arXiv:1201.3976

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Seridi, H., Sari, T., Khadir, T., & Sellami, M. (2006, July). Adaptive instructional planning in intelligent learning systems. In *Sixth International Conference on Advanced Learning Technologies, 2006* (pp. 133-135). 10.1109/ICALT.2006.1652386

Shaw, E., Johnson, W. L., & Ganeshan, R. (1999, April). Pedagogical agents on the web. In *Proceedings of the third annual conference on Autonomous Agents* (pp. 283-290). 10.1145/301136.301210

Skinner, B. F. (1958). Teaching machines. *Science*, *128*(3330), 969–977. doi:10.1126cience.128.3330.969 PMID:13592277

Specht, M., & Oppermann, R. (1998). ACE-adaptive courseware environment. *New Review of Hypermedia and Multimedia*, *4*(1), 141–161. doi:10.1080/13614569808914699

Specht, M., Weber, G., Heitmeyer, S., & Schöch, V. (1997, June). AST: adaptive WWW-courseware for statistics. In *Proceedings of Workshop "Adaptive Systems and User Modeling on the World Wide Web" at 6th International Conference on User Modeling, UM97, Chia Laguna, Sardinia, Italy* (pp. 91-95).

Stern, M., & Woolf, B. (1998). Curriculum sequencing in a web-based tutor. In Intelligent Tutoring Systems (pp. 574-583). Springer Berlin/Heidelberg. doi:10.1007/3-540-68716-5_63

Valigiani, G., Biojout, R., Jamont, Y., Lutton, E., & Collet, P. (2005, March). Experimenting with a real-size man-hill to optimize pedagogical paths. In *Proceedings of the 2005 ACM symposium on Applied computing* (pp. 4-8). 10.1145/1066677.1066683

Van den Berg, B., Van Es, R., Tattersall, C., Janssen, J., Manderveld, J., Brouns, F., . . . Koper, R. (2005, September). Swarm-based sequencing recommendations in e-learning. In *Proceedings. 5th International Conference on Intelligent Systems Design and Applications ISDA'05* (pp. 488-493). 10.1109/ISDA.2005.88

Vassileva, J. (1997). Dynamic course generation. *CIT. Journal of Computing and Information Technology*, *5*(2), 87–102.

Wang, T. I., Wang, K. T., & Huang, Y. M. (2008). Using a style-based ant colony system for adaptive learning. *Expert Systems with Applications*, *34*(4), 2449–2464. doi:10.1016/j.eswa.2007.04.014

Warendorf, K., & Tan, C. (1997, August). ADIS-An animated data structure intelligent tutoring system or Putting an interactive tutor on the WWW. In *Proceedings of Workshop" Intelligent Educational Systems on the World Wide Web" at AI-ED* (Vol. 97, pp. 54-60).

Weber, G., & Specht, M. (1997). User modeling and adaptive navigation support in WWW-based tutoring systems. In *User Modeling* (pp. 289–300). Vienna: Springer. doi:10.1007/978-3-7091-2670-7_30

Wu, D. (1991). Active acquisition of user models: Implications for decision-theoretic dialog planning and plan recognition. *User Modeling and User-Adapted Interaction*, *1*(2), 149–172. doi:10.1007/BF00154476

Xie, H., Zou, D., Wang, F. L., Wong, T. L., Rao, Y., & Wang, S. H. (2017). Discover learning path for group users: A profile-based approach. *Neurocomputing*, *254*, 59–70. doi:10.1016/j.neucom.2016.08.133

This research was previously published in the Journal of Cases on Information Technology (JCIT), 20(3); edited by Andrew Borchers; pages 1-20; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 36
Designing a Socially Open Narrative Generation System

Taisuke Akimoto
The University of Electro-Communications, Japan

Takashi Ogata
Iwate Prefectural University, Japan

ABSTRACT

The authors propose the design of a Socially Open Narrative Generation System (SONGS) that co-creates a collection of diverse narratives from a narrative generation program and people. This is a challenge of the social application of narrative generation technology used for vitalizing the social activity of producing and sharing narratives. The key idea is to connect and unify individual narrative productions by many agents, including a computer program and many humans, via a collection of narratives produced and accumulated by these agents. At the same time, SONGS is the practice of a computational approach to narratology as a model for the social process of narrative production. This chapter describes the key concepts and mechanism design of SONGS with several experimental programs.

INTRODUCTION

Narrative is a universal form of human knowledge and memory, and a strong way for people to communicate knowledge and information in their society. The computational modeling for narrative generation and understanding is a key topic for the next information society. Such modeling will contribute to the creation of human-friendly information technology and culture development. The fundamental interest of this study is to explore the possibilities of the social application of narrative generation technologies.

This chapter proposes the design of a Socially Open Narrative Generation System (SONGS) that will run on the Web as a space for co-creating a collection of diverse narratives from a narrative generation program and many humans. The ultimate goal of this study is to vitalize the social activity of producing and sharing narratives. Its long-term vision includes, in particular, the creation of a new type of platform

DOI: 10.4018/978-1-7998-0951-7.ch036

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

for expanding and supporting creative activity and a content industry with computational narrative creativity. At the same time, SONGS is the practice of a computational approach to narratology as a model for the social narrative production process.

The remaining part is organized as follows. First, the authors provide background information on intelligent information technologies for narrative. Second, the authors propose the basic design of SONGS. Third, in the subsequent two sections, the authors describe the two main components of the narrative generation and interface parts. Finally, the authors discuss future works and conclude the chapter.

BACKGROUND

This section takes two perspectives for providing background information. The first subsection reviews the history of the studies on narrative generation systems. For further information in this area, refer to the "Introduction" chapter of this book by Ogata or other surveys, such as Gervás (2009). The second subsection discusses the significance of narrative and narrativity for artificial intelligence (AI) and informatics.

Narrative Generation System

Narrative generation systems, or the computational modeling of narrative creativity, are a challenging AI topic. The fundamental objectives include the computer implementation of narrative creativity and elucidation of the human narrative ability or narrative nature. In the early stage, since approximately 1970s to 1990s, researchers attempted to apply AI methods to narrative generation models. These studies are broadly divided into three approaches. First, several researchers attempted to model the story generation process as the planning of character actions directed by goals, emotions, or other bases (Meehan, 1980; Okada & Endo, 1992; Riedl & Young, 2010). Second, the formulation of generative rules or schemes for narrative structures were addressed (Pemberton, 1989; Lang, 1999; Bringsjord & Ferrucci, 1999). These are mainly rooted in a cognitive model of the story scheme (Rumelhart, 1975). The third approach was derived from the framework of case-based reasoning (CBR) (Riesbeck & Schank, 1989). In this approach, narrative generation is formalized as the transformation or adaptation of a previous narrative or case to fit the current problem (Turner, 1994; Peinado & Gervás, 2005; Swanson & Gordon, 2012).

Since approximately 2000s, several researchers have introduced narratology—the theory of narrative mainly inspired by structuralism—to narrative generation studies. Ogata (2002, 2004a) proposed a research methodology called "expanded literary theory" that aims to create synergy between narrative generation technologies and narratology. Gervás, Lönneker-Rodman, Meister, and Peinado (2006) also discussed the importance of interdisciplinary approaches for both disciplines of AI and narratology. In this movement, the morphology of the folktale by Propp (1968) has been applied to several studies as a model or schematic knowledge of story composition (Ogata & Terano, 1991; Peinado & Gervás, 2005; Imabuchi & Ogata, 2012; Gervás, 2015). In addition, narratology provided a distinction between story—"[t]he [content] plane of [narrative]" (Prince, 2003, p. 93)—and discourse—"[t]he [expression] plane of [narrative]" (Prince, 2003, p. 21)—to narrative generation studies. Whereas most of the early systems focus primarily on the story aspect, several recent studies have applied the narrative discourse theory by Genette (1980) to the modeling of narrative discourse composition (Ogata, 2004b; Montfort, 2007; Bae, Cheong, & Young, 2011; Akimoto & Ogata, 2012, 2013).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

For the past few years, Ogata and colleagues have been developing an integrated narrative generation system (INGS) where the three generative modules for story structure, discourse structure, and surface expression are connected (refer to the "Introduction" chapter by Ogata for further information). This includes three narratology-based mechanisms that apply the theories of Propp (1968), Genette (1980), and Jauss (1982) as central elements for the story and discourse generation (Akimoto & Ogata, 2014). Although INGS is in an ongoing process of the research project, this systematic framework provides a solid foundation for the narrative generation mechanism of SONGS.

Meanwhile, narrative generation studies have great application potential for advanced information technologies, but this is almost unexplored. A possibility is to expand the creative activities of people or society. Narrative or narrativity is a common and core element of various artistic or entertainment works, such as novels, film, drama, manga (comic), anime (animation), video games, and commercial films. Narrative generation technologies can create new forms of expression and a creation method for the narrative. Interactive narrative systems are already a major genre of narrative content that uses computer technologies (Riedl & Bulitko, 2013), but this is based only on a semi-automatic narrative generation, and most content is designed by the human creators. Because of technological difficulties, there are very few applied studies on fully automatic narrative generation. However, the authors propose two exploratory applied narrative generation systems: "KOSERUBE," which automatically generates multimedia narrative expressions in the style of a folk tale (Akimoto et al., 2013), and "Narrative Forest," which figuratively expresses the floating narrative generation process within the system using the image of growing trees (Akimoto, Ono, & Ogata, 2013).

Narrative as Information, Knowledge, and Intelligence

Narrativity is a fundamental nature of the human mind that forms memory or knowledge. From the psychological perspective, Bruner (1991) discussed that the human experience and cognition of the real world is based on narrativity and is formed as narratives. In addition, a narrative contextually integrates a variety of things or knowledge and effectively provides rich information to people. From such nature, narratives provide different values to people and society. People tell and receive narratives in various scenes of daily life. Narrative is commonly involved in various genres of art and entertainment. People develop knowledge or cultural literacy and cultivate sensitivity through the accumulation of such artistic experiences. Narrative is an effective and versatile communication tool for various business fields. Quesenbery and Brooks (2010) described practical methods of storytelling for improving the design processes of services or artifacts. Information technology has great potential to expand human activities in a broad range of fields.

Meanwhile, narrative is an essential research subject for the computational implementation of intelligence, a common goal of AI. Schank and colleagues have addressed the story nature of human intelligence and proposed several key AI frameworks, including Conceptual Dependency, script, plan, goal, and CBR (Schank & Abelson, 1977; Riesbeck & Schank, 1989; Schank, 1990). These ideas have significant influence on narrative generation studies.

Several researchers have addressed the construction of large-scale commonsensical knowledge bases mainly based on the concept of script. Singh, Barry, and Liu (2004) designed StoryNet as a database of scripts constructed through crowdsourcing. Chambers and Jurafsky (2010) presented a database of narrative schemas that consists of the semantic knowledge of events and their temporal relationships. This was constructed with their knowledge acquisition algorithms for learning narrative schemas from

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

texts (Chambers, Wang, & Jurafsky, 2007; Chambers & Jurafsky, 2008, 2009). This type of knowledge is applicable to various tasks of natural language processing, such as understanding and summarization, and also provides a basic knowledge for narrative generation. McIntyre and Lapata (2009) and Li, Lee-Urban, Appling, and Riedl (2012) proposed a story generation method based on script similar to chains of events. The case base in the CBR-based narrative generator is also narratives.

Another important research topic is to explore how narrative-based knowledge works in a cognitive system. Although CBR provides a reasonable model, in recent years, several researchers have addressed modeling narrative-centered cognitive architecture and knowledge systems (Szilas, 2015; Anderson, 2015). From the perspective of cognitive systems, a role of narrative creativity is to construct internal narratives as knowledge or episodic memory from external information.

Similar to the topics discussed above, narrative computational modeling is a crucial problem for future intelligent information technology. First, it will expand the creation and sharing of narratives in society. Second, it is necessary for true intelligent systems that construct knowledge through interaction with the real world, subjectively cognize things, create new narratives or knowledge, and effectively communicate with others.

BASIC DESIGN OF SONGS

Figure 1 shows the overall structure of SONGS, which co-creates a collection of narratives (CNs) through the continual interaction of produced narratives from a narrative generation program and many human users. The central idea of the narrative generation program is to model narrative generation as a cyclic process for composing new narratives in relation to previously generated narratives. The authors call this narrative-based narrative generator (NBNG). CNs provide NBNG with comprehensive knowledge for composing new narratives. In addition, SONGS involves an interface that opens CNs to individual humans (users). The design for SONGS assumes that users will access CNs based on their values or objectives, and appreciate contained narratives or create new ones by recomposing those that already exist. Moreover, user-created narratives are added to CNs. In order to allow such usage, the interface provides multiple perspectives on CNs, in particular:

1. Collective view on CNs for exploring valuable narratives,
2. Surface view on a narrative for appreciating it, and
3. Internal view on a narrative structure for creating a new narrative by recomposing the structure.

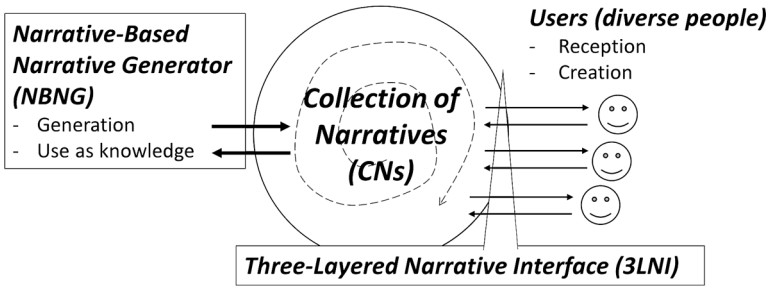
We call the interface that is based on the above framework a three-layered narrative interface (3LNI). The following four subsections describe the key concepts of SONGS.

Narratives as Comprehensive Knowledge for Narrative Generation

Narrative is a type of comprehensive knowledge structure that contains various types of information or knowledge. A narrative can contain a variety of things, including events, humans, objects, abstract concepts, time, space, and so on, within its contextual information or meaning. In addition, the composition of a narrative implicitly includes such abstract or ill-defined information as emotion, sense of value, and aesthetic, which are crucial aspects for generating valuable narratives for people.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Overall structure of SONGS



By focusing on the above narrative nature, SONGS positions CNs as the central knowledge for producing new narratives. The basic principle of using narratives as generative knowledge is to formalize a narrative generation process as recomposition of an existing narrative through structural operations, such as decomposition, transformation, and synthesizing with other narratives.

The point that SONGS positions individual and concrete narratives as the central knowledge is different from a major approach that generalizes narrative structures as a set of abstract rules or schemes—e.g., schematic methods based on Propp's morphology of the folktale (Ogata & Terano, 1991; Imabuchi & Ogata, 2012; Gervás, 2015). However, such abstract knowledge is essentially acquired empirically and dynamically from many narratives, similar to Propp (1968) deriving his morphological knowledge from the analyses of approximately 100 Russian fairy tales. An essential question is to explore a unified framework of the individuality/concreteness and generality/abstractness of narrative knowledge.

On the other hand, NBNG is partially similar to CBR-based narrative generators. Riesbeck and Schank (1989) described the basic procedure of CBR as "[finding] those cases in memory that solved problems similar to the current problem, and [adapting] the previous solution or solutions to fit the current problem, taking into account any difference between the current and previous situations" (p. 25). CBR-based narrative generators basically follow this procedure. However, NBNG uses CNs in flexible ways as types of materials for composing new narratives. In addition, narrative generation in NBNG is directed not as problem solving, but as the production or exploration of some differences and novelty from prior narratives. This concept is discussed in the next subsection.

Narrative Generation in Relation to Prior Narratives

Because the macroscopic goal of SONGS is to accumulate a wealth of narratives through the continual generation cycle, the microscopic goal in each generation process is to produce a narrative that has some novelty or difference from prior narratives. Namely, the prior narratives are not only material knowledge for narrative generation, but basing points for creating new narratives. CNs have a higher-level structure for organizing the narratives, and by using it, NBNG systematically controls the narrative production.

A relationship among narrative works is a crucial aspect for the process of producing narratives in the real world. Several literary models of relative narrative production have discussed in the field of narratology. In early twentieth century, the movement of Russian formalism discussed a model for the evolution of literature as a continual cycle of "automatization" of human perception and alternation of the immanent laws of a particular genre (Tynyanov, 1982). Jauss (1982) presented a new idea concerning the history of literature. Although he respected the previous research regarding Russian formalism,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

he criticized the literary theory of evolution as being dependent upon the dynamism of change in narrative techniques and descriptions. He sought to center his literary process model in the dynamism of the mutual relationship between a literary work/author and the readers. His idea involved modeling the literary history of various relationships between a reader's "horizon of expectation" and literary works, including those estrangements and the change of horizon through a continual reception process.

In another context, Kristeva (1980) described the theory of "intertextuality," arguing that a literary text is positioned as an organization and citation of related preceding texts, and also included subsequent works for the reader. There exist few studies that apply the concept of intertextuality to narrative generation methods. Nakashima and Ogata (2006) proposed several experimental programs that generate new stories through elemental substitution or associations of an input story.

Collective Narrative Generation

The co-creation in SONGS refers to the collective creation process as a continual chain reaction of narrative productions among different types of generative agents—i.e., NBNG and the users—via produced narratives or CNs. Narrative generation in SONGS is basically formalized as the recomposition or derivation of prior narratives. Computers where basic principles of narrative generation are programed can systematically produce a large number of different narratives in order to not be tied to commonsensical or stereotyped constraints on narratives. The distinct narrative creativity will bring inspiration or surprising ideas to the users. On the other hand, the cooperation of diverse users is important for both increasing the diversity of narratives and the generation ability of NBNG. How to attract and activate user creation activity is a key issue in designing SONGS.

Meanwhile, because SONGS aims to create new values for people or society, the problem is similar to the one for artifact design. Co-creation is an important approach in designing complex artifacts that involve pluralistic values related to various people (Fujita & Ota, 2015). Internet technology has expanded co-creation by many people. Tanaka (2011) introduced a novel design method called "Design as Re-Design." This refers to the continual upgrading and remixing of existing open data on the Web through "editing, reforming, or deriving" as a type of collective creation.

Collective creation is traditionally embedded in Japanese culture, and it is a distinguishing feature of contemporary subculture in Japan. Information technologies—e.g., online community spaces such as "Nico Nico Douga" (Niwango, n.d.) and "VOCALOID," a technology for singing voice synthesis (Yamaha, n.d.)—are increasing collective creation activity through "N-th order derivative creation" (Hamano, 2008). This activity is building a singular and rich culture. Co-creation is a valid approach for the problem of narrative generation, and SONGS aims to drive such type of activity by the computational narrative creativity.

Providing Diverse Values to Diverse People

The authors broadly categorize types of users as readers, those who appreciate narratives, and authors, those who use SONGS for creating original narratives. For instance, potential uses of SONGS include the following situations: a manga (comic) artist searches/makes primitive materials for his/her work, a commercial planer searches/makes ideas of scenario for advertising a product, an amateur author creates derivative works of computer-generated narratives, and a student learns multiple perspectives on a social issue from different narratives relevant to the issue.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

SONGS should provide some type of value to diverse users. The authors assume that narrative values include the following two aspects: giving some type of cognitive effect to readers—e.g., impression, surprise, and knowledge discovery—and giving some type of effect to the author's creation process or the result—e.g., inspiration, augmentation, extension, amplification, constraint, and inducing. In the above sense, narrative values vary depending on various factors, such as individual value, purpose or use situation, and historical or cultural background. Hence, it is difficult to computationally define and control narrative values.

The authors apply the computational design methodology called "Design as Search" introduced by Tanaka (2011) to the above problem. In the methodology, "design" refers to a process where the user searches or explores a preliminary modeled solution space for finding a solution that meets his/her need or purpose. In SONGS, the user searches valuable narratives from CNs. This is an application of the concept of "Design as Search." The system will be useful for individual users when there exist few valuable narratives for them, and they can find them. A central problem is designing an interface that allows the user to find valuable or desired narratives, including cases where the user has no explicit narrative needs. This is a common problem with general information retrieval. Although there is another method as a type of on-demand narrative generation, in many cases, users cannot represent their needs or have no clear narrative needs. In addition, it is difficult task to generate narrative that is suitable to the user's needs.

NBNG: NARRATIVE-BASED NARRATIVE GENERATOR

In designing NBNG, the authors start with by modifying and expanding a Jauss-based narrative discourse generation system proposed by Akimoto and Ogata (2012, 2013). A central idea of the system is to model a generative process of historical or continual literature changes based on the reception theory by Jauss (1982). Although the implemented system has many theoretical problems, the fundamental idea is applicable to NBNG. After an overview of the Jauss-based system, the authors describe the NBNG design.

Narrative Discourse Generation System

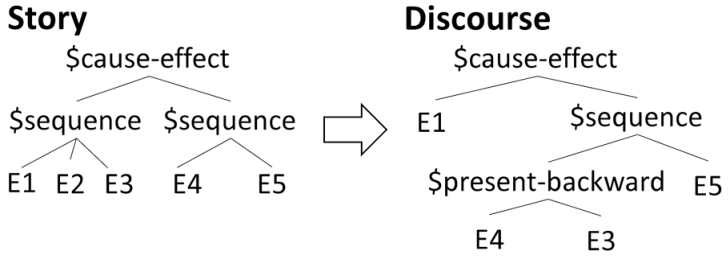
This subsection outlines the Jauss-based system based on Akimoto and Ogata (2012, 2013) and discusses the issues from the perspective of SONGS. The system was developed as a prototype version of the discourse generation part of INGS. The system generates a series of different "discourse" structures from an input "story" structure. A story represents the content information of a narrative. In particular, it represents the temporal sequence of events. A discourse means the structural aspect of how a story is expressed. Discourse generation refers to structural transformations from a story structure to a discourse one, and it is modeled based on the narrative discourse theory by Genette (1980). Jauss's theory is applied to the controlling model of continual discourse generation.

Modeling of Discourse Generation Based on Genette's Theory

Both story and discourse are described through the conceptual representation of a tree that consists of "events" and "relations." Figure 2 shows diagrams of story and discourse structures. An event is the basic unit of constituting a story or discourse, whereas a relation semantically combines the child nodes, i.e.,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Story and discourse structures



the events or sub-structures. In the figure, events and relations are indicated by "E*" and "$*", respectively. An event is represented with a case structure of verb concept. Figure 3 is an example of event representation. The system includes a simple natural language generator for transforming a conceptual representation to Japanese text. For instance, the event representation in Figure 3 is transformed into the sentence "蛇が皇女を誘拐した" ("a snake abducted a princess").

Structural transformation techniques, or "discourse techniques," are defined by applying the narrative discourse theory by Genette (1980). He categorized various classes of discourse structures through the analysis of a novel. His discourse classes consists of the following three broad categories: "tense," which concerns the relationship between the time of the story and that of the discourse; "mood," which concerns the modality for regulating narrative information; and "voice," which concerns the relationship among narration, story, and discourse. Each category is further divided into subcategories.

Through the computational formalizations of part of Genette's subcategories, the authors implemented the following 13 discourse techniques: "external analepsis," "complementary analepsis_ellipsis," "complementary analepsis_paralipsis," "repetitive analepsis," "external prolepsis," "complementary prolepsis_ellipsis," "complementary prolepsis_paralipsis," "repetitive prolepsis," "achronie," "pause," "implicit ellipsis," "repeating," and "paralipsis." In general terms, they include the following transformation types: temporal ordering of events, adding supplemental events for the story, omitting events, repeating events, and inserting descriptions of characters, objects, or locations. For instance, in the discourse structure of Figure 2, E2 is omitted by "implicit ellipsis" and E3 is reordered by "complementary analepsis_ellipsis."

Figure 3. Representation of an event (the words in angle brackets are English translations)

(event 誘拐する(1) <abduct>
(*type* action)
(*ID* 1)
(*time* (1 2))
(*agent* 蛇<snake>)
(*counter-agent* 皇女<princess>)
(*object* nil)
(*instrument* nil)
(*location* 庭<garden>)
(*from* nil)
(*to* nil))

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

One or more discourse techniques can be used in combination with generating a discourse structure. Different discourse structures are generated depending on the selected techniques and target part—i.e., a node of the tree structure—of each technique. Because each target part is determined randomly with several conditions, output structures by the same combination of techniques have relatively small differences.

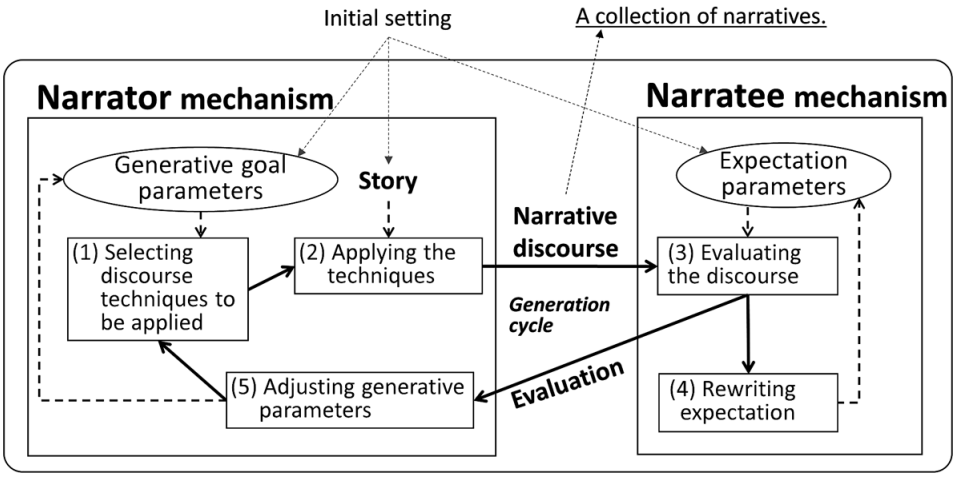
Continual Discourse Generation Based on Jauss's Reception Theory

Jauss (1982) developed seven theses for the essence of his reception theory. The authors focus mainly on theses II and III in modeling the continual generation process. Jauss's thesis II describes the "horizon of expectation" as being the reader's preliminary knowledge for ranking a literary work in the context of reading experience, and this can be acquired by known genres and exposure to famous literary works. In thesis III, Jauss proposes that readers understand literary artistic features as the distance between the work and their "horizon of expectation." If the distance is small, the work has strong recreational purpose and a familiar beauty. Those with large distances have various and distinguished influences for the reader: to one reader the work may be highly regarded creative work, but it could be a waste of time to another reader. The horizon can change so that the impression of a work becomes ordinary for a reader.

The authors modeled Jauss's historical view as an iterative mutual action, referred to as "generation cycle," between the two mechanisms of narrative generator, or "narrator," and evaluator, or "narratee." Figure 4 outlines the process flow. In the system, authors and readers in Jauss's theory respectively correspond to the narrator and narratee mechanisms. The notion of "horizon of expectation" is objectivized as the narratee's expectation represented as a set of parameters.

The input consists of a story, initial parameters of the narrator's generative goal and the narratee's expectation, total number of generating discourses, and other configurations. The parameters for the generative goal and expectation commonly consist of the following ten structural features: "supplement," "complexity," "suspense," "length," "hiding," "descriptiveness," "repetition," "diffuseness," "implication," and "temporal-independency." Each parameter takes a value of 1 (small), 2 (medium), or 3 (large). Through an iterative generation process, the system outputs a series of discourses. A text by the natural language generator and log data accompany each discourse.

Figure 4. Jauss-based narrative discourse generation system



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The generation cycle continues according to the interaction between the narrator and narratee. First, the narrator generates discourse from the input story based on the generative goal. The narrator has a set of rules for controlling the application of discourse techniques based on the generative parameters. The narratee then evaluates the results by comparing his/her expectation parameters and the narrator's generative ones. In the next cycle, the narrator applies corrections according to the evaluation in order to achieve the narratee's satisfaction. The process, however, eventually reaches a point where the narratee becomes tired or his/her satisfaction begins to fall. When this happens, "deviation" occurs and the narrator abandons a portion of the old generative parameters and moves to the generation of a discourse with the new generative goal. The narratee's expectations also change according to the reconstruction.

System Behavior

The authors show an example of executing the system for 10,000 steps. In Figure 5, the upper and lower texts respectively show the input story and generated discourse by English translations of Japanese sentences by the natural language generator. The authors added explanations on the discourse structure in angle brackets.

In order to analyze all generated discourses, the authors defined metrics to quantify the structural features of each discourse, referred to as discourse feature values (DFVs) (see Table 1). These features correspond to the aforementioned ten parameters in the system. These metrics are incorporated in the above system, and DFVs for each generated discourse are automatically logged in the generation cycle. Figure 6 shows the transitions of each DFV from step 1 to 1000. There are two types of changing patterns: micro level changes based on the same generative parameters that occur in each step, and macro level changes caused by the narrator's deviations.

Issues

The central problem of the above model is that the narrator and narratee mechanisms are disconnected from accumulated discourses through the generation cycle. In each step, the narrator generates a discourse from the input story according to a set of absolute parameters represented in three-point scale.

Figure 5. Generation example

Natural language expression of the input story (English translation):
A snake abducted a princess at a garden. An older woman sang a grief-stricken song at the garden. Ivan resolved to search the princess at a town. Ivan departed from the town. Ivan battled with the snake at the snake's land. Ivan injured his arm by the snake at the snake's land. Ivan won against the snake at the snake's land. Ivan abducted the princess at the snake's land. Ivan departed from the snake's land. The snake flied at the snake's land. The snake chased Ivan at the snake's land. Ivan hid himself on the back of a rock at the snake's land. The princess knew Ivan's wound at the snake's land. Dwarfs built a palace at the town. Ivan lived at the palace. Ivan married with the princess at the palace.

Natural language expression of a generated discourse (English translation):
A snake abducted a princess at a garden. [<future> The princess divorced with Ivan. The princess returned home from a palace.] An older woman sang a grief-stricken song at the garden. Ivan resolved to search the princess (at a town). Ivan departed from the town. [<future> Ivan won against the snake at the snake's land. [<future> Ivan lived at a palace.]] Ivan battled with the snake (at the snake's land). Ivan injured his arm by the snake. Ivan won against the snake. Ivan abducted the princess. The princess knew Ivan's wound. Dwarfs built a palace at the town. Ivan lived at the palace. Ivan married with the princess. [<backward> Ivan battled with the snake (at the snake's land). Ivan injured his arm by the snake. Ivan won against the snake.]

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. DFVs metrics

DFVs	Metrics
DFV_1: Supplement	Number of "events" positioned outside the story's temporal range that provide complementary information.
DFV_2: Complexity	Number of "relations" required to define temporal order transformations. A relation nested under one or more relations of temporal ordering is additionally weighted in proportion to the depth of nesting.
DFV_3: Suspense	Number of "events" moved from a backward position in the story and not accounted for before.
DFV_4: Length	Total number of "events" and "descriptions."
DFV_5: Hiding	Number of "events" contained in the story, but not accounted for in the discourse.
DFV_6: Descriptiveness	Number of "descriptions."
DFV_7: Repetition	Sum of the numbers of repetition of each "event."
DFV_8: Diffuseness	Subtraction of DFV_5 ("hiding") from DFV_7 ("repetition"), i.e., "diffuseness" is increased by "repetition" and decreased by "hiding."
DFV_9: Implication	Number of "relations" required to define the prospective temporal order transformation.
DFV_{10}: Temporal-independency	Number of "relations" using the "achronie" technique, which inserts events that do not have any temporal connection with the story.

Potential narratives to be generated are prelimited depending on finite combinations of parameters, independently of the past history of the system actions. Actually, the authors observed that the difference or novelty of the generated narrative from those prior, based on DFVs, decreases gradually through the generation cycle. Because the primary goal of NBNG is to continually increase narrative diversity, this model is not adequate for NBNG.

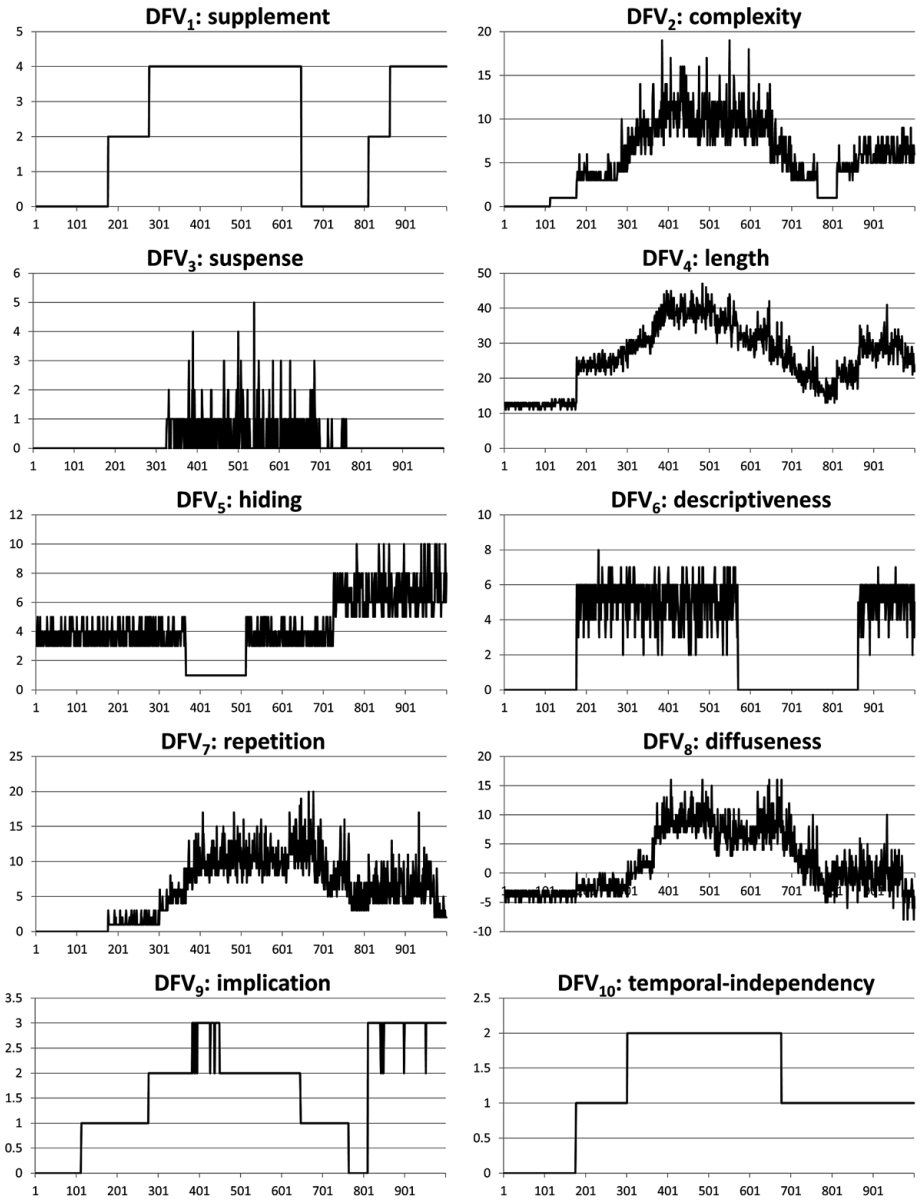
A root cause of the above problem is in the functional relationship between the narrator and narratee, i.e., the modeling of how the generation and reception interact in a continual narrative production process. The above system is implemented as the direct interaction of the generation and evaluation of the two agents. This is a personal interaction, rather than a social or historical process. However, the point of Jauss's theory is to provide the role of dynamically forming the "horizon of expectation" to the reception side (i.e., readers) and construct a literary history based on the historical change of the horizon. Namely, each work is historically positioned based on the relationship to the horizon at the time. In this manner, the reception side plays a central role in the construction of literary history. This idea is applicable to the model of production and organization of narratives in SONGS, but is missed in the above model. The authors designed a new model for NBNG mainly by modifying this problem.

Designing NBNG

A main modification from the previous model is to place the recipient side not as the direct evaluation mechanism for the narrative, but as the mechanism for organizing narratives in CNs. In particular, the recipient side attaches external attribute information to a generated narrative in order to position the narrative into the organization of CNs. On the other hand, the generation side generates new narratives with reference to CNs, based on a relative goal as difference from prior narratives. This way, an indirect interaction between the narrative generation and reception arise via CNs. In addition, the authors mainly apply the following modifications. First, the users are explicitly associated with the model. Second, the system treats the total process of narrative generation, not only discourse generation, but also story gen-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

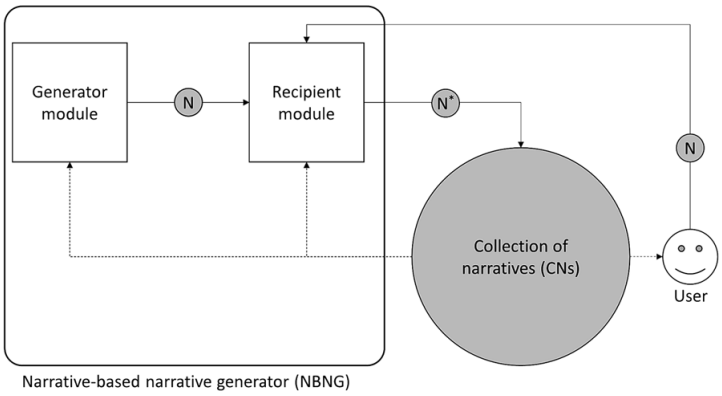
Figure 6. Change of generated discourses through generation cycle



eration. Third, the authors completely revise the terms from the previous model. In particular, because the mechanisms of narrative generation and reception in SONGS are positioned as being the outside agents of a narrative text, the authors change the names narrator—"[t]he one who narrates, as inscribed in the text" (Prince, 2003, p. 66)—and narratee—"[t]he one who is narrated to, as inscribed in the text" (Prince, 2003, p. 57)—to "generator" and "recipient," respectively. Figure 7 shows the overall architecture of NBNG, including the relationships to CNs and users. In the figure, "N" indicates a narrative and "N*" means a narrative accompanied by the external attribute information by the recipient. The rest of this section describes the main elements of NBNG.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. NBNG architecture



Narrative (N)

In CNs, each N corresponds to a conceptual representation as the formal description of content and structure of a given narrative. In the development of INGS, the authors defined a structural framework of narrative that consists of the three dimensions of story structure, discourse, and expression. A story represents the content information of a narrative and discourse is the structural aspect of how a story is expressed. A story is further divided into two types of structures: a temporal sequence of events and the background states. The latter corresponds to the structure of the story world that defines the information of characters, objects, and locations in a story.

According to the above framework, the authors defined a frame format, referred to as N-frame, for representing N in CNs (see Figure 8). Each italic element in the figure indicates a slot, a composing element of N. Namely, N is composed of conceptual representations of the story world (*StoryWorld*), chronologically structured events (*Story*), and discourse structure (*Discourse*). The representation method of each element is basically defined based on that in INGS. Note that in SONGS, the narrative expression generation is a function of the interface part.

Collection of Narratives (CNs)

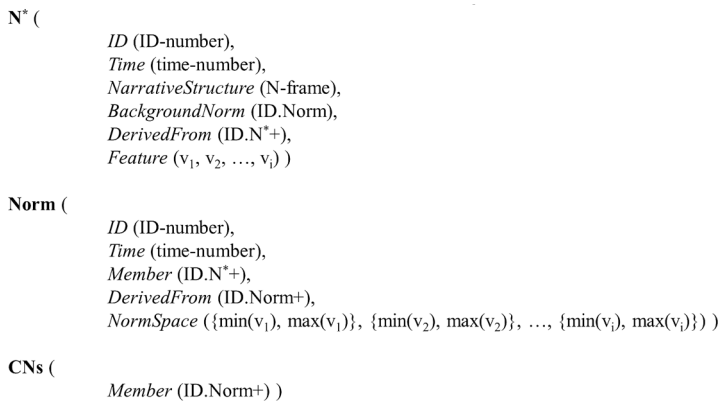
CNs are organized based on the two substructures of "N*" and "Norm." Each element is defined by a frame format referred to as N*-frame, Norm-frame, and CNs-frame (see Figure 9). In the figure, each italic element indicates a slot, tailing plus ("+") means "one or more," and a period (".") between two elements means "of"—e.g., ID.N* means *ID* of N*.

Figure 8. N-frame

N (
StoryWorld (conceptual representation of story world),
Story (conceptual representation of chronologically structured events),
Discourse (conceptual representation of discourse structure))

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. N-frame, Norm-frame, and CNs-frame*



In N*-frame, the *NarrativeStructure* slot contains N, and other slots correspond to the external attributes. *ID* indicates the unique number for N. *Time* marks the actual date when N is generated. It means the temporal position on the timeline of the continual production process. *BackgroundNorm* indicates the Norm in which N* is placed. *DerivedFrom* links prior N(s) used in generating N. *Feature* takes a set of feature values of N. It is similar to DFVs for the Jauss-based system.

Each Norm forms a subset of N*s that have some commonality in their features. In other words, it is an intermediate unit of CNs that corresponds to the notion of "horizon of expectation" in Jauss's theory. A Norm is used as the basis for both processes of the generator's narrative production and the recipient's narrative positioning. Norms and CNs are automatically updated when a new N* is added to CNs. In Norm-frame, *ID* and *Time* respectively indicate the unique number and temporal position. *Member* lists the contained N*s. *DerivedFrom* links the prior Norm(s) from which the Norm is derived. *NormSpace* represents the ranges of the feature values in the member N*s. CNs are only a set of all Norms.

Overall, CNs are organized based on the following four formation principles:

- **Timeline:** N*s and Norms are placed on a timeline (*Time*).
- **Norm-Based Hierarchy:** CNs are hierarchically grouped by Norms, in particular, *Member* in Norms and CNs and *BackgroundNorm* in N*s.
- **Derivative Network:** N*s and Norms are respectively linked based on their derivative relationships, in particular, *DerivedFrom* in N*s and Norms.
- **Feature Space:** N*s and Norms are positioned in a feature space based on their feature values, in particular, *Feature* in N*s and *NormSpace* in Norms.

Process Flow

Figure 10 shows a total picture of the system actions through generator to recipient.

The generator produces a new N through the following procedure:

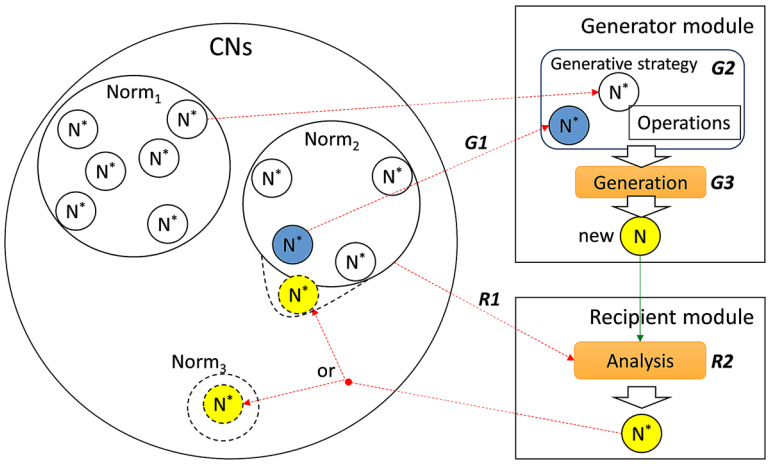
G1 - Setting a Basing N*: The generator selects N* for the basing point of generating a new N.

G2 - Planning a Generative Strategy: The generator makes a plan for how to recompose the basing N*.

First, with reference to the *BackgroundNorm* of the basing N*, the generator sets a basic objective

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. Total picture of system action



of either normative generation or deviation from the Norm. Deviation means an action to attempt to produce a new type of N by breaking the existing Norm. Based on the basic objective, the generator makes the following two choices: *what*, or target elements in the basing N*—i.e., constituents of *StoryWorld*, *Story*, and *Discourse*—and *how*, or operations for each target—e.g., transformation, deletion, addition, substitution, abstraction, concretization, and synthesizing with other N*s.

G3 - Generating N: The generator recomposes the basing N* in order to generate a new N based on the generative strategy.

The new N is sent to the recipient. In addition, the generator may send additional information to the recipient on background information of the generation process, i.e., the basing N* or *BackgroundNorm*. This helps the recipient to position N into CNs.

Next, the recipient composes N* through the following procedure:

R1 - Setting a basing Norm: The recipient selects a basing Norm in receiving the N based on the generator's additional information.

R2 - Composing N*: The recipient analyzes N and fills the slots of N*-frame. The values of *ID*, *Time*, and *NarrativeStructure* are simply determined. *DerivedFrom* is determined by the Generator's additional information. The values of *Feature* are calculated by analyzing the structure of N. *BackgroundNorm* is determined based on *Feature*. When N is located near the basing Norm or the inside, based on the features, it is incorporated in the basing Norm—i.e., *BackgroundNorm* is the basing Norm. Otherwise, N is regarded as a deviation from the basing Norm and is positioned as a new Norm—i.e., *BackgroundNorm* is a new Norm to be created. Finally the composed N* is sent to CNs.

After the process, CNs and Norms are updated. When N* is placed in an existing Norm, the latter is expanded. Otherwise, a new Norm is created. The above process is iteratively executed at a regular time interval—e.g., once a minute. On the other hand, the user also can create a new N.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

3LNI: THREE-LAYERED NARRATIVE INTERFACE

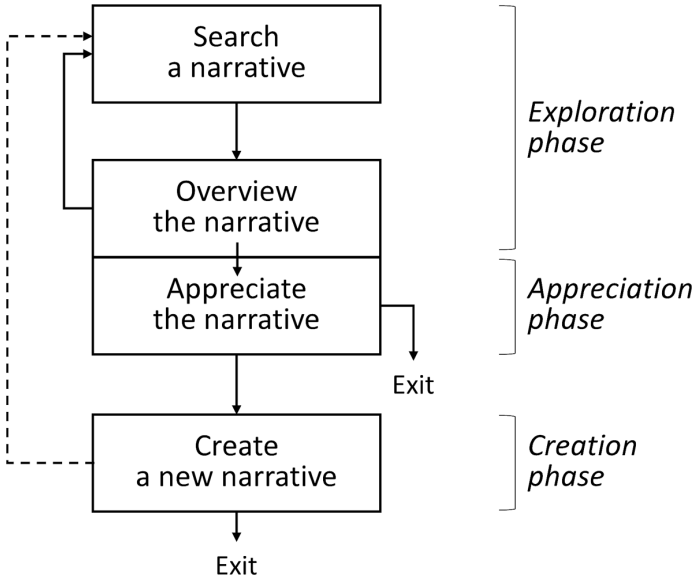
In this section, the authors propose the basic framework of 3LNI by considering a basic scenario for the use of SONGS. The central method is to visualize the following three aspects of CNs: organization of narratives, surface expressions of a narrative, and internal structures of a narrative. The authors also present a partial prototyping of 3LNI.

Basic Scenario of SONGS Usage

Figure 11 shows a basic scenario of user actions. This consists of the following three phases:

- **Exploration Phase:** First, users explore CNs to find a valuable narrative for their value or objective. In particular, users repeat the following two actions until a valuable narrative is found. First, the users search and select a narrative from CNs. Second, they overview the narrative in a short time. If the narrative seems valuable, the users proceed to the appreciation phase. Otherwise, they search another narrative.
- **Appreciation Phase:** The users closely read or see the selected narrative. This phase is always performed regardless of whether the user's objective is appreciation or creation. Subsequently, when the users have an objective or motivation of narrative creation, they proceed to the creation phase. Otherwise, the users exit from SONGS.
- **Creation Phase:** The users create an original narrative using the selected narrative as the material or draft. In the process, the users can obtain additional narratives or materials from CNs, as indicated by the dotted arrow line in Figure 11. Finally, the created narrative is added to CNs via the NBNG recipient module.

Figure 11. Basic scenario of SONGS usage



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

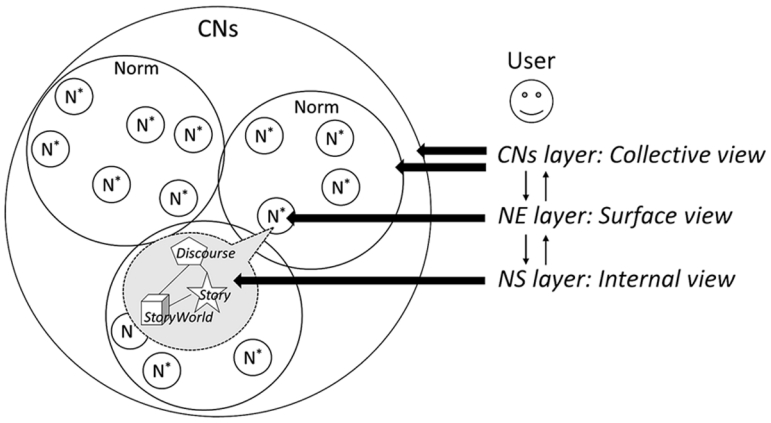
3LNI Framework

As the authors indicate in the above scenario, the interface needs to provide mainly the three functions of exploration, appreciation, and creation. The framework of 3LNI is designed to integrate these functions and the authors employ visualization—a well-known technique for providing human-friendly expressions in order to understand a huge amount of data or complex objects—for the central interface method. In particular, 3LNI divides and connects the following three perspectives on CNs: *CNs layer*, or collective view on whole and parts of CNs; *NE layer*, or surface view on a narrative expression; and *NS layer*, or internal view on a narrative structure. Via the above views, individual users share CNs, i.e., the object of co-creation, with NBNG and other users. Figure 12 shows the above concept of 3LNI.

The basic functions of each layer are as follows:

- **CNs Layer (Collective View):** This visualizes the organization of CNs into a two-dimensional space. Narratives or norms are represented as visual objects in the space. This includes various visualization methods depending on the CNs aspect or formation principle that is focused, in particular, timeline, norm-based hierarchy, derivative network, or feature space. The user explores and selects narratives with this macroscopic perspective.
- **NE Layer (Surface View):** This provides the surface expressions of a narrative. Because each narrative in CNs corresponds to a conceptual representation, surface expressions are generated in this layer. Natural language is a basic expression form. However, a narrative potentially allows various expressions depending on media—e.g., natural language, still or moving image, and music—and styles of expression. The user appreciates a narrative selected from the CNs layer.
- **NS Layer (Internal View):** This visualizes the internal structure of a narrative from multiple perspectives of the story world (*StoryWorld*), organization of events (*Story*), and discourse structure (*Discourse*). In addition, this layer includes editing functions that allow the user to visually recompose a narrative structure. The user sees the internal structures of a narrative from an analytical perspective and creates a new narrative by recomposing the visualized narrative structure.

Figure 12. 3LNI basic concept



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In the exploration phase, the user mainly employs the CNs layer to search a narrative, and the NE layer to overview it. In addition, the NS layer may provide multifaceted summaries of a narrative structure and help the user understand it quickly. Meanwhile, keyword retrieval is a popular method for searching information or content. However, the authors assume that, in many cases, the user has no clear requirement for narratives or cannot adequately represent the requirement through language. The authors will apply keyword retrieval methods as a subsidiary function for the exploration phase in the CNs layer.

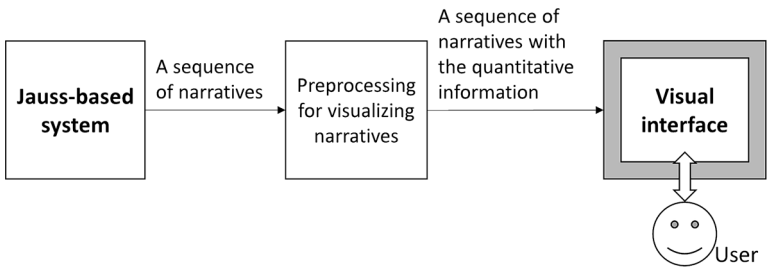
The above design positions the user creation phase in the NS layer. Although it is possible to design the creation phase at the surface expression level or NE layer, in order to share user-created narratives with NBNG, surface narratives must be transformed into a conceptual representation. This is a very difficult task for a computer. In addition, editing narrative expressions requires significant time and effort from the user. Hence, designing editing functions in the NS layer is a key issue for smoothly circulating the collective creation process between NBNG and users.

VISUAL INTERFACE FOR COMPUTER-GENERATED NARRATIVES

As a preceding idea for 3LNI, Akimoto and Ogata (2015a) proposed a visual interface for exploring narratives generated by the Jauss-based narrative discourse generation system introduced in the section "Narrative Discourse Generation System." The design concept partially overlaps the 3LNI framework. This subsection outlines the implemented visual interface in order to provide the readers with a concrete image of a part of 3LNI. The basic idea of the interface is to arrange and visualize many narratives into a two-dimensional space that allows users to explore and browse narratives based on their values or objectives. Figure 13 shows the relationship between the Jauss-based system and the visual interface.

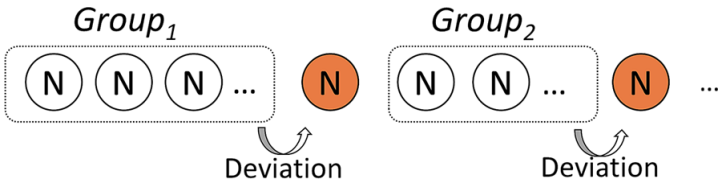
The composition of a series of narratives produced by the Jauss-based system is summarized in Figure 14. Such composition consists of a number of groups or subsets of narratives respectively formed by the same set of generative parameters. Namely, each group contains one or more narratives with similar features. In terms of the SONGS design, the entire set of narratives and groups correspond to CNs and norms, respectively. The number of narratives in each group is indefinite, but loosely controlled depending on the initial configurations of the Jauss-based system. A new group is produced via the narrator's deviation process that changes a parameter value. In Figure 14, "N" refers to a set of discourse structure, the natural language expression, and DFVs. Although the design of 3LNI includes the function of narrative expression generation in the NE layer, in this version, this is embedded in the Jauss-based system.

Figure 13. Process flow through narrative generation to visual interface



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 14. Composition of a series of generated narratives



The preprocessing part in Figure 13 calculates the statistics—minimum, maximum, and average for each DFV—of each narrative group and the entire set of narratives. The visual interface uses the group-based hierarchy and feature spaces to organize the narratives into a two-dimensional space. In particular, this consists of the following three perspectives:

V1: This plots the visual objects of the groups.
V2: This plots the visual objects of the narratives in a group.
V3: This represents a narrative by a natural language text and diagram of the structure.

In terms of the 3LNI framework, V1 and V2 correspond to the CNs layer, and V3 includes the NE and NS layers. Each perspective corresponds to an interface window, and these are mutually connected via the visual objects. Figures 15–17 show screenshots of each window for a series of 10,000 generated narratives with 271 groups. The input story is equal to that shown in Figure 5. In V1 and V2, two of the DFVs are used to arrange the discourses into a two-dimensional space. We indicate the features of the horizontal and vertical axes as DFV_x and DFV_y, respectively. Choosing different features for these axes results in different narrative arrangements. In Figures 15 and 16, DFV_x and DFV_y are DFV_4 (length) and DFV_2 (complexity), respectively. The user can change DFV_x and DFV_y using the "up"/"down" and "left"/"right" arrow keys, respectively.

In the V1 window, each group is represented by a blue rectangle with the group number. In the space, the left/lower ends and right/upper ends correspond to the minimum and maximum of DFV_x/DFV_y in the entire set of narratives, respectively. The area of a rectangle represents the ranges of DFV_x and DFV_y in the corresponding group. The users can move to the V2 window by clicking the left mouse button on a group number.

In the V2 window, each discourse is represented by a small yellow circle. The center of the space—i.e., the intersection of the horizontal and vertical axis lines—corresponds to the average of each DFV_x and DFV_y. Discourse objects are positioned based on their distance from the center, i.e., the averages. In addition, the deviated discourse is positioned as a small red circle. The user can move to the V3 window by clicking the left mouse button on a small circle. When the user clicks the right mouse button on a small circle, a sub-window for the miniature window L3 is displayed. In Figure 16, a sub-window of $Discourse_{3507}$ is shown. The user can return to the V1 window by clicking the "back" button.

The V3 window expresses a narrative using Japanese text and a diagram of the structure. The text in Figure 17 corresponds to the English expression of the generation example in Figure 5. The diagram represents the discourse structure using a graph. Each node (numbered circle) corresponds to an event, and the edges show the flow of the narration (for more detail, refer to Akimoto and Ogata (2015a)). The user can save the narrative text to a file by clicking the "save" button and return to the V2 window by clicking the "back" button.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 15. V1 window that contains 271 groups

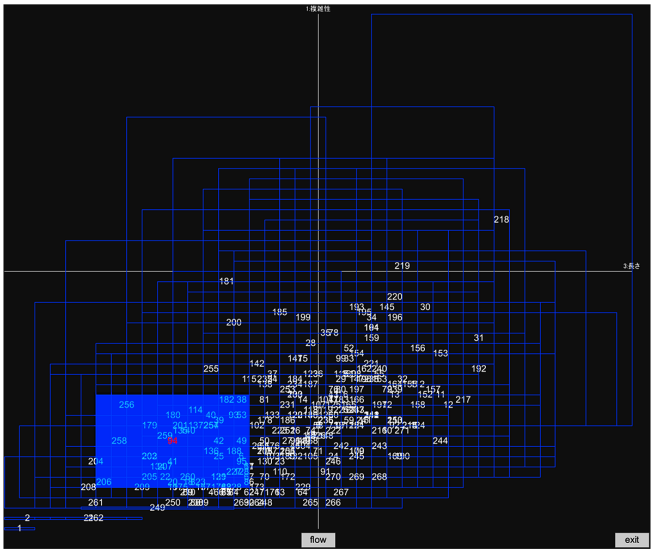
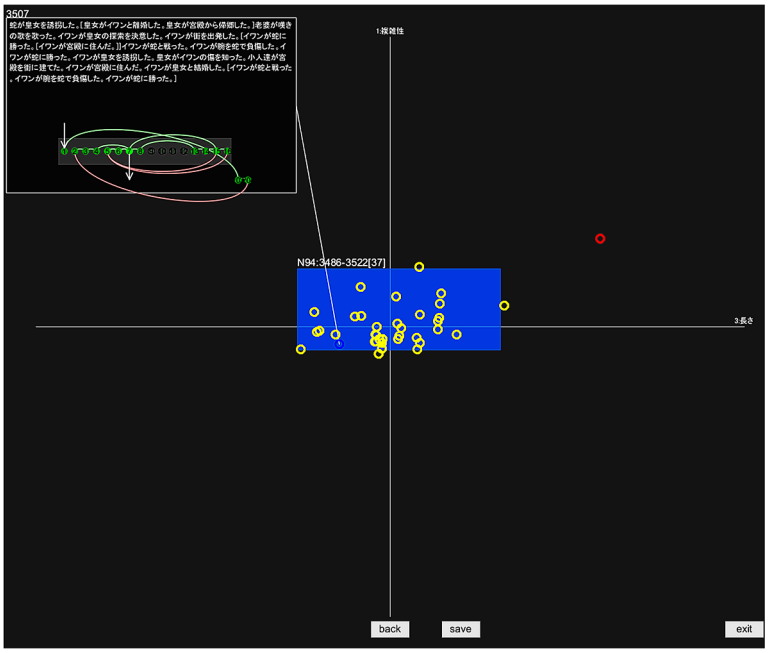


Figure 16. V2 window of group$_{94}$ that contains 37 narratives

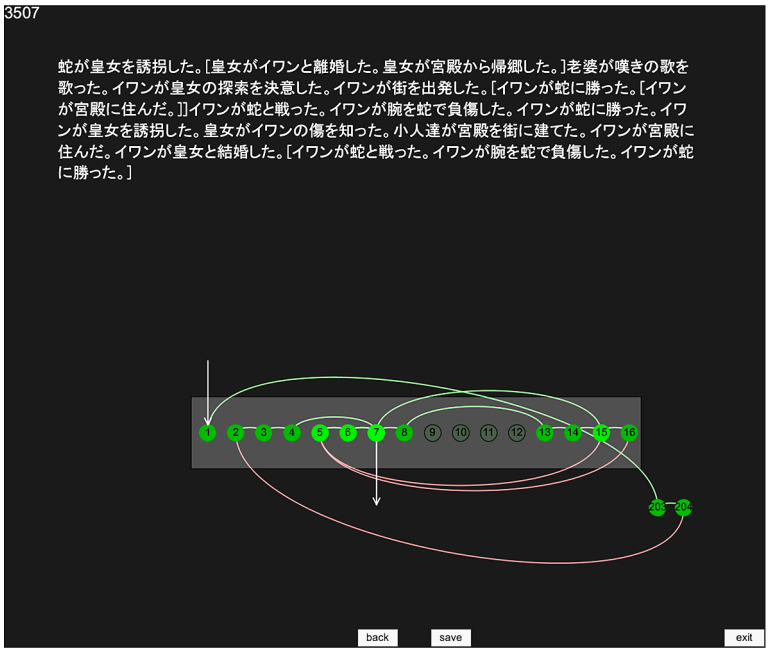


FUTURE RESEARCH DIRECTIONS

This chapter presented the core concepts and basic mechanism design of SONGS. The next challenge is to concretize the system design through the constructive cycle of experimental developments and the verifications or considerations of mechanisms. A main focus is to design a method for strategically

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 17. V3 window of narrative$_{3507}$



triggering various types of creative interactions between NBNG and human users. At the same time, it is important to discuss a macroscopic vision on the social impact of SONGS for determining the correct direction of system development. In the following subsections, first, the authors describe a plan for designing interaction between NBNG and human users. Second, the authors draw several long-term visions toward relating SONGS to the real world.

Toward Interaction Design between Generative Program and People

In order to produce a wealth of CNs, SONGS needs to be able to originate various types or modes of creative interactions between the generative program (i.e., NBNG) and many humans in the collective creation process. For this, it is important to design not only the mechanism of the program side, but also the behavior of the human side. If the program's narratives have an expansionary effect on human narrative creations—e.g., increasing the diversity of created narratives or triggering unique ideas—this will be a positive aspect of SONGS. In contrast, if the program's narratives do not give rise to such positive phenomena, the authors should consider the causes and redesign the system.

From the above perspective, the authors have conducted an experiment to investigate the effects of computer-generated narratives on human narrative creation using a narrative generation system that consists of the Jauss-based system and visual interface, i.e., preceding versions of NBNG and 3LNI (for details of the experiment, refer to Akimoto and Ogata (2015b)). In the experiment, 14 human subjects were divided into two groups: A and B. Each subject in group A used the system and created a variety of scenarios by obtaining narrative texts on the interface and then editing them to make them unique. Each subject in group B performed the same creation task, but without using this system. The authors then analyzed the structures of created scenarios and compared the differences between the scenarios

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

created by both groups. From the experiment results, the authors found a type of creation mode where humans strengthen the weakness of structural organization of computer-generated narratives. From a cognitive perspective, it seems that blanks or disconnections from the flow of a narrative text generated by the program strongly induce human creation in order to cause the narrative to be organized logically. This way, many scenarios were created that have unique storylines. A direction of future development is to design functions of the NS layer of 3LNI that allow this type of creation. However, the authors also need to consider strategies for triggering more diverse modes of the human creation processes.

Toward Relating SONGS to Real World

Next the authors take a long-term view and discuss possible impacts of SONGS to the real world. First, the central goal of SONGS is to activate creative activity and content industry in the real world. In order to consider this possibility, the authors discuss the relationship between the characteristics of the Japanese content industry based on Koyama (2009) and those of the SONGS design. Koyama described the characteristics and strengths of the Japanese-style content industry model by comparing with the Hollywood style. According to Koyama, the characteristics of the Hollywood style model can be summarized as follows: Groups of many professionals mass-produce huge works and maximize the benefits that arise from "hits." In this type of model, because the professionals excessively pursue the maximization of benefits while avoiding risks, diversity of subjects or narratives in the works tend to decrease. In contrast, the Japanese style model has the following characteristics: a) the boundary between professionals and amateurs is vague and there are many and diverse creators, b) the boundary of creators and consumers is vague and derivative, or collective creation activity is lively, c) the market consists mostly of small to medium-scaled works, d) there is a high degree of freedom for creation/expression activity, and so on. The unique and diverse creative culture in Japan is underpinned by these types of characteristics.

These characteristics of the Japanese style model are, at a fundamental level, largely similar to the design concepts of SONGS. In addition, SONGS drives the co-creative cycle through the computer program with a distinguishing narrative creativity. It has the potential of creating a new platform for expanding the creative culture and content industry. In order to obtain several users and gain powerful influence on the real world, it is naturally important to enhance the system's basic performance, such as the quality of produced contents, interface usability, degree of freedom for user creation, etc. The accumulation of steady efforts to achieve these is required.

Another possibility or secondary application of SONGS is for CNs to be a type of knowledge resource for intelligent information technologies in various fields. As described in the background section, narrative or story-based knowledge is focused as an important aspect of intelligence. Because each narrative in CNs consists of conceptual information by formal descriptions, CNs have high availability for computers or AI systems. For instance, lately, humanoid robots or AI applications have been active in various fields, such as entertainment, nursing care, education, etc. CNs can help the dialogue or storytelling tasks of such AI systems in various situations by providing diverse narratives. In addition, narrative knowledge or creativity is necessary for human-like intelligent machines. For instance, CNs or NBNG can be applicable to robots that compose high-level future goals or action plans, similar to the way in which humans makes decisions or life plans based on their narrative knowledge, such as past experiences, history, and so on.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

In this chapter, the authors proposed a design of SONGS that continually co-creates a wealth of narratives between a narrative generation program and many human users. This is based on the following four design concepts:

- SONGS uses narratives as comprehensive knowledge for producing new narratives.
- SONGS controls the continual narrative production process with relation to prior narratives.
- The generative program and many humans collectively produce narratives through the interaction of produced narratives.
- SONGS provides diverse narratives for people with diverse values or objectives.

In the next study stage, through experimental developments, the authors will elaborate the designs of NBNG for the narrative generation engine, and 3LNI for the visual interface. An important basis for the narrative co-creation is to formalize the narrative structure and its operations. The authors will apply their fundamental study of INGS to the modeling of structural representation and operations of narrative in SONGS. Although this chapter focused on a particular system, the basic framework or idea can be applicable to other creative activities of structural artifacts, such as music, architecture, and graphic arts. SONGS can be a form of connection between computational creativity and the real world.

REFERENCES

Akimoto, T., Imabuchi, S., Endo, J., Ono, J., Kurisawa, Y., Kamada, M., & Ogata, T. (2013). Development of a generation/expression system of narratives in the style of a folk tale, KOSERUBE Version 1. *Transactions of the Japanese Society for Artificial Intelligence*, *28*(5), 442–456. doi:10.1527/tjsai.28.442

Akimoto, T., & Ogata, T. (2012). A narratological approach for narrative discourse: Implementation and evaluation of the system based on Genette and Jauss. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 1272-1277), Austin, TX: Cognitive Science Society.

Akimoto, T., & Ogata, T. (2013). Towards a discourse mechanism in narrative generation system: Proposal of a system introducing narrative discourse theory and reception theory. *Cognitive Studies*, *20*(4), 396–420.

Akimoto, T., & Ogata, T. (2014). An information design of narratology: The use of three literary theories in a narrative generation system. *The International Journal of Visual Design*, *7*(3), 31–61.

Akimoto, T., & Ogata, T. (2015a). A visual interface for a collection of computer-generated narratives. *Information Engineering Express*, *1*(4), 43–52.

Akimoto, T., & Ogata, T. (2015b). An investigation of effects of computer-generated narratives on humans' narrative creations. In *Proceedings of 32nd Annual Meeting of the Japanese Cognitive Science Society* (pp. 462-471). Chiba, Japan: Academic Press.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Akimoto, T., Ono, J., & Ogata, T. (2013). Designing information contents with automatic narrative generation system. In *Proceedings of the 5th International Conference of International Association of Societies of Design Research* (pp. 5424-5435). Tokyo, Japan: Academic Press.

Anderson, T. S. (2015). From episodic memory to narrative in a cognitive architecture. In *Proceedings of 6th Workshop on Computational Models of Narrative* (pp. 2-11). Atlanta, GA: Academic Press.

Bae, B. C., Cheong, Y. G., & Young, M. R. (2011). Automated story generation with multiple internal focalization. In *Proceedings of 2011 IEEE Conference on Computational Intelligence and Games* (pp. 211-218). Seoul, South Korea. 10.1109/CIG.2011.6032009

Bringsjord, S., & Ferrucci, D. A. (1999). *Artificial intelligence and literary creativity: Inside the mind of BRUTUS, a storytelling machine*. Mahwah, NJ: Lawrence Erlbaum.

Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry*, *18*(1), 1–21. doi:10.1086/448619

Chambers, N., & Jurafsky, D. (2008). Unsupervised learning of narrative event chains. In *Proceedings of 46th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (pp. 789-797). Columbus, OH: Academic Press.

Chambers, N., & Jurafsky, D. (2009). Unsupervised learning of narrative schemas and their participants. In *Proceedings of Joint Conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing* (pp. 602-610). Suntec, Singapore. 10.3115/1690219.1690231

Chambers, N., & Jurafsky, D. (2010). A database of narrative schemas. In *Proceedings of Seventh International Conference on Language Resources and Evaluation* (pp. 1614-1618). Malta: Academic Press.

Chambers, N., Wang, S., & Jurafsky, D. (2007). Classifying temporal relations between events. In *Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions* (pp. 173-176). 10.3115/1557769.1557820

Fujita, T., & Ota, J. (2015). *Jinkoubutsu kougaku nyuumon* [Introduction to artifact engineering]. Tokyo, Japan: University of Tokyo Press.

Genette, G. (1980). *Narrative discourse: An essay in method* (J. E. Lewin, Trans.). Ithaca, NY: Cornell University Press. (Original work published 1972)

Gervás, P. (2009). Computational approaches to storytelling and creativity. *AI Magazine*, *30*(3), 49–62.

Gervás, P. (2015). *Computational drafting of plot structures for Russian folk tales*. Cognitive Computation; doi:10.100712559-015-9338-8

Gervás, P., Lönneker-Rodman, B., Meister, J. C., & Peinado, F. (2006). Narrative models: Narratology meets artificial intelligence. In *Proceedings of Satellite Workshop: Toward Computational Models of Literary Analysis, 5th International Conference on Language Resources and Evaluation* (pp. 44-51). Genoa, Italy: Academic Press.

Hamano, S. (2008). *A-kitekucha no seitaikei* [Ecosystem of architecture]. Tokyo, Japan: NTT Publishing.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Imabuchi, S., & Ogata, T. (2012). A story generation system based on Propp theory: As a mechanism in an integrated narrative generation system. In H. Isahara & K. Kanzaki (Eds.), *Advances in Natural Language Processing: Proceedings of the 8th International Conference on Natural Language Processing (LNAI)* (*vol. 7614*, pp. 312-321). Berlin, Germany: Springer. 10.1007/978-3-642-33983-7_31

Jauss, H. R. (1982). *Toward an aesthetic of reception* (T. Bahti, Trans.). Minneapolis, MN: University of Minnesota Press. (Original work published 1970)

Koyama, Y. (2009). Futatsu no kontentsu sangyou sisutemu [Two content industrial systems]. In H. Deguchi, H. Tanaka & Y. Koyama (Eds.), Kontentsu sangyou ron: Konkou to denpa no Nippon gata moderu [Studies of content industries: Mixture and diffusion of Japanese content] (pp. 61-90). Tokyo, Japan: University of Tokyo Press.

Kristeva, J. (1980). *Desire in language: A semiotic approach to literature and art*. New York, NY: Columbia University Press. (Original work published 1969)

Lang, R. R. (1999). A declarative model for simple narratives. In *Narrative Intelligence: Papers from the 1999 AAAI Fall Symposium* (Technical Report FS-99-01, pp. 134-141), North Falmouth, MA.

Li, B., Lee-Urban, S., Appling, D. S., & Riedl, M. O. (2012). Automatically learning to tell stories about social situations from the crowd. In *Proceedings of the Third Workshop on Computational Models of Narrative* (pp. 142-151). Istanbul, Turkey: Academic Press.

McIntyre, N., & Lapata, M. (2009). Learning to tell tales: A data-driven approach to story generation. In *Proceedings of Joint Conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing* (pp. 217-225). Suntec, Singapore. 10.3115/1687878.1687910

Meehan, J. R. (1980). *The metanovel: Writing stories by computer*. New York, NY: Garland Publishing.

Montfort, N. (2007). *Generating narrative variation in interactive fiction* (Doctoral dissertation). Available from ProQuest. (AAI3271851)

Nakashima, M., & Ogata, T. (2006). Narrative generation system and intertextuality: A consideration on the concept and prototype systems. In *Proceedings of the 20th Annual Conference of the Japanese Society for Artificial Intelligence* (2E2-2). Tokyo, Japan: Academic Press.

Niwango. (n.d.). *Niconico*. Retrieved from http://www.nicovideo.jp/

Ogata, T. (2002). Expanded literary theory: Cognitive/computational expansion of literary theories and narratology. In *Proceedings of 17th Congress of the International Association of Empirical Aesthetics* (pp. 163-166). Osaka, Japan: Academic Press.

Ogata, T. (2004a). A computational approach to literary and narrative production: Toward computational narratology. In *Proceedings of 18th Congress of the International Association of Empirical Aesthetics* (pp. 509-516). Lisbon, Portugal: Academic Press.

Ogata, T. (2004b). Toward the consistent simulation of narrative discourse based on narratology. In *Proceedings of the Ninth International Symposium on Artificial Life and Robotics* (*Vol. 2*, pp. 579-584). Oita, Japan: Academic Press.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ogata, T., & Terano, T. (1991). Explanation-based narrative generation using semiotic theory. In *Proceedings of Natural Language Processing Pacific Rim Symposium 1991* (pp. 321-328). Academic Press.

Okada, N., & Endo, T. (1992). Story generation based on dynamics of the mind. *Computational Intelligence*, *8*(1), 123–160. doi:10.1111/j.1467-8640.1992.tb00341.x

Peinado, F., & Gervás, P. (2005). Creativity issues in plot generation. In *Workshop on Computational Creativity, Working Notes, 19th International Joint Conference on A.I.* (pp. 45-52), Scotland, UK: Academic Press.

Pemberton, L. (1989). A modular approach to story generation. In *Proceedings of the Fourth Conference on European Chapter of the Association for Computational Linguistics* (pp. 217-224). 10.3115/976815.976845

Prince, G. (2003). *A dictionary of narratology* (Revised Ed.). Lincoln, NE: University of Nebraska Press.

Propp, V. Y. (1968). *Morphology of the folktale* (L. Scott, Trans.). Austin, TX: University of Texas Press. (Original work published 1928)

Quesenbery, W., & Brooks, K. (2010). *Storytelling for user experience: Crafting stories for better design*. New York, NY: Rosenfeld Media.

Riedl, M. O., & Bulitko, V. (2013). Interactive narrative: An intelligent systems approach. *AI Magazine*, *34*(1), 67–77.

Riedl, M. O., & Young, R. M. (2010). Narrative planning: Balancing plot and character. *Journal of Artificial Intelligence Research*, *39*, 217–267.

Riesbeck, C. K., & Schank, R. C. (1989). *Inside case-based reasoning*. Mahwah, NJ: Lawrence Erlbaum.

Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science*. Waltham, MA: Academic Press. doi:10.1016/B978-0-12-108550-6.50013-6

Schank, R. C. (1990). *Tell me a story: Narrative and intelligence*. Chicago, IL: Northwestern University Press.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Mahwah, NJ: Lawrence Erlbaum.

Singh, P., Barry, B., & Liu, H. (2004). Teaching machines about everyday life. *BT Technology Journal*, *22*(4), 227–240. doi:10.1023/B:BTTJ.0000047601.53388.74

Swanson, R., & Gordon, A. S. (2012). Say Anything: Using textual case-based reasoning to enable open-domain interactive storytelling. *ACM Transactions on Interactive Intelligent Systems, 2*(3), Article 16.

Szilas, N. (2015). Towards narrative-based knowledge representation in cognitive systems. In *Proceedings of 6th Workshop on Computational Models of Narrative* (pp. 133-141). Academic Press.

Tanaka, H. (2011). A-kitekuchuraru ko-dingu. [Architectural coding]. In Y. Karasawa, H. Tanaka, C. Dominick, R. Fujimura, & S. Matsukawa (Eds.), *Sekkei no sekkei* [Design of design] (pp. 53–110). Tokyo, Japan: INAX Publishing.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Turner, S. R. (1994). *The creative process: A computer model of storytelling and creativity*. Mahwah, NJ: Lawrence Erlbaum.

Tynyanov, Y. (1982). Bungaku no sinka. [Evolution of literature]. In *Rosia forumarizumu bungaku ronshuu 2* [A collection of literary essays in Russian formalism 2] (T. Kodaira, Trans.) (pp. 104–130). Tokyo, Japan: Serika Shobo. (Original work published 1929)

Yamaha. (n.d.). *What is VOCALOID*. Retrieved from http://www.vocaloid.com/en/about/

KEY TERMS AND DEFINITIONS

3LNI: Three-layered narrative interface, the user interface for accessing CNs. 3LNI provides multiple perspectives on CNs: collective, surface, and internal.

CNs: Collection of narratives. Narratives produced in SONGS are accumulated in this part and shared by the narrative generation program (NBNG) and human users.

Co-Creation: Cooperative process for producing novel and valuable artifacts or information through mutual action among two or more agents.

Computational Creativity: Research field that addresses computer-based modeling, design, simulation, and implementation of creativity. Research subjects include various creative areas, such as arts, entertainment, artifacts, content, ideas, and knowledge.

Narrative: Product that represents one or more events by some media for humans, or some formal descriptions for computers. In our study, a narrative is constructed as the totality of a story, discourse, and expression.

Narrative Discourse: The structural aspect of how a story is expressed. In our study, a discourse is represented not as surface expression, but as a formal description.

Narrative Expression: Surface expression of a narrative by a medium such as natural language, still or moving image, and music, or their combination.

Narrative Generation System: Computer program that produces some types of narratives.

NBNG: Narrative-based narrative generator, the generative program in SONGS. The narrative generation is modeled as a cyclic process for composing new narratives in relation to previously generated narratives in CNs.

Story: The content information of a narrative, in particular, a chronologically ordered sequence of events.

Story World: The background world information of a story, including characters, objects, places or spaces, timelines, etc.

This research was previously published in Computational and Cognitive Approaches to Narratology edited by Takashi Ogata and Taisuke Akimoto; pages 91-117; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 37
Conceptual Graphs Based Approach for Subjective Answers Evaluation

Goonjan Jain
Jawaharlal Nehru University, India

D.K. Lobiyal
Jawaharlal Nehru University, India

ABSTRACT

Automated evaluation systems for objective type tests already exist. However, it is challenging to make an automated evaluation system for subjective type tests. Therefore, focus of this paper is on evaluation of simple text based subjective answers using Natural Language Processing techniques. A student's answer is evaluated by comparing it with a model answer of the question. Model answers cannot exactly match with the students' answers due to variability in writing. Therefore, researchers create conceptual graphs for both student as well as model answer and compute similarity between these graphs using techniques of graph similarity measures. Based on the similarity, marks are assigned to an answer. Lastly, in this manuscript authors compare the results obtained by human graders and the proposed system using Pearson correlation coefficient. Also, comparison has been drawn between the results of proposed system with other existing evaluation systems. The experimental evaluation of the proposed system shows promising results.

INTRODUCTION

Tests are an indispensable part of our education system and are used to gauge student's knowledge. Two basic categories of tests are Objective and Subjective tests. Objective tests require student to only select a correct response from several alternatives or utmost to supply a word or short phrase to answer a question. On the contrary, subjective tests require test taker to form an original answer based on their own understanding of topic.

DOI: 10.4018/978-1-7998-0951-7.ch037

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Subjective tests are useful in measuring educational objectives such as analysis, and synthesis. They permit freedom of response that in turn allow students to present their ideas and knowledge they have in the subject area covered by the question. With this, subjective tests also promote development of problem solving skills and eliminate guessing.

Despite the unique benefits of subjective tests, their automation is very limited when compared with automation of objective tests. Major reason for this is that there are downfalls when it comes to marking subjective exams electronically. They have lots of complexities like, different ways of answering the same question; some question requires diagrams or equations. Some questions require answer in more than one word and length of answer may even vary from individual to individual. On the contrary, it is simple and efficient to mark objective item tests because a computer can easily process multiple choice responses.

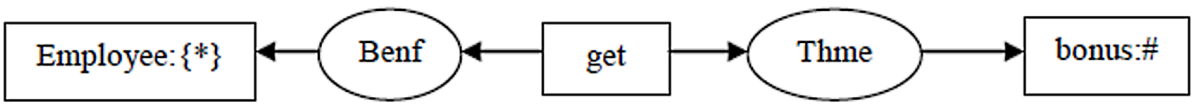
In this paper, authors propose a unique automatic grading system that can assign grades to short subjective answers provided by student. The system does so by using a model or template answer as a reference. Proposed technique is unique in itself because none of the existing techniques leverages semantics of the text to evaluate it. This limitation significantly hampers effectiveness of technique as questions can be answered in many ways based on candidate's understanding, strength of vocabulary and expressive nature. In addition to this, considering only syntax and morphology will not be sufficient to accurately mark a subjective answer. A correct answer can take many shapes as long as its semantic and syntactic structure is correct. Therefore, it is very important to understand the way the answers are paraphrased for representing the importance of words. This can be achieved by understanding semantics of text and the authors have employed Conceptual Graphs for this.

"Conceptual Graphs (CG) are a system of logic which is based on the existential graphs, the semantic networks of artificial intelligence. Conceptual graphs are described as the logical form that state relationships between concepts and thus represents a meaning." (Sowa, 1984) They were originally developed as a representation that could help in understanding the meaning of natural language text. They have been successful in capturing the nuances in natural language along with the ability to be implemented in computer software. Conceptual Graph is a knowledge representation technique that expresses meaning in a precise, readable and computationally tractable way and therefore is an efficient tool for information and knowledge processing.

The CGs are labelled graphs and have two types of nodes – concept and relation nodes. A concept node is a discrete unit of perception and represents entities, attributes, states, and events (Sowa, 1984). There are two types of fields in a concept node - type and referent. The type field specifies the class of entity represented by the concept and an instance of the concept type is identified by a referent. The interconnection of concept nodes is shown by relation nodes. Each conceptual relation has n arcs which are linked to some concept. Every conceptual relation has a relation type and a valence which tells the number of arcs belonging to the relation. Consider an example sentence "Employees get the bonus". The CG representation for this sentence is shown in Figure 1.

Since CGs are a good way to capture semantics of text, authors have used an NLP approach along with conceptual graphs to understand the semantics of students' answers and then evaluate it.

Figure 1. Conceptual Graph Representation for Sentence "Employees Get the Bonus"



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Next section presents a brief description of some of the approaches that have been proposed for automatic evaluation of subjective answers. It also summarizes the work that has been done in the field of conceptual graphs. Detailed design and implementation methodology of proposed work is explained in the following section. The results of evaluation of proposed work by undertaking various experiments and its comparison with other available methods are reported in later section. Finally, the work has been concluded along with its possible future extensions.

RELATED WORKS

Work has been done in automatic scoring of answers in recent years which is evident from development of various systems like TCT (Larkey, 1998), Intelligent Essay Assessor (Foltz, Laham, & Landauer, 1999), Automark (Mitchell et al., 2002), C-rater (Leacock, Chodorow, 2003), OXFORDUCLES (Sukkarieh, Pulman, & Raikes, 2003), JESS (Ishioka, Kameda, 2004) etc. Various techniques have been employed by these systems like Latent Semantic Analysis (LSA) or any variant of it, text mining, Natural Language Processing (NLP), Information Extraction (IE), keyword analysis, BLEU algorithm or a hybrid of any of the above.

Auto-assessor developed in year 2011 by Laurie and Maiga uses NLP techniques. Both the student's answer and correct answer are reduced to their canonical forms and these forms are compared with each other to determine their level of equivalency and accordingly scores are awarded for student's response. Ade-Ibijola et al. (2012) developed Expert System for Essay Scoring (ES4ES) for scoring free text answers. It uses a built-in Inference Engine designed with Shallow Natural Language Processing Techniques using Tokenization, Statistical Keyword Analysis and Domain-Specific Dictionary. It further uses a Fuzzy-Scoring Model to assess student's answers to open-ended questions and thus computes the relative correctness of student's answers to model answers. A system developed by A. Dhokrat et al... in 2012 generates a sequence state transition for candidate's answer which is then compared with the model answer. Exact match or at least 50% match in key terms marks the answer to be correct. Papri Chakraborty (2012) in proposed that an answer is considered to be correct if it is syntactically and morphologically same as the model answer. A. Kaur et al. in 2013 proposed an approach that converts reference and student's answers into a graphical form and then matches their nodes and labels by applying different similarity measures. These measures include partial and full string match and use of WordNet measures like Similar_to, Is_apart_of, Subset_of or Superset_of or Derived_from, Opposite_of etc. B. Kaiche et al. (2014) proposed a method for automatic evaluation of multiple sentence descriptive answers by using an algorithm based on Pattern Matching Technique. In 2014, Patil et al. proposed a system to evaluate multiple sentences answer by determining the semantic meaning of the answer. Tutor identifies the keywords from correct answer with their respective weights. Both students' and correct answers are processed and then these answers are compared word-by-word along with the POS tag and word position. Willis (2015) developed a system called *Amati* that develops a set of automatic marking rules which can be applied to larger question sets as well as for automatic marking. A unique superlative model has been proposed by Jayshankar et al. in 2016 that improves the accuracy by using word clouts.

Since the inception of Conceptual Graphs by Sowa in 1984, they have been highly used in different fields and for different purposes. Their capability to capture the semantics of natural language text makes them usable in varied domains. Various methods have been proposed by researchers for automatic generation of CG from natural language text. Sowa (1986) himself attempted one. In his work, he describes

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

a semantic interpreter where it accesses a lexicon of canonical graphs for each word of input and joins them to generate conceptual graph. In his book titled *Knowledge Representation*, Sowa (1999) explained logical representation of knowledge and various computational issues related to knowledge representation. This book concerns about knowledge representation using logic which includes vocabulary, syntax, semantics and rules of inference using conceptual graphs. Shih-Yao Yang, Von-Wun Soo in 2012 proposed to extract CG from plain text in patent claims by combining techniques such as finite state machines, POS tags, domain ontology and dependency tree. F. Rotella et al. (2013) proposed automated learning of CG with respect to missing or partial knowledge that does not allow extracting a full conceptual graph and only provides sparse fragments thereof.

A number of researchers have proposed computationally tractable CG matching algorithms. M.Montes et al. (2000) used variation of Dice Coefficient that took bipartite nature of CG into consideration to compare CGs. Montes et al. (2001) proposed another method for flexible comparison of CGs. They improved the flexibility by considering synonymy and subtype/supertype relationships between concepts and relations. Another matching algorithm based on Conceptual sub-graph weight self-adjustment was proposed by Zeng et al. in 2014. To compute the Conceptual Graph similarity, it gives homologous weight to subgraphs based on the proportion of how much information a subgraph contains in the whole Conceptual Graph.

DESIGN AND IMPLEMENTATION OF THE PROPOSED SYSTEM

In the proposed method, researchers find confidence level for a given answer by comparing student's answer to model answer and therefore check text relatedness between the two answers. As a first step, researchers perform preprocessing of the student's answer to correct spelling mistakes before performing other tasks. The model answer available is also processed using similar NLP tools. To understand semantics of answers, candidate's and model answers are represented in CG notation. CGs are compared by taking their corresponding structures into consideration and eventually producing an output in a form that shows semantic closeness between the candidate's answer and given model answer.

Proposed method is shown diagrammatically in Figure 2. The proposed method consists of mainly two tasks:

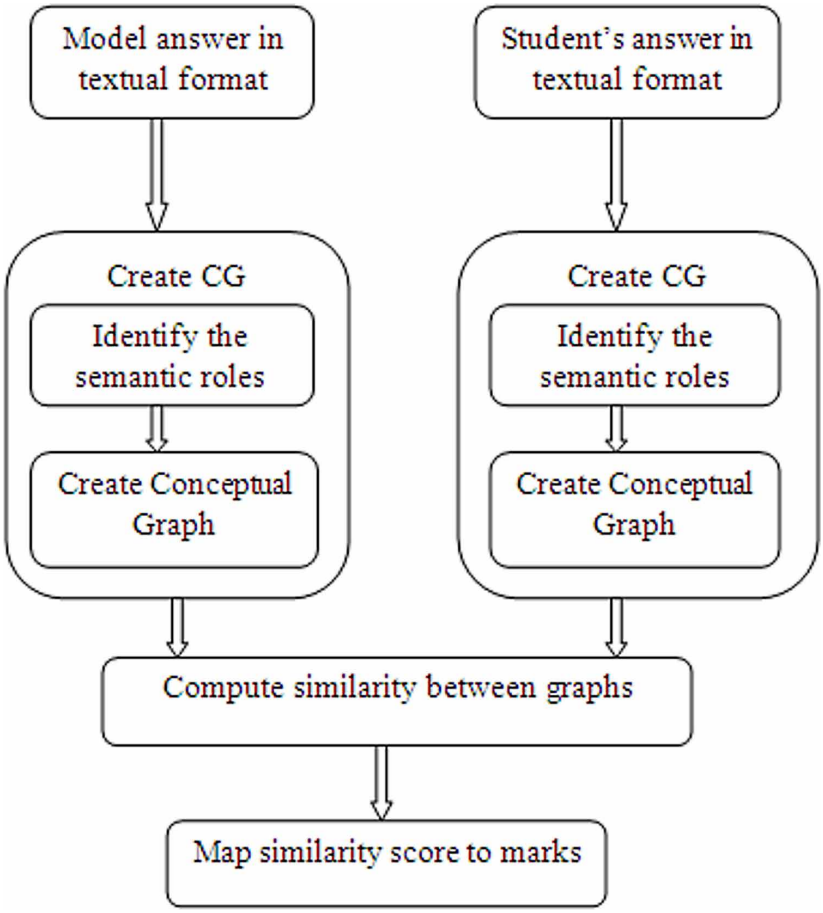
1. Create Conceptual Graphs for model and student answer.
2. Find similarity between these graphs using conceptual graph similarity measure and assign marks accordingly.

Creation of Conceptual Graph

The authors have used NLP techniques to create CG representation of text. A sentence is considered to be consisting of many fragments or syntactic patterns. These segments are related to each other through relationships identified by verbs. Therefore, to construct a CG for a sentence, program has to identify the syntactic patterns of a sentence and the semantic roles that each word in the sentence plays with respect to every identified verb. The proposed system uses an external Semantic Role Labeling (SRL) tagger to tag words of a sentence with corresponding semantic roles. In the process of identifying semantic roles, tagger performs syntactic parsing of a sentence and identifies constituents or syntactic patterns that constitute the sentence.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Flowchart of the Proposed System



The process of creating a CG from a text is as follows:

1. Use SRL tool to tag each word of the sentence with their corresponding semantic roles.
2. Create a CG for this semantic role tagged sentence.

Various SRL tools are freely available like Mate Tools[1], SEMAFOR[2], SENNA[3], ClearNLP[4] etc. to tag semantic roles in a sentence. In the proposed work, authors have used *"SENNA- Semantic Extraction using a Neural Network Architecture"* for tagging sentences with semantic roles.

SENNA is a software which outputs a host of Natural Language Processing (NLP) predictions: part-of-speech (POS) tags, chunks (CHK), name entity recognition (NER), semantic role labels (SRL) and syntactic parsing (PSG) (Collobert et al., 2011; Collobert, 2011). A sentence is taken as input and the algorithm gives semantic roles for each word of the input sentence for every identified verb. Several stages of processing are involved in SRL systems. In the first stage, a parse tree is produced and then parse tree nodes that represent the arguments of a given verb are identified, and finally these nodes are classified in to the corresponding SRL tags. *Chunking* (*CHK)* labels segments of a sentence with syntactic constituents such as noun or verb phrases (NP or VP) (Collobert et al., 2011). IOBES (inside, other,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

beginning, end, and single) encoding scheme has been adopted to mark chunk boundaries, i.e., identify location of the current word in the chunk. It associates a single tag with each input token. To mark a verb phrase containing a single word, it uses tag "*S-VP*". Tags "*B-VP*", "*I-VP*", and "*E-VP*" mark the first, intermediate and last words of the verb phrase. Tag "*O*" marks those words that are not members of a chunk. Before transforming a sentence into its CG, anaphora resolution is performed.

A sentence is transformed into its CG representation through following steps:

- Build CG representation for every syntactic constituent of a sentence.
- Link together all the conceptual graphs representing the constituents to form a single graph, thus, representing the complete sentence. (Hensman, 2004; Hensman et al., 2004)

First step is to build a conceptual graph for every constituent or phrase. Here, assumption is to represent each token in a sentence by a separate concept, therefore all nouns, verbs, adjectives and adverbs are represented using individual concepts. Determiners, numbers etc. are used as referents of the relevant concept. However, this work does not consider referents while creating CGs (Hensman, 2004; Hensman et al.., 2004). Not using referents for CG also keeps the system domain independent as the system need not maintain any ontology or knowledge base. For a sentence like "Tommy likes milk" the CG is Tommy:Dog <- like -> milk. So, the system needs to maintain an ontology where it knows that Tommy is a dog.

Algorithm 1 depict steps to be followed for creation of CG of a sentence already processed through SENNA. If CHK-tag = '*S*', a concept with the corresponding token as concept type is created. If CHK-tag = '*B*', it marks the beginning of a new chunk, hence, look forward in the following lines till corresponding '*E*' tag is found which marks end of the chunk or phrase. While traversing through a chunk, track of POS labels of each token is kept and concatenated with each other as they make the phrase structure of the identified chunk. Finally, concepts based on the involved syntactic pattern are created. Some of the syntactic patterns for noun phrases are shown in Table 1.

Algorithm 1. Creation of Conceptual Graph

```
Input: Text file containing syntactic information of a sentence in a tabular
format
Output: Conceptual Graph for a sentence
1.        begin
2.        while file is not empty
3.                 begin while
4.                         if CHK-tag ='S' and CHK = "PP"
5.                                Do nothing
6.                         else if CHK-tag = 'S'
7.                                Create a concept for the corresponding to-
ken
8                         else if CHK-tag = 'B'
9.                                 Traverse the file till corresponding 'E'
tag is found and based on the phrase structure formed by corresponding POS
tags create a CG
```


EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

```
10.         end while
11.         for each of the CGs created
12.                 Locate the semantic role of the head node and link the head
node to the verb concept using the identified semantic role
13. end
```

Each of the syntactic patterns needs different processing. For example, for pattern (3), concepts that represent noun token and adjective token are created and then linked using an attribute (*attr*) relation. Determiners in a sentence provide information on the type of referents. However, referents are not considered while creating CGs in this work. Following this notation, for pattern (4), a single concept for noun without any referent is created.

In case of prepositions, they mark the beginning of semantic role played by rest of the relevant phrases in the sentence for the identified verb. Therefore, if for a token, CHK-tag = '*S*' and CHK = "*PP*", no concept is created for it.

After separate graphs for each constituent is constructed, they are linked together to get a complete conceptual graph for the sentence. As each of these constituents describes one or another aspect of the verb, therefore their CGs are linked to the concept representing the verb. The concept representing head of the constituent is linked to the verb concept with an appropriate relation.

Compute Similarity of Conceptual Graphs

To evaluate candidate's answer, firstly the proposed method compares CGs of candidate's answer and model answer and then based on the closeness amongst graphs, marks are calculated.

The researchers have used the method proposed by Montes et al. (2000) to compute similarity of CGs. The algorithm for comparison of CG representations of two sentences is implemented in two main steps:

1. Find intersections (I_1, I_2 …) of the given two conceptual graphs to identify the common elements.
2. Measure similarity between the two graphs as the relative size of each one of their intersection graphs.

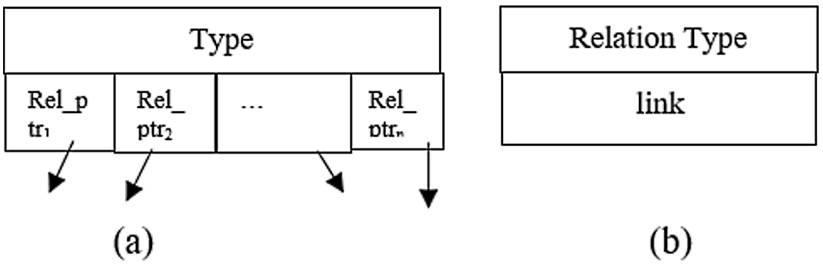
The value of similarity score varies between 0 to 1, where 0 indicates that given two graphs are totally dissimilar and 1 indicates that the two CGs are semantically equivalent.

Table 1. Some of the Syntactic Patterns Headed by Noun Phrase (NP)

[1] [9] NP ->NP PP (1)
[2] NP -> NNP (NNP ...)(or NNPS) (2)
[3] NP -> JJ NN (3)
[4] NP -> DT NN (4)
[5] NP -> NN (5)
[6] NP -> NNS (6)
[7] NP -> DT JJ NN (7)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Structure of (a) CONCEPT and (b) RELATION Nodes



Intersection of Two Conceptual Graphs

The intersection graph contains all concept nodes common in both the graphs and relation nodes appearing in both the graphs that relate the same concept nodes. In general, intersection of graphs may consist of more than one subgraph. In such a case, similarity measure is applied to each of them separately, and the highest value is taken as the similarity value.

To implement intersection of CGs, create two kinds of nodes – CONCEPT and RELATION nodes. CONCEPT node consists of a *type* field which stores the concept type of the concept and an array of pointers *(rel_ptr$_0$, rel_ptr$_1$ … rel_ptr$_n$)* to RELATION nodes that link concept to relations. In this work, authors are not taking referent field into consideration. Further, RELATION node consists of a *relation type* field and a *link* pointer which connects relation node to a concept node, assuming only binary relations (see Figure 3).

Algorithm 2. Intersection of Conceptual Graphs

Input: Student's answer CG and model answer CG.
Output: Intersection subgraphs of two inputted graphs.
Algorithm starts with the root node of student's CG

Step 1: Compare a concept from student's graph with all the concepts in model graph. If a match is found, continue with step 2 else goto step 4.
Step 2: Create a new concept. Compare the corresponding preceding concept as well as the relation joining the two concepts of student graph as well as already initiated intersection subgraph. If they are same, then link the newly created concept to the existing intersection subgraph.
Step 3: Else initiate a new intersection subgraph with the newly created concept.
Step 4: Repeat steps 1 to 4 with the next concept in student's graph till no more concept in student graph is left for comparison.

Similarity Measure of Conceptual Graphs

In the second step, correspondence between graphs is measured. Since conceptual graphs are bipartite in nature, the similarity *s* is calculated as a combination of conceptual similarity, and relational similarity (Montes et al., 2000).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. The conceptual similarity, s_c, measures degree of similarity between concepts in both sentences.
2. The relational similarity, s_r, measures similarity of information communicated by these concepts in the two sentences.

The conceptual similarity s_c provides the number of concepts that are common in both the graphs. This is calculated by using an expression analogous to Dice coefficient (Dice, 1945):

$$s_c = \frac{2n\left(G_i\right)}{n\left(G_s\right)+n\left(G_m\right)} \tag{1}$$

where $n(G_s)$, $n(G_m)$ and $n(G_i)$ are the number of concept nodes contained in student's, model and intersected conceptual graphs respectively. The value of this expression lies between 0 and 1. Value is 0 when the given graphs have no concept in common between them and is equal to 1 when the two graphs have identical set of concepts. Method to count the number of concepts in a graph is given in Algorithm 3.

Algorithm 3. Count Number of Concept Nodes in a CG

Input: Nodei - initial concept node of a CG. No_of_concepts - initially 0.
Output: No_of_concepts - number of concepts in a graph

```
CountConcepts(CONCEPT *Node_c)
begin
        if  Node_c is not NULL
                No_of_concepts++
                Recursively call CountConcepts for all concepts related via
relation pointers rel_ptr in Node_c
end
```

To identify relationship between contexts of common concepts, relational similarity s_r is calculated. This measure provides the similarity of information communicated in both sentences about these concepts and is calculated using Equation 2.

$$s_r = \frac{2 * m\left(G_i\right)}{m_{G_i}\left(G_s\right)+m_{G_i}\left(G_m\right)} \tag{2}$$

where $m(G_i)$ is the number of relation nodes in conceptual graph G_i, and $m_{Gi}(G_s)$, $m_{Gi}(G_m)$ is the number of edges in G_s and G_m which have at least one end belonging to G_i respectively. To implement this measure, mathematical method to calculate the value of $m_Y(G)$ for a subgraph $Y \subseteq G$ is used as Equation 3:

$$m_Y\left(G\right) = \sum_{c \in Y} deg_G c - m\left(Y\right) \tag{3}$$

where $\deg_G c$ is the degree of a concept node c which is equal to the number of edges connected to c in the conceptual graph G.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Algorithm 4 and 5 describe methods to find the number of relation nodes in a graph and the degree of a concept node in a CG respectively.

Algorithm 4. Count Number of Relation Nodes in a Conceptual Graph

Input: $Node_j$ – initial concept node of a CG. No_of_relations – initially 0
Output: No_of_relations - number of relations in a graph

```
CountRelations(CONCEPT *Node_r)
begin
        if Node_j is not NULL
                for all non-null relation pointers rel_ptr in Node_r
                                No_of_relations ++
                                Recursively call CountRelations for all con-
cepts related via relation pointers rel_ptr in Node_r
End
```

Algorithm 5. Find Degree of a Concept Node in a Conceptual Graph

This algorithm finds degree of a concept which is sum of incoming and outgoing edges in a concept.

```
Input:          CONCEPT Node_d - node whose degree is to be calculated. concept_
degree -  initially 0
Output: concept_degree- degree of a concept node in a graph
CountDegree(CONCEPT *Node_d)
begin
        if Node_d has an incoming edge
                        Concept_degree++
                for all non-null relation pointers rel_ptr in Node_d
                        Concept_degree++
end
```

The cumulative similarity measure *s* is defined as

$$s = s_c * (a + b * s_r) \tag{4}$$

Coefficients *a* and *b* are smoothing coefficients to level the effect of s_r on cumulative measure, *s*. The values of coefficients *a* and *b* depends on the structure of the graphs G_s and G_m. Value of *a* is calculated using the following formula:

$$a = \frac{2n\left(G_i\right)}{2n\left(G_i\right) + m_{G_i}\left(G_s\right) + m_{G_i}\left(G_m\right)} \tag{5}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

where $n(G_i)$ is the number of concept nodes in intersection graph *Gi* and $m_{G_i}(G_s) + m_{G_i}(G_m)$ is the number of relation nodes in G_s and G_m that are connected to the concept nodes appearing in G_i. Value of b is given as below:

$$b = 1 - a \tag{6}$$

Once the intersection of graphs is computed, then using Equations 1 to 6, similarity score between CGs is calculated.

Calculation of Marks

After calculation of similarity values for all the intersection subgraphs, the highest value amongst them is taken as the similarity score of the original graphs. Marks are assigned based on this similarity score using the equation given2 below:

Marks assigned = Total marks allocated * similarity score (7)

ILLUSTRATIVE EXAMPLE

The above method can be explained through examples where model answer and candidate's answers are taken as input. For an example:

Question: Why are leg bones so large compared to other bones in the human body?
Model answer: The leg bones need to be large because the leg bones carry a lot of weight.
Candidate Answer 1 (C1): Because leg bones have to hold up a lot of weight so they have to be big.
Candidate Answer 2 (C2): They support your body.
Candidate Answer 3 (C3): The leg bones help stand up. They hold the most weight.

To evaluate candidate's answer, first step is to run these sentences through SENNA – SRL tagger. Conceptual Graphs for all these sentences is generated using Algorithm 1. The generated CGs are shown in Figure 4, Figure 5a, Figure 6a and Figure 7a for model and candidates' answers respectively. To calculate the resemblance between these graphs, intersection of graphs is computed as per Algorithm 2 and the intersected graph is shown in Figure 5b, Figure 6b and Figure 7b respectively.

Similarity values calculated using Equations 1 to 6 are given in Table 2. After similarity is measured, marks are calculated using Equation 7 assuming total marks assigned to the question be 10.

EXPERIMENTS AND EXPERIMENTAL RESULTS

To evaluate performance of the proposed approach, number of experiments are conducted. In this section, (i) evaluation metrics, (ii) experimental setup and procedure, and (iii) experimental results, are described.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Conceptual Graph for Model Answer

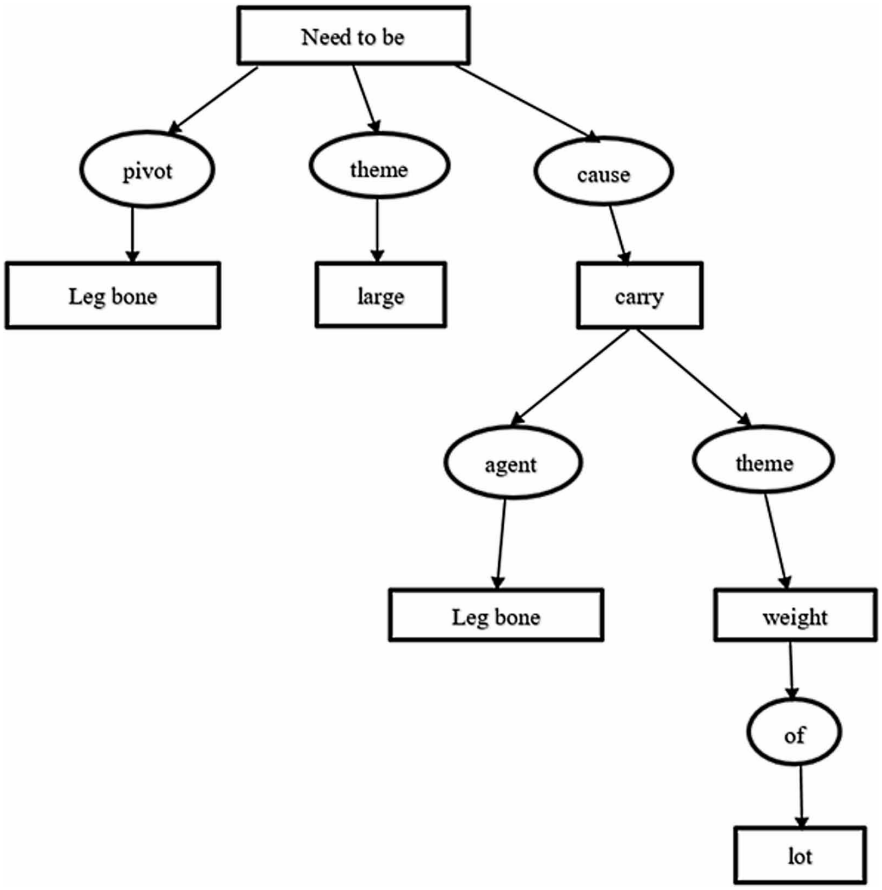
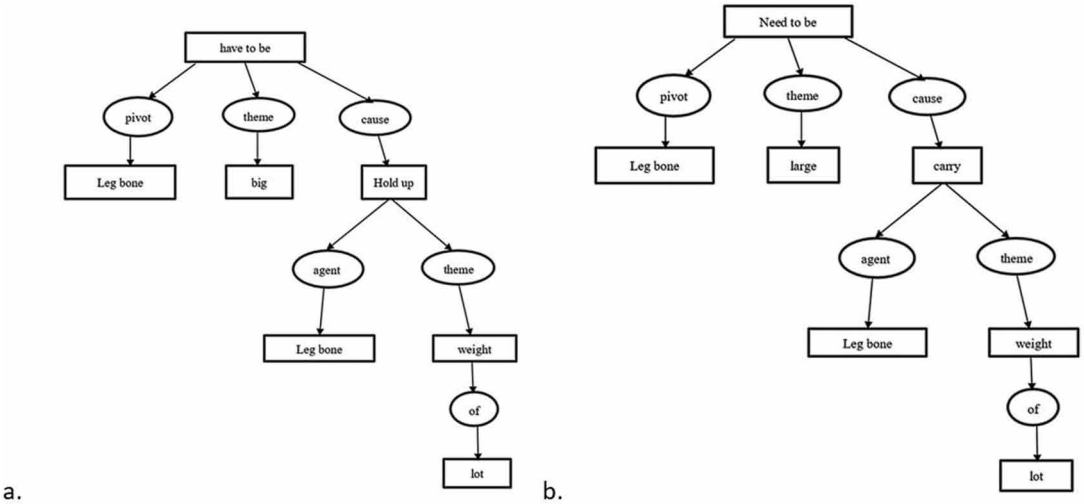


Figure 5. a) CG for Candidate Answer, C1 b) CG of Intersection of Model Answer and C1



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. a) CG for Candidate Answer, C2 b) CG of Intersection of Model Answer and C2

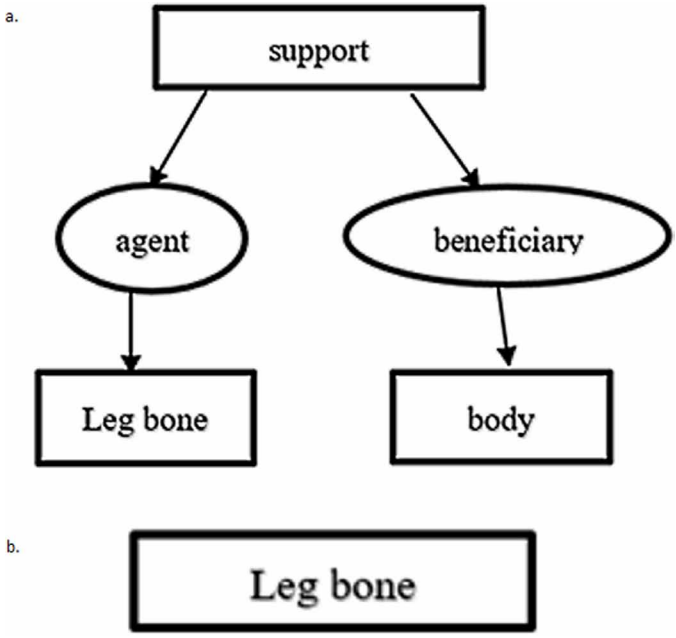


Table 2. Similarity Values Between the Graphs

Candidate Answer	s_c	s_r	A	b	s	Obtained Marks
C1	1.00	1.00	.67	.33	1	10 (correct)
C2	.4	.67	.57	.43	.34	3.4 (incorrect)
C3	.57	.85	.53	.47	.53	5.3 (correct)

Evaluation Metrics

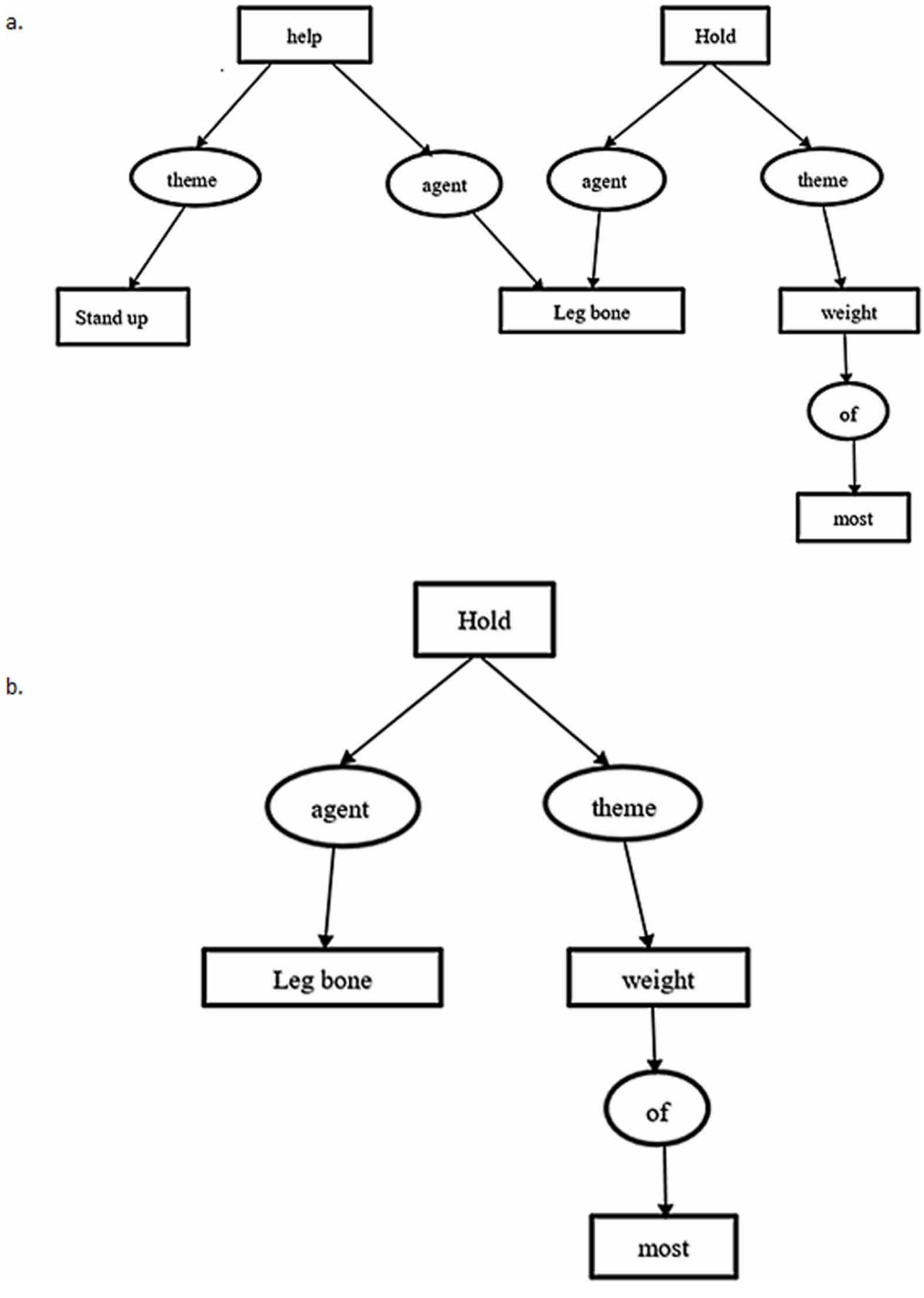
To evaluate performance of this system, authors compare the degree of agreement between human assessment and assessment of this proposed system using inter-rater agreement. Pearson Correlation Coefficient[3] has been used to evaluate the performance of proposed system.

"Pearson Correlation Coefficient is a statistical measure of the strength of linear relationship between two variables". It measures the standard correlation, that is, how much the human assessor scores or true scores (X) are related with the scores assigned by the system (Y). It is calculated by using Equation 8 below. The value of this measure lies between +1 and -1 inclusive where +1 signifies total positive correlation, 0 means no correlation and -1 shows total negative correlation.

$$Correlation\left(X,Y\right)=\frac{Covariance\left(X,Y\right)}{Standard\,Deviation\left(X\right)*Standard\,Deviation\left(Y\right)} \tag{8}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. a) CG for Candidate Answer, C3 b) CG of Intersection of Model Answer and C3



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Experimental Setup and Procedure

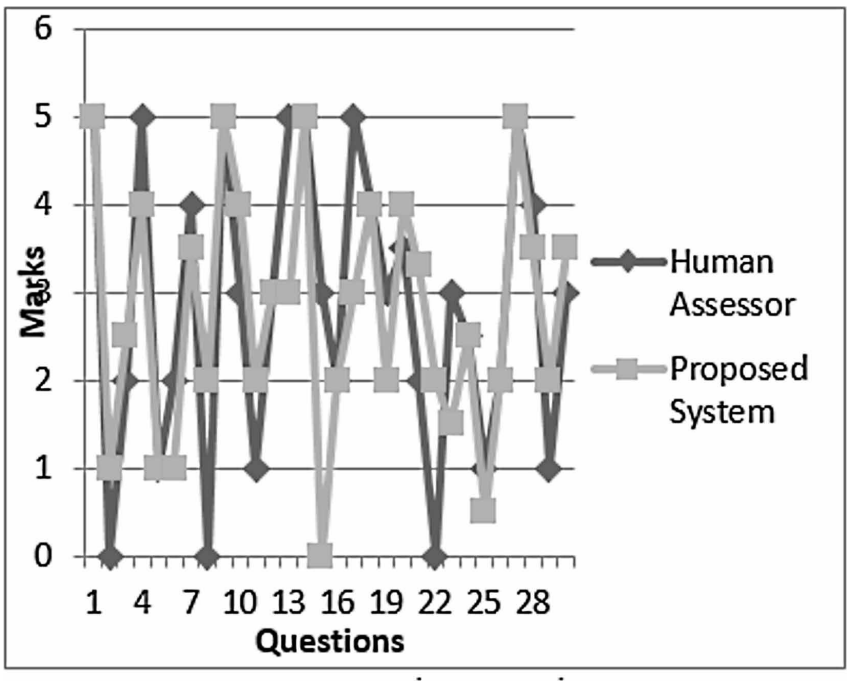
The proposed algorithm has been implemented in C language. Empirical approach is considered for evaluating the system where users compose input for the developed system.

Initially, all experiments have been carried out with students belonging to same standard in primary school. A questionnaire was prepared consisting of 30 questions belonging to different topics. All questions carried equal weightage of 5 marks each. This questionnaire was distributed among 24 students and they were asked to answer these questions in a single sentence. Model answers for these questions were generated with the help of subject experts. The students' responses were then assessed by the newly developed assessment system as well as by human assessors. Marks were obtained from both assessments. Analysis of obtained marks showed that proposed system fully agreed with human graders about 44.7% of the time.

Experimental Results

The evaluations are run using Pearson's correlation coefficient filling one vector with marks assigned by human assessor and other with the marks generated by proposed assessment system. The correlation between the marks from human assessor and proposed system is measured and its value is $r = 0.73$ which suggests a strong positive correlation, i.e. higher score by human assessor is associated with higher score by the proposed system. A graphical description of this correlation is presented in Figure 8.

Figure 8. Plot Between Human Assessor Scores and Proposed Assessment System Score Using 30 Test Samples



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Overview of Reviewed Automatic Assessment Systems

System	Techniques	Evaluation	Language
Proposed System	NLP	Corr: 0.73	English
C-Rater	NLP	Agr: 0.83	English
E-Rater	NLP and VSM	Agr: 0.87	English
Automark	IE	Corr: 0.93	English
ES4ES	IE	Corr: 0.71	English

The overlapping points on the graph indicate the number of times the marks assigned by a teacher were exactly the same as the marks assigned by the system. Next, authors compare the performance of the proposed system with the existing 4 different automated assessment systems also. Table 3 gives overview of the techniques, evaluations results and language used of the reviewed automatic assessment systems.

SemEval-2013 – Task 7: The Joint Student Response Analysis and 8th Recognizing Textual Entailment Challenge

Effectiveness of proposed system is checked by performing student response analysis on SemEval-2013 Task 7 test data set. The data for competition was taken from Student Response Analysis corpus (Dzikovska et al... 2012) This corpus consists of two subsets: Beetle data (Dzikovska et al... 2010) and SciEntsBank data (Nielsen et al... 2008). With the aim to support those systems that generalize across problems and domains, three distinct test sets viz. Unseen Answers (UA), Unseen Questions (UQ) and Unseen Domain (UD) were created (Dzikovska et al... 2013).

Researchers performed experiments on 2-way test set i.e. proposed system marks answer as either correct or incorrect which is an umbrella term for partially_correct_incomlplete, contradictory, irrelevant and non-domain. Experiments undertaken established that answers with grades between 0 to 0.4 are incorrect and between 0.4 to 1 are marked correct. Experiments were performed on all three categories of test sets (UA, UQ and UD) on SciEntsBank data set and results obtained are shown in Figure 9. Results are tabulated in percentage of answers that were correctly marked as correct and incorrect in all three test sets. The results in Figure 9a demonstrates that system works fairly well with all the three test sets.

Comparison of proposed system with original SemEval-2013 competition results is shown in Figure 9b. The evaluation metric used was macro-average that is the average value of each metric i.e. recall, precision and F1 across classes without taking class size into consideration. It is defined as $1/N_c \sum_c \text{metric}(c)$, where N_c is the number of classes which in our case is 2. It is evident that this algorithm performs well and is quite competitive with other approaches. Though proposed system is a good performer in UA and UQ category but is one of the top performer in UD category. The reason for good performance in UD is the approach followed where proposed method does not rely on any ontology or knowledge base for CG creation and thus is domain independent. This gives the system an edge over other presented systems. A list of some of the examples taken from SemEval-2013 test sets is tabulated in Table 4 along with the evaluation results obtained using the proposed system.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. a) Summary of Percentage of Answers Correctly Marked as "Correct" and "Incorrect" in Each Test Set Category. b) Comparison of Performance of Proposed System with Systems Participated in SemEval 2013 Task 7

Test Set	Correct (%)	Incorrect (%)
UA	62.5	59.21
UQ	63.19	70.32
UD	72.44	72.97

Run	UA	UQ	UD
$CELI_1$	.588	.619	.615
$CNGL_2$	.591	.561	.556
$CoMeT_1$	**.768**	.579	.670
CU_1	.603	.638	.673
ETS_1	.705	.688	.683
ETS_2	**.762**	.602	.543
$LIMSIILES_1$	.583	.629	.648
Proposed System	**.674**	**.639**	**.692**
$SoftCardinality_1$	.715	**.737**	**.705**
$UKP\text{-}BIU_1$	.726	.669	.666
Baseline Lexical	.617	.630	.650
Baseline Majority	.362	.367	.367

Table 4. List of Some Examples Marked Using Proposed System

Question	Reference/Model Answer	Student Answer	Answer Accuracy as per test set	Marks and Result (Correct/ Incorrect)
You used several methods to separate and identify the substances in mock rocks. How did you separate the salt from the water?	The water was evaporated, leaving the salt.	We evaporated the water.	Correct	.94 (correct)
		Get all the water out.	Incorrect	.31(Incorrect)
You can tell if a rock contains calcite by putting it into a cold acid (like vinegar). Describe what you would observe if you did the acid test on a rock that contains this substance	Many tiny bubbles will rise from the calcite when the calcite contacts the cold acid.	If you see bubbles coming from the rock it contains calcite.	Correct	.82 (correct)
		It would fizzle.	Correct	0 (Incorrect)
You used 3 scratch tools in class to test minerals for hardness: your fingernail, a penny, and a paperclip. If a mineral can be scratched by a penny, you can be sure that a (paperclip) will also scratch it. Explain how you know that tool would scratch the mineral.	A paperclip is harder than a penny, so if a penny can scratch a mineral, a paperclip can also scratch the mineral.	Because the paperclip is an easier item to scratch a rock or mineral than a penny.	Incorrect	.14 (Incorrect)
		I knew that because the paperclip is harder than the penny.	Correct	.86 (correct)
What happens to earth materials during erosion?	Earth materials are worn away and moved during erosion.	They all settle to the bottom.	Incorrect	0 (Incorrect)
		During erosion earth materials break away and flow away.	Correct	.66 (correct)
Why does the level of the water in the straw go up when the bottle-and-straw thermometer is heated?	Water expands when water is heated.	Because hot water expands and cold water contracts.	Correct	.61 (correct)
		The hot water heats the thermometer.	Incorrect	.42 (correct)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

Computer Based Assessment has been an interesting research area since 1970's. It helps in the evaluation of student's performance without bias and wastage of time. Automatic evaluation of descriptive answers would be beneficial for schools and universities as well as it provides ease to evaluators. The algorithm proposes a solution for automating the evaluation process of simple textual descriptive answers. The overall performance of this system was found to be 0.73. Major constraint of proposed system is that it is not considering answers that include mathematical formulas, diagrams, and examples.

Since the current work is a modest attempt to develop a CG based evaluation system, there is definitely a good scope for future research work in this area.

One of the possible works is to have an exhaustive analysis for similarity matching by assigning weightage to each term or word to have a better understanding of the relevance of each word in the answer. This may thus help in better evaluation of answers. The proposed algorithm can be modified for evaluation of complex, compound and multiple sentences answers.

ACKNOWLEDGMENT

This work was supported in part by the Council of Scientific and Industrial Research (CSIR) with file no. 09/263(1018)/2014-EMR-I.

REFERENCES

Ade-Ibijola, A. O., Wakama, I., & Amadi, J. C. (2012). An expert system for automated essay scoring (aes) in computing using shallow nlp techniques for inferencing. *International Journal of Computers and Applications*, *51*(10).

Chakraborty P. (2012). Developing an intelligent tutoring system for assessing students' cognition and evaluating descriptive type answers. *International Journal of Modern Engineering Research*, *2*(3), 985-990.

Collobert, R. (2011). Deep Learning for Efficient Discriminative Parsing. In *Proceedings Of International Conference on Artificial Intelligence and Statistics (AISTATS)* (pp. 224- 232).

Collobert R., Weston J., Bottou L., Karlen M., Kavukcuoglu K. and Kuksa P. Natural Language Processing (Almost) from Scratch. *Journal of Machine Learning Research*, *12*(1), 2493-2537.

Cutrone, L., & Chang, M. (2011). Auto-assessor: computerized assessment system for marking student's short-answers automatically. In *Proceedings of the IEEE International Conference on Technology for Education* (pp. 81-88). 10.1109/T4E.2011.21

Dhokrat, A., Gite, H. R., & Mahender, C. N. (2012). Assessment of Answers: Online Subjective Examination. In *Proceeding Workshop on Question Answering for Complex Domains COLING 2012*, Mumbai (pp. 47–56).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dice, L. R. (1945, July). Measures of the Amount of Ecologic Association between Species. *Ecology*, *26*(3), 297–302. doi:10.2307/1932409

Dzikovska, M. O., Moore, J. D., Steinhauser, N., Campbell, G., Farrow, E., & Callaway, C. B. Beetle II: a system for tutoring and computational linguistics experimentation. In *Proceedings of ACL 2010 System Demonstrations*, 2010, pp. 13-18.

Dzikovska, M. O., Nielsen, R. D., & Brew, C. (2012). Towards effective tutorial feedback for explanation questions: A dataset and baselines. In *Proceedings of 2012 Conference of NAACL: Human Language Technologies* (pp. 200–210).

Dzikovska, M. O., Nielsen, R. D., Brew, C., Leacock, C., Giampiccolo, D., Bentivogli, L., . . . Dagan, I. (2013). SemEval-2013 Task7: The Joint Student Response Analysis and 8th Recognizing Textual Entailment Challenge. In *Proceedings of the Second Joint Conference on Lexical and Computational Semantics (*SEM)* (pp. 263–274).

Foltz P. W., Laham D., Landauer T. K. (1999, April). The Intelligent Essay Assessor: Applications to Educational Technology. *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning*.

Hensman, S. (2004). Construction of Conceptual Graph representation of text. In *Proc. HLT-SRWS '04 Student Research Workshop at HLT-NAACL* (pp. 49-54). 10.3115/1614038.1614047

Hensman, S., & Dunnion, J. (2004). Constructing Conceptual Graphs using Linguistic Resources. *In Proceedings 7th International Conference, TSD 2004*, Brno, Czech Republic (pp. 81-88).

Ishioka, T., & Kameda, M. (2004). Automated Japanese essay scoring system: jess. In *Proceedings 15th International Workshop on Database and Expert Systems Applications* (pp. 4-8). 10.1109/DEXA.2004.1333440

Jayshankar, S., & Sridaran, R. (2016). Superlative model using word cloud for short answers evaluation in eLearning. *Education and Information Technologies*, *21*(3), 1–20.

John, F. (1999). *Sowa, Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Brooks Cole Publishing Co.

Kaiche, B., Kalan, S., More, S., & Shelukar, L. (2014). Online descriptive examination and assessment system. *International Journal of Emerging Technology and Advanced Engineering*, *4*(3), 660-664.

Kaur A., Sasikumar M., Nema S, Pawar S. (2013). Algorithm for Automatic evaluation of single sentence descriptive answer. *International Journal of Inventive Engineering and Science*, *1*(9), 112-121.

Larkey, L. S. (1998). Automatic Essay Grading Using Text Categorization Techniques. In *Proceedings of 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Melbourne, Australia (pp. 90-95). 10.1145/290941.290965

Leacock C., Chodorow M. (2003). C-rater: automated scoring of short-answer questions. *Computer and the Humanities, 37*(4), 389-405.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Mitchell, T., Russell, T., Broomhead, P., & Aldridge, N. Towards robust computerized marking of free text responses. In *Proceedings of 6th computer Assisted Assessment Conference, 2002*.

Montes-y-Gómez, M., Gelbukh, A., López-López, A., & Baeza-Yates, R. Flexible Comparison of Conceptual Graphs. In *Proceedings of 12th International Conference, DEXA 2001* Munich, Germany, 2001, pp. 102-111.

Montes-y-Gómez, M., López-López, A., & Gelbukh, A. (2000). Information Retrieval using Conceptual Graph Matching. In *Proceedings of DEXA-2000, 11th conference on Database and Expert Systems Applications*, Greenwich, England (pp. 312-321).

Nielsen, R. D., Ward, W., Martin, J. H., & Palmer, M. (2008). Annotating students' understanding of science concepts. In *Proceedings of the Sixth International Language Resources and Evaluation Conference, (LREC08)*.

Palmer, M., Gildea, D., & Kingsbury, P. (2005). The Proposition Bank: A Corpus Annotated with Semantic Roles. *Computational Linguistics Journal*, *31*(1), 71-106.

Patil, S., & Patil S. (2014). Intelligent tutoring system for evaluating student performance in descriptive answers using natural language processing. *International Journal of Science and Research*, 3(9), 779-783.

Rotella, F., Ferilli, S., & Leuzzi, F. (2013). An Approach to Automated Learning of Conceptual Graphs from Text. In *Proceedings 26th Int. Conf. on IEA/AIE 2013*, Amsterdam, The Netherlands (pp. 341-350).

Sowa, J. F. (1984). *Conceptual Structures: Information Processing in Mind and Machine, The Systems Programming Series*. Addison-Wesley.

Sowa J. F., Way E. C. (1986). Implementing a semantic interpreter using conceptual graphs. *IBM Journal of Research and Development*, *30*(1), 57-69.

Sukkarieh, J. Z., Pulman, S. G., & Raikes, N. (2003). Auto marking: using computational linguistics to score short, free text responses. In *Proceedings 29th annual Conference of IAEA*.

Willis, A. (2015). Using NLP to support scalable assessment of short free text responses. In *Proceedings of Tenth Workshop on Innovative Use of NLP for Building Educational Applications* (pp. 243-253). 10.3115/v1/W15-0628

Yang, S., & Soo, V. (2012, June). Extract conceptual graphs from plain texts in patent claims. *Engineering Applications of Artificial Intelligence*, *25*(4), 874–887. doi:10.1016/j.engappai.2011.11.006

Zeng, H., Xiong, L., & Chen, J. (2014). The Chinese Conceptual Graph Matching Algorithm Based on Conceptual Sub-Graph Weight Self-adjustment. In *Proceedings of the 9th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing* (pp. 197-200).

ENDNOTES

[1] Mate Tools. Available: https://code.google.com/p/mate-tools/

[2] SEMAFOR. Available: http://www.ark.cs.cmu.edu/SEMAFOR/

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

[3] SENNA. Available: http://ml.nec-labs.com/senna/

[4] ClearNLP. Available: http://clearnlp.wikispaces.com/

[5] Pearson Correlation. Available: http://en.wikipedia.org/wiki/Pearson_product-moment_correlation_coefficient

This research was previously published in the International Journal of Conceptual Structures and Smart Applications (IJCSSA), 5(2); edited by Simon Polovina and Simon Andrews; pages 1-21; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 38
Wild Architecture:
Explaining Cognition via Self-Sustaining Systems

Vincent T. Cialdella
Illinois State University, USA

Emilio J. C. Lobato
Illinois State University, USA

J. Scott Jordan
Illinois State University, USA

ABSTRACT

In this chapter, the authors focus on cognitive architectures that are developed with the intent to explain human cognition. The authors first describe the mission of cybernetics and early cognitive architectures and recount the popular criticism that these perspectives fail to provide genuine explanations of cognition. Moving forward, the authors propose that there are three pervasive problems that modern cognitive architectures must address: the problem of consciousness, the problem of embodiment, and the problem of representation. Wild Systems Theory (Jordan, 2013) conceptualizes biological cognition as a feature of self-sustaining embodied context that manifests itself at multiple, nested, time-scales. In this manner, Wild Systems Theory is presented as a particularly useful framework for coherently addressing the problems of consciousness, embodiment, and representation.

INTRODUCTION

For some time now, one of the leading assumptions in the development of cognitive architectures has been Marr's (1982) tri-level theory of explanation; the idea that the proper approach in developing a cognitive architecture (and thus explaining cognition) is to

DOI: 10.4018/978-1-7998-0951-7.ch038

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. Determine the computations necessary to completing a cognitive task (e.g., sorting a list of numbers from lowest to highest),
2. Generate a representation of the inputs, outputs, and algorithms an information-processing system would need to complete the task, and
3. Actually build (i.e., implement) a system capable of executing the algorithms.

The purpose of the present paper is to examine issues that have proven challenging to Marr's implementation approach to explaining cognition. Three particular challenges are the issues of consciousness, embodiment, and representation. After examining these challenges, we will present an approach to describing cognitive architectures (Wild Systems Theory—WST, Jordan, 20213) that addresses each challenge, while simultaneously shifting the focus of modeling from looking to biology for inspiration, to looking at a more fundamental property that biological systems share with many other types of systems, including chemical, psychological, and cultural—specifically, the ability of certain far from equilibrium systems to generate catalysts that feedback into and sustain the processes that produced them; what Kauffman (1995) refers to as 'autocatalytic' systems, and what Jordan (2013) refers to as self-sustaining, or *wild* systems.

PROBLEMS WITH "IMPLENTATION AS EXPLANATION"

While Marr's (1982) approach to developing cognitive architectures has inspired research that has given rise to a host of new technologies, there are those who have expressed doubts regarding his assertion that implementation constitutes explanation. In his seminal paper, *Quantitative analysis of purposive systems: Some spadework at the foundation of scientific psychology*, William T. Powers (1973) expressed his belief that psychologists were working under the confused assumption that control-theoretic concepts had been developed to explain the behavior of organisms. According to Powers, control-theoretic concepts were developed so that engineers could develop systems capable of doing what organisms do; namely, maintain ordered states with their environment by offsetting environmental disturbances to those ordered states. Powers referred to this ability to maintain ordered states as *input control*, and he further stated that in order for engineers to be able to build input control systems, they had to develop a conceptual scheme that captured the dynamics of organismic input control in a way that allowed the dynamics to be transformed into functioning, artificial, input-control systems, what have come to be known as servomechanisms. Thus, the conceptual scheme created by engineers includes phrases such as reference signal, which represents the state the system is to sustain. For example, when one dials in a room temperature on a thermostat, one is basically specifying the reference signal for the system. Or, said another way, one is setting the input level the system should keep constant. In order to keep its input at a pre-specified level, the system must be able to counteract environmental events that move the input (i.e., sensed temperature) away from the value of the reference signal. For example, when one opens a window and allows cold air into the room, the air temperature will change. If it varies from the pre-specified level, the system must be able to offset such disturbance. The system's ability to generate disturbance-offsetting events is referred to as output. In the case of the thermostat, the system's output is the turning on of the furnace. As the system continues to generate output (e.g., the furnace stays on) the difference between the input and the reference signal decreases. From the perspective of an engineer, who is actually trying to build such a system, this means the system has to be able to compare its current

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

input state (i.e., sensed room temperature) to its reference signal. Engineers refer to the system component that accomplishes this function as a comparator. They also refer to the result of the comparison as an error signal. Collectively then, servomechanisms are constructed such that the error signal drives the system's disturbance-offsetting outputs (i.e., the difference between the sensed temperature and the reference signal determines when and for how long the furnace is active). As a result of this architecture, the error signal actually serves to negate itself (i.e., make itself become zero), and this explains why engineers also refer to servomechanisms as negative-feedback control systems.

The reason for analyzing the nature of servomechanistic concepts so thoroughly was to underline Powers' (1973) point that concepts such as *reference signal*, *disturbance*, *input*, *output*, *comparator*, *error-signal*, and *negative-feedback* were created by engineers so they could create systems capable of doing what organisms do (i.e., maintain ordered relationships). That is, the concepts were created for the purpose of *implementation*, not *explanation*. Implicit in Powers' argument is a take on the relationship between implementation and explanation that is counter to Marrs' (1982) position.

From Cybernetics to Cognitive Architecture

Following the cybernetic movement referred to by Powers (1973), cognitive psychologists began to work toward developing computational systems (i.e., architectures) meant to model human cognition. While some architectures were and are developed with the sole intention of enabling artificially intelligent behavior, others were and are developed in the hopes of explaining the architecture that underlies human cognition. Thus, as was the case in Powers' (1973) critique of the cybernetics movement, there are those whose approach to developing cognitive architectures entails two related, but distinct, goals: the technological goal and the explanatory goal. The former refers to the practicality of using artificial systems to complete tasks that require human level competence or, perhaps, a level of performance that goes beyond the limitations of humans (Duch, Oentaryo, & Pasquier, 2008). Cognitive scientists operating within the field of Artificial Intelligence (AI) are particularly interested in building systems that can realize the same cognitive functions that are achieved by humans. To some extent, the focus on modeling and building cognitive systems renders this approach an engineering perspective on cognition. On the other hand, the explanatory goal refers to the possibility of using what was learned while building an artificial cognitive system to explain the achievements of natural cognitive systems such as humans. That is, success in building an artificially cognitive system might be taken as an indication that human cognition is actualized by the same mechanisms that were used to produce artificial cognition (Anderson et al., 2004).

According to contemporary AI researchers, *cognitive architecture* can be described as the invariant properties that underlie cognition (Douglass & Mittal, 2013). While contexts inevitably vary, the principles by which a system engages the world exhibit consistency; that is, they are invariant. This consistency may be attributed to the system's stable architecture. Examples of cognitive architectures include Soar (Laird, 2008; Laird, Newell, & Rosenbloom, 1987), LIDA (Baars & Franklin, 2009; Franklin et al., 2013), and ACT-R (Anderson, 1996; Anderson et al., 2004). While these architectures may, to some extent, exhibit similarities, they diverge in their implementation, the authors' motivations and goals for their architectures, and the phenomena to which they are commonly applied. Laird (2012) described this divergence in reasons for development: Some developers design their architectures specifically with a more technical, engineering focus in order to support the development of agents, whereas others are designed with the goal of explaining cognition. In addition to their variable purpose, architectures also

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

differ in their fundamental assumptions about cognition; they may differ in the mechanisms by which they learn, the way in which knowledge is represented, or even the types of knowledge that can be represented.

To be sure, it is often the case that the same architecture is used to pursue both the technical and the explanatory goal. In the past, Soar has been used to model human behavior (Lewis, 1993; Miller & Laird, 1996), but the developers are now predominantly concerned with demonstrating its real-time functionality in artificial systems (Laird, 2012; Langley, Laird, & Rogers, 2009). In this sense, the developers are largely interested in achieving the technological goal of AI. In contrast with Soar, a majority of work in ACT-R seems to reflect an interest in the explanatory goal of AI. Although it has seen applied use in intelligent tutors (Koedinger, Anderson, Hadley, & Mark, 1997) and robotics (Trafton et al., 2013), ACT-R has largely been used to model a wide variety of data obtained from work in experimental psychology (e.g., Halbrügge, Quade, Engelbrecht, 2015; Wirzberger & Russewinkel, 2015). Despite the use of some architectures to achieve both technical and explanatory goals, we now focus on one particular architecture (i.e., ACT-R) and the issues that have arisen as research have attempted to use it as an explanation of cognition.

ACT-R as Explanation

The ACT-R architecture (Anderson et al., 2004) is composed of a number of components called modules, which each handle different types of information. Examples of ACT-R modules include an intentional module that maintains goal states and intentions, a declarative module that handles the retrieval of information from memory, a visual module that supports information about entities in the visual field, and a manual module that affords control over hands. Each module contains a buffer that maintains information relevant to the task currently being addressed. Two kinds of knowledge are represented in ACT-R: declarative knowledge and procedural knowledge. Declarative knowledge is represented using *chunks*: knowledge units that accommodate key-value pairs. Procedural knowledge is represented using *productions*: condition-action rules.

During run-time, buffers of each module can be manipulated by way of a central production system that recognizes when the chunk in a buffer meets certain conditions. When a pattern is matched, the production system can make changes to that buffer or the buffers of other modules. Buffers are updated by the productions that 'fire' when their conditions are met. For example, if a chunk in the visual buffer matches conditions specified in a production, the model might alter the chunk in the manual buffer to make a request to move the hands toward an identified object. In ACT-R, only one production can fire at a time. To resolve this conflict, a utility calculation determines which production will be most likely to achieve the current goal and consequently fire. Apart from the inter-module communication that is afforded through the central production system, these modules operate independently of one another.

According to ACT-R's implementers, the ACT-R architecture affords the development of software agents as well as a unified theory of human cognition (Anderson et al., 2004). That is, the authors strive toward the explanatory goal of cognitive architectures. As any good theory of human cognition should, the theory is used to explain a large body of data obtained in experimental psychology. In a comprehensive presentation of the ACT-R theory and related empirical work, Anderson and colleagues (2004) presented data from a number of behavioral and neuroimaging studies to identify the neural correlates of various ACT-R modules by tracking how and when information is updated in buffers while engaging in tasks. Since then, ACT-R models have been used in concert with fMRI data to localize the updating of working memory and the retrieval of declarative memory by using the activity of ACT-R's modules

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

as a predictor of neural activity in a regression analysis (Borst & Anderson, 2013). Another study has simulated the process of learning a new song, demonstrating robust primacy and recency effects that are similar to those observed in human subjects (Chikhaoui et al., 2009). Results from modeling efforts have also motivated the development of richer theoretical accounts of associative memory (Thomson, Pyke, Trafton, & Hiatt, 2015).

In other experiments, researchers have measured participants' ability to estimate the amount of time that has passed while multitasking. Comparing these results with ACT-R models has highlighted the importance of considering the influences of multiple contextual factors (Moon & Anderson, 2013). Researchers investigating action control while driving an automobile have evaluated the performance of novice and experienced drivers by using ACT-R to model the influence of driving experience on collision avoidance and lateral control skills (Cao, Qin, Jin, Zhao, & Shen, 2014; Cao, Qin, Zhao, & Shen, 2015). Additionally, comparison of models with human reaction-time data in a task-switching paradigm has generated possible, concrete accounts of psychological constructs such as flow (Altmann, & Trafton, 2007). In sum, by comparing the data exhibited by ACT-R models to the data collected from human participants, researchers are not only afforded the ability to generate possible explanations of observed phenomena but are also able to enhance their theoretical perspective.

Despite ACT-R's ability to computationally model experimental human data, there are those who express concerns regarding its ability to constitute an explanation of human cognition. For example, it is recognized (e.g., Bösser, 1987, 2013) that there exists more than one set of values for chunking parameters in Soar that can fit a single data set. This also applies to the parameterization in ACT-R. Others have criticized the practice of parameter estimation on the grounds that any pattern of data can be fit (Roberts & Pashler, 2000), and these criticisms have been appropriately responded to (see Anderson et al., 2004). Johnson-Laird expresses further suspicions regarding the explanatory capabilities of models:

The reason for the suspicion is complex. In part it derives from the fact that any large-scale program intended to model cognition inevitably incorporates components that lack psychological plausibility.… Certain aspects of any such program must be at best principled and deliberate simplifications or at worst ad hoc patches intended merely to enable the program to work. The remedy… is not to abandon computer programs, but to make a clear distinction between a program and the theory that it is intended to model. (1980, p. 110, emphasis in original)

Choices about what should be modeled (e.g., the types of knowledge) and how to model it will necessarily reflect the constraints and affordances of the technology that are made available to the developers. In order to obtain a functioning model, decisions might be made during implementation so as to accommodate for technological limitations. For this reason, as Johnson-Laird advises, one should be wary of equating a description (i.e., the model) with the explanation (i.e., the theory).

To be sure, the authors of ACT-R do not present their architecture as *the* truth about human cognition. Anderson and colleagues note, "just because such architectures aspire to account for the integration of cognition, they should not be interpreted as claiming to have the complete truth about the nature of the mind" (2004, p. 1057). Instead, the hallmark of cognitive architectures is their ability to force researchers to think about how all these various, independently studied functions can all be brought together under one cohesive framework. Focusing on the architecture of cognition draws attention to the problems that are neglected when theorists fail to concern themselves with the implementation of cognition.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONTEMPORARY ISSUES WITH IMPLEMENTATION

Contemporary attempts to explain human cognition via implemented architectures have begun to place more emphasis on the biological substrate of human cognition (Franklin, Strain, McCall, & Baars, 2013). The BICA (Biologically Inspired Cognitive Architectures) movement, as represented by the journal of the same name, describes its mission as follows: "The focus of the journal is on the integration of many research efforts in addressing the challenge of creating a real-life computational equivalent of the human mind." Along the way, BICA researchers have also had to address the issue of whether the goal is to implement an architecture that can produce the same output as a biological organism, or is the goal to implement an architecture that serves to explain biological cognition? If the goal is the former, then issues with implementation are predominantly pragmatic. Specifically, the challenge of implementation largely, if not entirely, concerns finding the algorithms that will provide the right kind of output. If the goal, however, is explanatory, then all of the issues in the behavioral, cognitive, and social sciences necessarily become issues for instantiating a particular model of cognition. In the following sections, we outline three general topics that we feel need to be addressed by any explanatory cognitive architecture. These topics are consciousness, embodiment, and representation.

The Problem of Consciousness

A more recent cognitive architecture has been developed in response to the pursuit of biologically inspired artificial intelligence systems: the Learning Intelligent Distribution Agent (LIDA; Franklin et al., 2007). LIDA was developed to implement the Global Workspace Theory (GWT) of consciousness to integrate both conscious and unconscious aspects of biological, in particular human, cognition (Baars & Franklin, 2009). GWT proposes that cognition results from the culmination of many specialized and distributed processes within the brain. Conscious cognition is enabled and constrained by unconscious cognitive processes, in particular perceptual processes, which are recursively influenced by the focal content of conscious cognition feeding back to these unconscious processes (for review, see Baars, 1988). What LIDA attempts to do is to computationally model and implement GWT. Specifically, the LIDA model focuses on the notion of cognition as resulting from cognitive cycles, which are rapid sensory samplings by the cognitive agent of the environment (both external and internal) to update the agent's representation of the environment. During this rapid sampling process, the agent's unconscious processing determines what aspects of the environment are sufficiently salient enough to draw conscious attention toward, which is then fed back to the lower-level processes so that subsequent appropriate action selection and execution can occur.

The developers of the LIDA model have also recently proposed LIDA as a referent architecture for the broader community of scholars interested in artificial intelligence by outlining the conceptual commitments of the model (Franklin et al., 2013). The major draw of doing so is that LIDA, which is an outgrowth of GWT, is committed to the importance of consciousness as a necessary component of cognition, in contrast to other architectures like ACT-R or Soar (Chella, 2013; Franklin et al., 2013).

The Problem of Embodiment

Curiously, the authors of the LIDA model, in explicating the conceptual commitments of the system, express a commitment to embodiment as significant for their model without necessarily insisting on physical, robotic implementation (Franklin et al., 2013). This is due to their broad interpretation of embodied cognition as

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

just the structural coupling of the agent with its environment. Their view is that general intelligence is a sufficiently abstract enough entity to not necessitate robotic instantiation, and embodiment still occurs as long as the agent and environment are coupled in some way (Franklin, 1997; Franklin et al., 2013). Commenting on this curiosity of a disembodied form of embodiment, Wiltshire and colleagues (2013) noted that physical embodiment of an AGI model such as LIDA would enhance the capacities of the model. Morphology prescribes and constrains an agent's ability to perturb and interact with its environment (Anderson, 2003), particularly with regards to more complex and dynamic forms of agent-environment interactions such as social cognition. Architectures that aim to be used for the development of artificially intelligent agents might be more successful if they were to directly address the complexity of embodied cognition.

To be sure, this is easier said than done. The complexity of the embodied cognition perspective is partially due to the lack of a firm or precise description of what constitutes a body (Anderson, 2003). The most radical claim of the embodied cognition perspective is that "the brain is not the sole cognitive resource we have available to use to solve problems" (Wilson & Golonka, 2013, p. 1). As dramatic as that claim is, which itself cannot be understated currently, the claim does require further elaboration on what constitutes a body. This issue is addressed only indirectly, by describing what bodies do, such as "move and act in rich real-world surroundings" (Clark, 1998, p. 506) or describing the role of the body in constraining the form that behavior takes via an organism's perceptual-motor capabilities (Anderson, 2003; Varela, Thompson, & Rosch, 1991; Wilson & Golonka, 2013).

The Problem of Representation

In addition to consciousness and embodiment, another on-going critique of the cognitive architecture literature is its total reliance on the notion of "representation." Common to most such criticisms is a commitment to some form or another of dynamical systems theory and the assertion that the work of cognition is to be found in the self-organizing, synergetic, circular causality of continuous interactions between brain, body and world (Clark, 1997; Jordan, 2003; Kelso, 1995; O'Regan & Noe, 2001; Tschacher & Dauwalder, 2003; Thelen & Smith, 1994; van Rooij, Bongers & Haselager, 2002; Varella, Thompson & Rosch, 1991; van Gelder, 1998). A classic example of a self-organizing system is a convection roll. If one applies heat to a pan of cool oil, a rather complex system of convection rolls will emerge as the heated oil rises to the top and the cooler oil sinks to the bottom.

The resulting convection rolls are what physicists call a collective or cooperative effect, which arises without any external instructions… Such spontaneous pattern formation is exactly what we mean by self-organization: the system organized itself, but there is no 'self', no agent inside the system doing the organizing. (Kelso, 1995, pp. 7-8)

Van Gelder (1998) strongly rejects the computationalist notion that behavior is controlled (i.e., regulated) by internal structures such as codes and plans. Instead, he proposes it is possible for a system to give rise to regulatory behavior solely on the basis of self-organizing dynamics. He presents the Watt Governor as an example.

A Watt Governor is a mechanism that keeps the velocity of a fan constant. It does so because the velocity of the fan's flywheel influences itself recursively. That is, the faster the flywheel moves, the more it slows itself down. This is because the movements of the flywheel cause two weighted arms, attached to the

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

flywheel via a spindle, to rise as the flywheel moves faster. These movements cause a valve, connected to a steam-engine driving the flywheel movement, to close, thereby reducing the amount of steam released which, in turn, slows the flywheel and allows the weighted arms to lower. Because of the constant, mutual influence all the components have on one another, fan velocity can be held fairly constant. Van Gelder proposes the Watt Governor as a prime example of regulated behavior emerging out of purely self-organizing processes. At no point in the Watt Governor systems is there a 'code', a 'representation', or information being processed. As a result, he argues it should be possible to model cognitive systems solely in terms of self-organizing dynamics. (Jordan & Ghin, 2007. P. 183-184)

Jaeger (1998) later criticized theorists who asserted that dynamical systems such as the Watt Governor constituted the proper way to explain cognition. Specifically, he believed that the current state of mathematics in dynamical systems theory was not sophisticated enough to effectively model organisms:

First, DST principles like attractors and bifurcations are not of much help in wild systems with fast stochastic input varying on the system's own characteristic time scale. Second…DST handles high-dimensional domains by reducing them to low-dimensional descriptions…this reduction to some collective parameters is helpful in some respects but still poses a limit to the study of high-dimensional systems. (1998, p. 643-644, emphasis added)

Collectively, critiques of the ability of currently available cognitive architectures to explain cognition imply that while modeling efforts such as those in the ACT-R and LIDA communities have demonstrated considerable progress toward the technological goal of AI, the extent to which the efforts display progress toward the explanatory goal remains unclear. That is, such efforts surely have the potential to yield artificial systems that can do what we do, but successful modeling does not necessarily demonstrate that these artificial systems do what we do *in the way that we do it*. At their core, cognitive architectures engage cognition as a purely computational process. This conception of biological cognition has existed for decades (e.g., Miller, Galanter, & Pribram, 1960), arising in part from a desire to broaden psychological science as more than strict behaviorism. However, this assumption of cognition as computation is not universally accepted (e.g., Gibson, 1979), and there exist useful and productive lines of research that incorporate non-computational explanatory mechanisms in their theorizing of human cognition, either as true alternatives to purely computational models of information processing (e.g., Marsh, Richardson, & Schmidt, 2009) or as additional components into pluralistic frameworks that propose both computational and non-computational mechanisms to explain aspects of human cognition (e.g., Wiltshire et al., 2014; see also Gallagher, 2015). It is our contention that progress in the development of cognitive architectures that are intended to explain human cognition, rather than "simply" mimic the output of human cognition, would benefit greatly from attempting to model such non-computational components. In particular, we propose that the Wild Systems Theory (WST) put forward by Jordan (2013) serves as an especially useful framework for such an endeavor. In the following section, we will present and describe WST as a way to conceptualize human cognition that may aid in the development of a biologically inspired cognitive architecture that supports intelligence by means similar to those which result in natural cognitive systems.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

WILD ARCHITECTURE

Jordan and Day (2015) describe organisms as multi-scale self-sustaining systems. This means that the work that organisms do is self-sustaining, or autocatalytic. That is, the work produces products that further promote the work. Put simply, an autocatalytic reaction between chemicals *A* and *B* will produce a catalyst for the *A*-*B* reaction (Kauffman, 1995). Such reactions continue to occur by the nature of their self-sustaining work. As networks of autocatalytic reactions are sustained, the stability maintained by this level of autocatalysis enables the emergence of larger-scale autocatalytic processes. When one level of autocatalytic order is achieved, another can emerge. One might conceptualize living systems as large-scale organizations of autocatalytic processes or multi-scale systems of self-sustaining work. The tree-climbing behavior of a monkey (i.e., the work) to obtain a banana will yield energy in the form of food that will further promote the energy-obtaining work of the monkey. Importantly, within each organism, self-sustaining work is occurring at multiple time scales simultaneously (Jordan, 2013). The monkey exhibits self-sustaining work at the behavioral level while the organs, tissues, and cells exhibit self-sustaining work at various different biochemical levels and time scales.

As these living systems emerge, they alter their environment by the nature of their autocatalytic work, providing a new context in which other self-sustaining systems can emerge. For example, the emergence of plants afforded the emergence of a life form that harvests plants for energy, and the emergence of herbivores afforded the emergence of a life form that harvests herbivores for energy. Given that organisms emerge out of the context in which they are embedded, the work they do necessarily reflects the kind of energy transformation that is required for self-sustainment in that context; the fuel source dictates the consumer (Jordan & Ghin, 2006). An organism (i.e., a self-sustaining energy transformer) sustains itself using the energy that is made available in its environment. Conceptualizing organisms in this way, nature is illustrated as a self-organizing energy-transformation hierarchy in which each organism alters its context such that it can continue to sustain itself and consequently creates a context in which another organism can obtain, transform, and dissipate the newly available energy. At each level, energy transformation is not what a system *does*; it is what a living system *is*.

Following from the notion that the fuel source dictates the consumer, Jordan and Day (2015) argue that all organisms constitute embodiments of their context. That is, organisms embody the dynamics in which they emerged and in which they continue to sustain themselves. As an example of an embodied context, the herbivore embodies the contextual constraints that determine what kind of work must be done for an organism to sustain itself in an environment furnished with plants. A carnivore embodies the constraints of its context, namely that it must be able to catch a moving energy source in order to sustain itself. For the carnivore, this requires anticipation of where the prey *will be* rather than where the prey *is* in order to be successful.

Jordan and Day (2015) contend that this anticipatory form of aboutness (i.e., embodiment) is what one might refer to as mind. However, it is important to note that there is a fundamental similarity between the abstract, anticipatory aboutness that we experience as a mental life and the aboutness of a plant. Both emerge out of the constraints that determine the work that must be done for an organism to sustain itself in a context. It just so happens that, because of the constraints of the context that fostered the emergence of humans, humans embodied the dynamics of temporally distal events. This allows humans to live rather anticipatory lives in which we can organize ourselves about abstract events such as ‘tomorrow’ (Jordan & Vinson, 2012).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

What Does WST Afford AI Research?

From the WST perspective, one is afforded a new way of describing what it is that intelligent behavior entails: multi-scale, self-sustaining energy transformation. In architectures like ACT-R, wherein the basic mechanisms work in concert to produce a rich combination of symbolic and sub-symbolic knowledge (through the acquisition and curation of knowledge units), Wild Systems Theory (WST) proposes that knowledge exists at every scale of work because every level is about the context in which that order is maintained. In this way, knowledge is not something that intelligent agents *have*; it is what they *are*. This leads to two related statements that could be taken as invariant principles for a biologically inspired, artificial cognitive system: (1) behavior embodies contextual constraints and affordances; a system's work reflects the kind of energy transformation that is necessary for sustainment in that context, and (2) the system consists of multiple levels (e.g., sub-personal; personal; supra-individual; see Bohl & van den Bos, 2012; Wiltshire et al., 2014) of self-sustaining work; each level of order is achieved and maintained by work that adheres to principle (1). We argue that these two statements coherently capture what it is that intelligent systems are: multi-scale, self-sustaining embodiments of context. Thus, adherence to these principles effectively integrates all of human behavior (and cognition).

In order to better clarify how we believe the notion of multi-scale, self-sustaining embodiment of context provides a biologically inspired approach to cognition and cognitive architecture, we need to more thoroughly describe how the architecture of self-sustaining systems qualifies as "cognitive." Take, for example, a rock. A rock can be considered an embodiment of its context. That is, animals, plants, and rocks could all be considered instantiations of the demands or forces that their environments exert on them. In this sense, the rock is much like animals and plants; the rock also constitutes embodied context. So what makes the embodied context of living organisms different from the embodied context of a rock? Jordan and Vinson (2012) assert that the internal dynamics of an organism are coupled to the dynamics of the organism as a whole such that changes occurring at the micro scale bring about events at the macro scale that recursively affect the micro scale and serve to sustain both the micro and macro scales of organization. For example, energy depletion at the cellular level will influence behaviors at the organismic level to increase the amount of available energy at the cellular level. This recursion enables the sustainment of the internal and whole-organism dynamics. Like organisms, the rock is an embodiment of its context (i.e., the constant tension between the strong and weak atomic forces, electromagnetism, and gravity), but the micro-macro coupling does not serve to sustain the order achieved at either the micro or macro scale. In WST, it is the self-sustaining recursion of the micro-macro coupling that distinguishes the embodied context of organisms from that of rocks and affords a particularly interesting means of being 'about' the world. It is through this self-sustaining work that organisms maintain relationships with their environment. Rocks, desks, and computers do not share this feature with organisms. From this, we argue that the WST framework provides a unique avenue for addressing the problems of consciousness, embodiment, and representation that still challenge the development of cognitive architectures.

WST and Consciousness

Some researchers are not only interested in the processes that give rise to intelligent behavior, but are also interested in identifying those properties that afford a phenomenal consciousness (Reggia, 2013). Such an investigation first requires an adequate description of what consciousness is. Following from the conceptualization of organisms as self-sustaining embodiments of context, Jordan and Ghin (2006)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

propose that consciousness is the 'aboutness' inherent in self-sustaining embodied context. That is, if a self-sustaining system constitutes an embodiment of its context, it follows that system is naturally and necessarily *about* its context. The authors consider the aboutness of a single-celled organism a "proto-consciousness" because it is the evolutionary precursor to human, embodied aboutness. This suggests that phenomenal consciousness is achieved by humans through their work's natural property of being *about* sustaining a relationship with the world. Accordingly, any system—natural or artificial—that has self-sustaining embodied aboutness would have what we call consciousness.

The problem of implementing consciousness is important to a few AI researchers (e.g., Baars, 1998; Baars & Franklin, 2009). Alternatively, others might operate under the assumption that one can ignore the "Hard Problem" of consciousness (Chalmers, 1995) and solely address the comparatively "easy" work of building an artificial system that can perceive, deliberate, and act. Such an approach could go on doing its work and create artificial systems that behave intelligently without a phenomenal consciousness. By contrast, from the perspective of WST, the problem of artificial cognition and the problem of artificial consciousness are, essentially, one and the same. In other words, if one produces an artificial system that is intelligent by the same means as organisms, one necessarily has implemented a phenomenal consciousness because the system would constitute the self-sustaining kind of embodied aboutness. To be sure, the distinction between the first perspective and WST is not intended as a criticism of the former but, rather, as an opportunity to clarify the difference between creating an artificial system that can do what we do and creating an artificial system that can do what we do *in the way that we do it*. In WST, consciousness is necessary for cognition because it is what we are (i.e., self-sustaining embodied aboutness).

WST and Embodiment

Hypothetically, one can attach arms and legs to a modern computer and program various behaviors into this robot. On the condition that its battery is nearly depleted, the robot might act so as to recharge its battery by navigating to a charging station. To be sure, one can refer to such behavior as self-sustaining. However, we would not grant this system the same self-sustaining, embodied aboutness that we ascribe to organisms. While the behavior of such a robot is self-sustaining at the organismic level, the stability at every other level of scale is taken for granted and is relatively independent of the self-sustaining work of the system as a whole. The chassis, for example, does not generate dynamics that sustain its work or serve to sustain the work of the robot. On the other hand, human neuromuscular architecture works to sustain its own integrity while simultaneously sustaining the integrity of the human. Changes that occur at the level of our neuromuscular architecture give rise to changes at the level of the organism, which, in turn, influences the neuromuscular architecture. This recursion simultaneously sustains the work of the neuromuscular architecture and the human organism, preserving their stability. Such self-sustaining micro-macro coupling would not be observed in the self-charging robot as described. One might say that the robot we described merely *has the ability* to recharge its own battery (i.e., obtain energy) while obtaining energy (i.e., self-sustainment) is what organisms *are*.

In sum, the human body is not merely a "container" for cognition. A body is yet another scale of self-sustaining work that emerges from and is sustained by all the levels of work of which it is constituted. In addition, the various levels of self-sustaining work that constitute a body are simultaneously embodiments of the contextual constraints that need to be addressed in order to afford sustainment. As a result, the human body can be coherently conceptualized as an amalgamation of internal and external

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

dynamical constraints, such as the multi-scale constraints that have to be addressed in order to propel oneself toward an energy source. Accordingly, unlike other embodiment perspectives which emphasize the importance of the body but fall short of defining exactly what a body is (Anderson, 2003), WST addresses the issue of embodiment by unifying bodies and cognition under the conceptual framework of self-sustaining systems.

Furthermore, this perspective clarifies exactly why bodies matter for cognition. Put simply, the body is an ever-present context for the brain and serves as a definable border between the entity and the external environment, albeit a border that is amenable to change (the development from a tadpole into a frog or the growth of a human newborn into an adult human as examples of typical changes to bodies' borders; loss of a limb due to an accident, surgery, or predation as an example of atypical changes to a body's border). The nature of the body as a defined border allows for cognition and behavior to reflect the simultaneous incorporation of or approach towards parts of the environment that facilitate self-sustainment as well as the rejection or avoidance of parts of the environment that destabilize self-sustainment. The self-sustaining work of a body is among the vast constellation of contextual constraints that a brain has to embody in order to sustain itself and the body to which it is recursively coupled. In short, bodies matter to brains because brains, as embodiments of context, are inherently *about* bodies.

WST and Representation

We have noted that there are those who abstain from the use of representation in describing the work of cognitive systems (e.g., Van Gelder, 1998). Specifically, researchers committed to a dynamical account of cognition tend not to refer to internal states with causal powers and, instead, view cognition as a self-organizing process spanning brain, body, and world. While WST does not refer to representations, it does not deny that cognition is constituted of content-bearing internal states. In fact, given the conception of cognitive systems as embodiments of contexts, cognition necessarily entails content (Jordan & Ghin, 2006). The internal dynamics of an organism must be about the contexts in which they are sustained, and for this reason, the internal dynamics of the organism are inherently meaningful. In other words, the internal work of an organism naturally refers to its context (i.e., it is an internalization of the contingent constraints). By replacing the notion of representation with the notion of *self-sustaining embodied context*, WST provides a means of addressing how the internal can be related to the external: "In short, the *inside* is an embodiment of the *outside*. Thus, there is no need to ask how internal states map onto external states….synchronies between internal and external events constitute resonances" (Jordan & Ghin, 2006, p. 58). The use of the word resonance—instead of representation—highlights the dynamic, interdependent relationship of the internal and external. The benefit of approaching the problem of representation from the perspective of WST is that one does not need to ask how cognitive systems represent their environments. Ultimately, self-sustaining systems entail embodied content. To clarify, from this perspective, representation is ubiquitous. Representation (i.e., embodiment) is not unique to brain dynamics or the putatively cognitive processes. It is a feature of all embodiments of contexts as they all effectively "represent" their contexts. The multi-scale self-sustaining processes that constitute a human body represent the phylogenetic, cultural, and ontogenetic contexts in which the body emerged and in which it sustains itself (Jordan & Vinson, 2012). As an embodiment of its context, it naturally and necessarily is *about* its contexts. In other words, one might say that the human body maintains "internal" states that reflect the contingent constraints. What distinguishes brains from other embodiments of context is the time scale at which they can reorganize themselves or change what they are "representing." The

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

importance of treating representation as a natural characteristic of embodied contexts is that it leads to the realization that bodies, like cognitions, are information-laden because they are saturated with *aboutness*. Strictly speaking, information is not unique to brains or cognitive systems; all embodiments of context entail information. Another way to say this is that self-sustaining embodiments are inherently semiotic systems (Emmeche, 2002; Jordan, 2008)

CONCLUSION

The description of Wild Systems Theory (Jordan, 2013; Jordan & Day, 2015) introduced an alternative way to conceptualize what it is that intelligent behavior entails. While other approaches (e.g., Soar, LIDA, and ACT-R) have made considerable progress toward developing architectures that enable intelligent behavior, Wild Systems Theory provides a glimpse at the fundamental principles that might support artificial intelligence if one is interested in creating systems that exhibit intelligent behavior *in the same way* as natural cognitive systems. In this chapter, we have set the foundation for the development of a biologically inspired cognitive architecture—a Wild architecture—one that is grounded in the notion of self-sustaining embodiment of context. Further, the proposal of consciousness as a property of self-sustaining systems sheds light on the nature of artificial systems that might potentially constitute phenomenal consciousness.

REFERENCES

Altmann, E. M., & Trafton, J. G. (2007). Timecourse of recovery from task interruption: Data and a model. *Psychonomic Bulletin & Review*, *14*(6), 1079–1084. doi:10.3758/BF03193094 PMID:18229478

Anderson, J. R. (1996). A simple theory of complex cognition. *The American Psychologist*, *51*(4), 355–365. doi:10.1037/0003-066X.51.4.355

Anderson, J. R., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, *111*(4), 1036–1060. doi:10.1037/0033-295X.111.4.1036 PMID:15482072

Anderson, M. L. (2003). Embodied Cognition: A Field Guide. *Artificial Intelligence*, *149*(1), 91–130. doi:10.1016/S0004-3702(03)00054-7

Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. Cambridge, UK: Cambridge University Press.

Baars, B. J., & Franklin, S. (2009). Consciousness is computational: The LIDA model of global workspace theory. *International Journal of Machine Consciousness*, *1*(01), 23–32. doi:10.1142/S1793843009000050

Bohl, V., & van den Bos, W. (2012). Toward an integrative account of social cognition: Marrying theory of mind and interactionism to study the interplay of Type 1 and Type 2 processes. *Frontiers in Human Neuroscience*, 6. PMID:23087631

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Borst, J. P., & Anderson, J. R. (2013). Using model-based functional MRI to locate working memory updates and declarative memory retrievals in the fronto-parietal network. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(5), 1628–1633. doi:10.1073/pnas.1221572110 PMID:23319628

Bösser, T. (1987). Learning in man-computer interaction: A review of the literature. In *Esprit Research Reports* (Vol. 1). Heidelberg, Germany: Springer-Verlag. doi:10.1007/978-3-642-83233-8

Bösser, T. (2013). A discussion of 'The Chunking of Skill and Knowledge' by Paul S. Rosenbloom, John E. Laird & Allen Newell. In B. A. Elsendoorn & H. Bouma (Eds.), *Working Models of Human Perception* (pp. 411–418). San Diego, CA: Academic Press.

Cao, S., Qin, Y., Jin, X., Zhao, L., & Shen, M. (2014). Effect of driving experience on collision avoidance braking: An experimental investigation and computational modelling. *Behaviour & Information Technology*, *33*(9), 929–940. doi:10.1080/0144929X.2014.902100

Cao, S., Qin, Y., Zhao, L., & Shen, M. (2015). Modeling the development of vehicle lateral control skills in a cognitive architecture. *Transportation Research Part F: Traffic Psychology and Behaviour*, *32*, 1–10. doi:10.1016/j.trf.2015.04.010

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies*, *2*, 200–219.

Chella, A. (2013). LIDA, Committed to Consciousness. *Journal of Artificial General Intelligence*, *4*(2), 28–30.

Chikhaoui, B., Pigot, H., Beaudoin, M., Pratte, G., Bellefeuille, P., & Laudares, F. (2009). Learning a song: An ACT-R model. *International Journal of Computer, Electrical, Automation*, *Control and Information Engineering*, *3*, 1784–1789.

Clark, A. (1997). *Being there: putting brain, body, and world together again*. London, UK: MIT Press.

Clark, A. (1998). Embodied, situated, and distributed cognition. In W. Bechtel & G. Graham (Eds.), *A companion to cognitive science* (pp. 506–517). Malden, MA: Blackwell.

Douglass, S. A., & Mittal, S. (2013). A framework for modeling and simulation of the artificial. In A. Tolk (Ed.), *Ontology, Epistemology, and Teleology for Modeling and Simulation* (pp. 271–317). Heidelberg, Germany: Springer-Verlag; doi:10.1007/978-3-642-31140-6_15

Duch, W., Oentaryo, R. J., & Pasquier, M. (2008). Cognitive architectures: Where do we go from here? In P. Wang, B. Goertzel, & S. Franklin (Ed.), *Proceedings of the First AGI Conference* (pp. 122-136). Memphis, TN: IOS Press.

Emmeche, C. (2002). The chicken and the Orphean egg: On the function of meaning and the meaning of function. *Sign Systems Studies*, *30*, 15–32.

Franklin, S. (1997). Autonomous Agents as Embodied AI. *Cybernetics and Systems*, *28*(6), 499–520. doi:10.1080/019697297126029

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Franklin, S., Ramamurthy, U., D'Mello, S. K., McCauley, L., Negatu, A., Silva, R., & Datla, V. (2007). LIDA: A computational model of global workspace theory and developmental learning. In *AAAI Fall Symposium on AI and Consciousness: Theoretical Foundations and Current Approaches* (pp. 61-66). Arlington, VA: AAAI Press.

Franklin, S., Strain, S., McCall, R., & Baars, B. (2013). Conceptual commitments of the LIDA model of cognition. *Journal of Artificial General Intelligence*, *4*(2), 1–22. doi:10.2478/jagi-2013-0002

Gallagher, S. (2015). The new hybrids: Continuing debates on social perception. *Consciousness and Cognition*, *36*, 452–465. doi:10.1016/j.concog.2015.04.002 PMID:25952957

Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Halbrügge, M., Quade, M., & Engelbrecht, K.-P. (2015). A predictive model of human error based on user interface development models and a cognitive architecture. In N. A. Taatgen, M. K. van Vugt, J. P. Borst, & K. Mehlhorn (Ed.), *Proceedings of the 13th International Conference on Cognitive Modeling* (pp. 238-243). Groningen, The Netherlands: University of Groningen.

Jaeger, H. (1998). Todays dynamical systems are too simple (Commentary on Tim van Gelders The dynamical hypothesis in cognitive science). *Behavioral and Brain Sciences*, *21*, 643–644. doi:10.1017/S0140525X98401730

Johnson-Laird, P. N. (1980). Mental models in cognitive science. *Cognitive Science*, *4*(1), 71–115. doi:10.120715516709cog0401_4

Jordan, J. S. (2003). Emergence of self and other in perception and action. *Consciousness and Cognition*, *12*(4), 633–646. doi:10.1016/S1053-8100(03)00075-8 PMID:14656506

Jordan, J. S., & Ghin, M. (2007). The role of control in a science of consciousness: Causality, regulation, and self-sustainment. *Journal of Consciousness Studies*, *14*(1-2), 177–197.

Jordan, J. S. (2008). Toward a theory of embodied communication: Self-sustaining wild systems as embodied meaning. In I. Wachsmuth, M. Lenzen, & G. Knoblich (Eds.), *Embodied Communication in Human and Machines* (pp. 53–75). Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780199231751.003.0003

Jordan, J. S. (2013). The wild ways of conscious will: What we do, how we do it, and why it has meaning. *Frontiers in Psychology*, *4*. doi:10.3389/fpsyg.2013.00574 PMID:24027543

Jordan, J. S., & Day, B. (2015). Wild systems theory as a 21st century coherence framework for cognitive science. In T. Metzinger & J. M. Windt (Eds.), *Open MIND: 21*. Frankfurt am Main, Germany: MIND Group; doi:10.15502/9783958570191

Jordan, J. S., & Ghin, M. (2006). (Proto-) consciousness as a contextually-emergent property of self-sustaining systems. *Mind & Matter*, *4*, 45–68.

Jordan, J. S., & Vinson, D. (2012). After nature: On bodies, consciousness, and causality. *Journal of Consciousness Studies*, *19*(5-6).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kauffman, S. (1995). *At Home in the Universe*. New York, NY: Oxford University Press.

Kelso, J. A. S. (1995). *Dynamic patterns*. Cambridge, MA: MIT Press.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence*, *8*, 30–43.

Laird, J. E. (2008). Extending the Soar cognitive architecture. In P. Wang, B. Goertzel, & S. Franklin (Ed.), *Proceedings of the First AGI Conference* (pp. 224-235). Memphis, TN: IOS Press.

Laird, J. E. (2012). *The Soar cognitive architecture*. Cambridge, MA: MIT Press.

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). Soar: An architecture for general intelligence. *Artificial Intelligence*, *33*(1), 1–64. doi:10.1016/0004-3702(87)90050-6

Langley, P., Laird, J. E., & Rogers, S. (2009). Cognitive architectures: Research issues and challenges. *Cognitive Systems Research*, *10*(2), 141–160. doi:10.1016/j.cogsys.2006.07.004

Lewis, R. L. (1993). An architecturally-based theory of sentence comprehension. *Proceedings of the Fifteenth Annual Conference of the Cognitive Science Society* (pp. 108-113). Boulder, CO: Lawrence Erlbaum.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco, CA: W. H. Freeman.

Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science*, *1*(2), 320–339. doi:10.1111/j.1756-8765.2009.01022.x PMID:25164936

Miller, C. S., & Laird, J. E. (1996). Accounting for graded performance within a discrete search framework. *Cognitive Science*, *20*(4), 499–537. doi:10.120715516709cog2004_2

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart and Winston. Inc. doi:10.1037/10039-000

Moon, J., & Anderson, J. R. (2013). Timing in multitasking: Memory contamination and time pressure bias. *Cognitive Psychology*, *67*(1-2), 26–54. doi:10.1016/j.cogpsych.2013.06.001 PMID:23892230

ORegan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*(5), 939–1031. doi:10.1017/S0140525X01000115 PMID:12239892

Powers, W. T. (1973). Quantitative analysis of purposive systems: Some spadework at the foundation of scientific psychology. *Psychological Review*, *85*(5), 417–435. doi:10.1037/0033-295X.85.5.417

Reggia, J. A. (2013). The rise of machine consciousness: Studying consciousness with computational models. *Neural Networks*, *44*, 112–131. doi:10.1016/j.neunet.2013.03.011 PMID:23597599

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review*, *107*(2), 358–367. doi:10.1037/0033-295X.107.2.358 PMID:10789200

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Thomson, R., Pyke, A., Trafton, J. G., & Hiatt, L. M. (2015). An account of associative learning in memory recall. *Proceedings of the 37th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Trafton, J. G., Hiatt, L. M., Harrison, A. M., Tamborello, F., Khemlani, S. S., & Schultz, A. C. (2013). ACT-R/E: An embodied cognitive architecture for human robot interaction. *Journal of Human-Robot Interaction*, *2*(1), 30–55. doi:10.5898/JHRI.2.1.Trafton

Tschacher, W., & Dauwalder, J.-P. (Eds.). (2003). *The dynamical systems approach to cognition: concepts and empirical paradigms based on self-organization, embodiment, and coordination dynamics*. Singapore: World Scientific. doi:10.1142/5395

van Gelder, T. J. (1998). The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences*, *21*(05), 1–14. doi:10.1017/S0140525X98001733 PMID:10097022

van Rooij, I., Bongers, R. M., & Haselager, F. G. (2002). A non-representational approach to imagined action. *Cognitive Science*, *26*(3), 345–375. doi:10.120715516709cog2603_7

Varella, F., Thompson, E., & Rosch, E. (1991). *The embodied mind: cognitive science and human experience*. Cambridge, MA: MIT Press.

Wilson, A. D., & Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology*, 4. PMID:23408669

Wiltshire, T. J., Lobato, E. J., McConnell, D. S., & Fiore, S. M. (2014). Prospects for direct social perception: A multi-theoretical integration to further the science of social cognition. *Frontiers in Human Neuroscience*, 8. PMID:25709572

Wiltshire, T. J., Lobato, E. J. C., Jentsch, F. G., & Fiore, S. M. (2013). Will (dis) embodied LIDA agents be socially interactive? *Journal of Artificial General Intelligence*, *4*(2), 42–47.

Wirzberger, M., & Russwinkel, N. (2015). Modeling interruption and resumption in a smartphone task: An ACT-R approach. *i-com, 14*, 147-154. doi:10.1515/icom-2015-0033

This research was previously published in Advanced Research on Biologically Inspired Cognitive Architectures edited by Jordi Vallverdú, Manuel Mazzara, Max Talanov, Salvatore Distefano, and Robert Lowe; pages 41-62; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 39
Predictive Analysis of Emotions for Improving Customer Services

Vinay Kumar Jain
Jaypee University of Engineering and Technology, India

Shishir Kumar
Jaypee University of Engineering and Technology, India

ABSTRACT

Human emotions plays an important role in everyday communication. Emotions are formed by the combination of cues such as relative actions, facial expressions, and gestures and reactions. Emotions are also present in written texts like in social media, chats, customer reviews. By getting inspired by works done in the domain of sentiment analysis, this chapter explores advances to automatic detection of emotions in text which help in Improving Customer Services. This chapter presents a framework for automatic detection of emotions in customer reviews based on different emotions theories in the fields of psychology and linguistics. This framework uses advanced Machine Learning (ML) techniques with Natural Language Processing (NLP) methods for better understanding of emotion detection and recognition in customer reviews. The text under study comprises data collected from leading Indian e-commerce portals like Flipkart, Snapdeal and Amazon, which contains text rich in emotions. The advantages and application based emotion detection framework has been incorporated with suitable examples.

INTRODUCTION

Language is the most important tool for communication and help to convey information in a society. Language provides suitable means to express emotions in several forms. Natural Language Processing techniques are used to identify meaningful inferences content in text. Multiple applications related to information retrievals such as topic-based modeling, text categorization, question-answering systems, has been focused on the information contained in the data (Jain & Kumar, 2015).

DOI: 10.4018/978-1-7998-0951-7.ch039

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

During the last few decades around the world, the growth of the internet has increased the communication levels and internet users widely used multiple platforms to post their opinions online. In current scenario, people increasingly use micro-blogging platforms and social networking sites, to share and retrieve relevant information in real-time which contains opinions and sentiments (Chesley et al.,2006). Intelligent computational tools have been used to extract meaningful inferences and help companies to utilize this information for their productivity.

The e-shopping websites provides the feature where a customer can give feedback in the form of review to the product he purchased. There are massive amount of reviews available for every product in multiple e-commerce portals. A simple example related to customer review related to smart phone has been presented in Figure 1. Most of the users view the rating given to the product and read some of the reviews before decision making. The ratings are not enough to decide whether a product is feasible to buy or not. Therefore getting a conclusion from reviews is manually a tough task.

Much of the current work for analysis of customer reviews has been focused on polarity orientation of textual data. Limited work has been done in the direction of type of emotions present in written text by the customers. Recognizing emotions that has been expressed by customer reviews could provide the general intention of customer towards products and services. This chapter presented an emotion detection based system to provide benefit to the customer in decision making before the purchase of product or services. The Emotions categories in customer reviews definitely help companies to enhance their marketing strategy in a better way. .

The methods related to automatic emotion detection can be useful in many real-time applications with the help of human psychological techniques (Dung and Cao,2012). In recent years, the growth of e-commerce has brought a boom in the business market. The concept of online shopping has become one of the biggest successful businesses. With the availability of products on the websites, the interest of public has grown towards online shopping. Whenever any user is interested in purchasing a product online, the user completely depends on the reviews that have been given for a particular product. This review can be oral or written. On internet most of the emphasis is given on written review that a product receives, and this helps the user to make a decision whether to buy a product or not. So, it is necessary to build a system that finds out the emotions present every review by users so that they can but the product It also helps the company owners to determine whether the product they launched in the market is been accepted by users or not. In concern of social media or the customer reviews data it has more complications due different geographical location. A wide variety of work has been carried out in the field of sentiment analysis and opinion mining but limited work has been carried out in the field emotion detection and recognition. This chapter presents a framework for automatic detection of emotions in customer reviews using different emotions theories in the fields of linguistics and psychology. This framework uses Machine Learning methods and Natural Language Processing (NLP) and for better understanding of emotions present in customer reviews. A simple example of customer reviews which contains emotion bearing words has been presented in Figure 1.

BACKGROUND

Emotional states have multiple cognitive bases which are formed by number of factors. The scope of this chapter is limited to provide important features relevant to recognize emotion and the process of determining the emotional orientation of customer reviews. Emotion present in multiple languages or

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. A sample of reviews collected from E-commerce website

Very nice phone in terms of looks, weight, but the Flyme OS is little tricky and not at all user friendly, If you can bear the same you may go for it rest of all is fine..

Nice phone. camera, battery backup, look etc all are perfect at this price range.

hangs, touch is too bad at end of screen...never trust this model... above all where is the support...amazon just washed it by hands return policy over so can't make replacement !! Who will repair or replace !!

Ultimate mobile phone,have used HTC, SAMSUNG, APPLE and MOTOROLA but this one is not behind them in any way.

This is very fast and all mobile future very different then other mobile I love miezu

All features is gud in this phone but i have only one problem that i can't use 2g network only 3g and wifi is working

I have also a k3 and a mi 4i.performance wise k3 is best.but in camera look sound mi is best meizu is gd in performance touch experience design.....thanks

simply in a language can be expressed in several ways and understanding of this expression can be interpreted differently by different readers.

During the last few decades, Computational techniques such as Natural language processing, Artificial Intelligence (AI), Machine learning has been used to develop intelligent systems that interpreted human emotions (Jain et al., 2016). Several real-time processes have been modeled to develop effective intelligent systems by providing learning, perception and reasoning. Development of affective interfaces is an important research area in emotions. Affective interfaces are the system which provides emotional inputs; emotional responses using this facilitate online communication with the help of animated affective agents. These affective interfaces provide better user experience in following areas such as Computer-Mediated Communication (CMC) and Human-Computer Interaction (HCI). Text-to-Speech (TTS) synthesis systems are also used these interface for developing real-time systems.

Recognizing the polarity of sentiment present in written text such as in social media posts, blogs and forums is the most popular in sentiment and opinion mining domain (Pang and Lee, 2008) .Among the less paying attention area related to sentiment is the identification of emotions present in customer reviews and social media data. Computational techniques related to emotion analysis present in text mainly deals with their emotion modalities and emotion-annotated data (Jain et al., 2016).However, only limited work has been done in developing automatic emotion recognition system which benefits customer and e-commerce companies in decision-making. The focus of this chapter is to give advantages of recognizing emotions present in customer reviews which provide an insight related to customer's intentions.

THEORIES OF EMOTION

Theories of emotion mainly categorized in terms of the context within which the explanation is developed. There are different emotion theories are used in physiology and other relevant fields (Cowie et al.,2011).Recently, researchers have investigated a number of aspects related to human emotions in multiple domains. Authors categorized emotions terms in multiple emotion classes based on emotion-bearing words which are accepted across the world (Picard, et al [4]).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Multiple milestone work has been carried out by researchers to provide right judgment related to emotion categories (Tomkins 1962; Plutchik 1980; Izard 1977; Ortony et. al., 1988; Raghavan, 2007 and Ekman, 1992). Emotion categories recognized by the different researchers has been presented in Table 1.

According to the popular emotion theory (Rasa theory) given by Bharta Muni, emotions are the primary gastric juices which represent every part of the world into the sentiments (Raghavan,2007). Rasa emotional theory is based on the Natyashastra (the Textbook on Drama) and given by Bharata Muni. According to Natyashastra(Raghavan,2007)., there are eight rasas corresponding to Bhava (mood). Ekman(Ekman,1992) has defined six basic emotions on the basis of distinctive facial expressions which are universally accepted has been presented in Table 1.

Watson and Tellegen in 1985 has presented a compressive theory i.e The Circumplex Theory of Affect which identifies two main classes as positive and negative affects which range from high to low(Watson and Tellegen, 1985). Osgood's in 1957 has proposed theory of Semantic Differentiation. (Osgood et al., 1957).In this theory words are assigned by emotive meanings using three factors such as activity factor, evaluative factor and potency factor.

Clore et al., proposed number of words which convey emotion explicitly, while other are depending on the context (Clore et al.1987). The classification of affective words into two classes 'indirect affective words' (implicit) and 'direct affective words' (explicit) has been performed by Strapparava and Valitutti (Strapparava and Valitutti, 2006).The chapter has utilized both these types of words given by Strapparava and Valitutti (Strapparava and Valitutti,2006).

EMOTION RECOGNITION

Emotion recognition in text is just one the numerous dimensions of the task of making the computers make sense of and respond to emotions. Emotion represents a complex, subjective experience which involves thinking, excitement, feeling and activation. This can be seen using native languages, physiological changes, and behavioral changes. The word "affect" is often used alternatively with "emotion" in the literature. Human emotion can be sensed from such cues as facial expression, gestures, speech and text. Research in emotion has focused on all these aspects (Cowie et al.,2011).

Table 1. Emotion categories or classes identified by authors

Muni (Raghavan,2007).,	Tomkins (1962)	Izard (1977)	Plutchik (1980)	Ortony et.al. (1988)	Ekman (1992)
		Disgust			
Hasya (Mirth)	Joy	Enjoyment	Joy		
Krodha (Anger)	Anguish	Sadness	Sorrow	Joy	Fear
Rati (Love)	Shame	Fear	Anger	Sadness	Anger
Bhaya (Terror)	Fear	Anger	Fear	Fear	Happiness
Jugupsa (Disgust)	Anger	Surprise	Disgust	Anger	Sadness
Soka (Sorrow)	Disgust	Shame	Surprise	Disgust	Surprise
Vismaya (Astonishment)	Surprise	Shyness	Acceptance	Surprise	Disgust
Utsaha (Energy)	Interest	Guilt	Anticipation		
		Interest			

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Emotion detection has been focused on the sentence-level analysis or document level analysis for learning emotions in customer reviews. Emotional words are used as keywords to identifying emotion in input sentences. Osgood used multidimensional scaling to visualize the affective words for calculating similarity ratings between them. Three dimensions were used by Osgood which includes "evaluation", "potency" and "activity", where evaluation can be used to measure how much a word can refer to a pleasant or an unpleasant event (Osgood et al., 1957).

Read et al. have carried out appraisal annotation of a corpus of book reviews; a genre that provides ample instances of the various kinds of appraisal classes (Read et al.2007). They found that out of the three subsystems of the Appraisal Framework, the attitude subsystem's instances (which include emotion expressions) were the easiest to identify. Mihalcea and Strapparava present results in favor of automatic recognition of humor in texts (Mihalcea and Strapparava, 2005). They perform experiments to identify humorous one-liners, which are one-line sentences generally characterized by simple syntax and use of rhetoric which gives them a humorous connotation. Neviarouskaya et al. propose a system for augmenting online conversations with a graphical representation of the user, which displays emotions and social behavior in accordance with the text (Neviarouskaya et al,2007). This system performs automatic estimation of affect in text on the basis of symbolic cues such as emoticons, popularly used IM (Instant Messaging) abbreviations, as well as word, phrase, and sentence-level analysis of text.

Ghazi et al. proposed a multiple levels of hierarchy classification to classify emotions words (Ghazi et al., 2010). Strapparava et al. has developed a linguistic resource to lexically represent affective knowledge named WordNet – Affect (Strapparava et al., 2006). The WorldNet-Affect contains affective words and available freely as open-source lexical recourses. Wang et al [63] proposed a novel approach to detect emotion from Chinese language. The algorithm proposed was segment based fine grained emotion detection model that is a supervised learning approach.

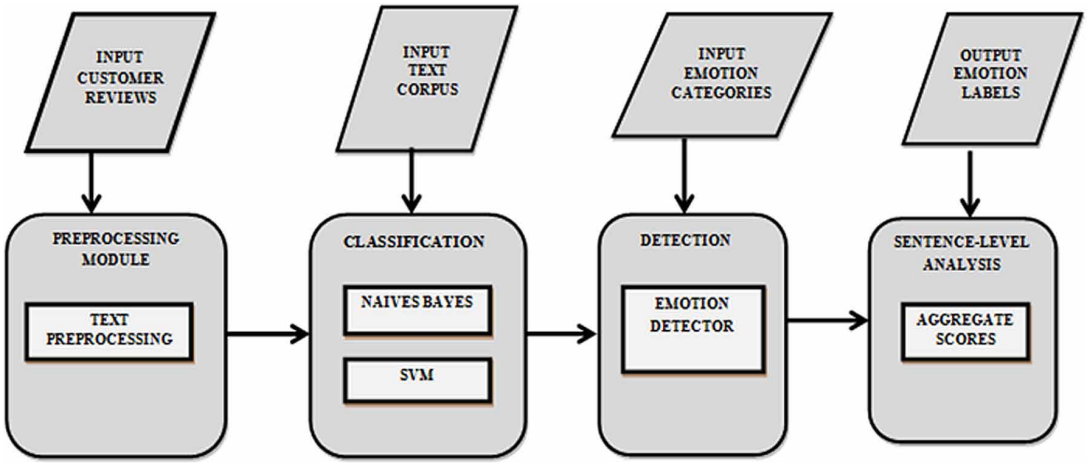
Lie Dey proposed a system for extracting emotions from the data collected from chat messenger (Dey et al.,2014). The author built a lexicon of emotion conveying words to extract emotions from the data sets. Shaheen et al. had proposed a framework for classification of emotions based on generalized concepts extracted from the English sentences (Shaheen et al.,2014). Perikos and Hatzilygeroudis developed a system that can automatically recognize emotions in natural languages (Perikos and Hatzilygeroudis, 2013). They used Tree tagger and Stanford Parser. They also used WordNet Affect lexical resource in order to spot the emotional words. They then analyzed emotional words dependencies in order to specify the strength of emotional word's strength and also determined overall sentence emotional statues that are based on dependency graph of sentence.

EMOTIONS DETECTION FRAMEWORK

The objective of this framework is to provide an insight to the process of automatically recognize emotions from customer reviews. Experimental study using appropriate corpus of text (customer reviews) has been presented using this framework for emotion recognition. Research in automatic text-based emotion analysis is limited due to unavailability emotion-annotated data for written text. To solve this problem, data has been collected from the top Indian online e-commerce websites like Snapdeal, Flipkart and Amazon under multiple categories such as LED television, smart phones, laptops, tablets, cameras.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. An overview of emotion detection framework



Emotion detection in text is referring to as a classification problem. In this process multiple nominal labels are assigned to a text or sentence from a group of emotion labels. Figure 2 presents an overview of emotion detection framework for customer reviews data analysis. Presented framework for emotion detection for customer reviews has followed the methodology given as:

Let $\boldsymbol{c}$ is a customer review corresponding to a product or services and $\boldsymbol{k}$ is an emotion label. Let $\boldsymbol{e}$ be a set of $\boldsymbol{n}$ probable emotion categories where $\boldsymbol{e} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3 \ldots\ldots\ldots \boldsymbol{e}_n\}$. The classification process has been used to label $\boldsymbol{c}$ with the finest promising emotion label $\boldsymbol{k}$, where $\boldsymbol{k} \in \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3 \ldots\ldots\ldots \boldsymbol{e}_n, \text{neutral}\}$.

The framework of emotion detection is divided into four modules:

Pre-Processing Module

In this module customer's reviews data has been pre-processed to remove unwanted noise present in the data. Segmentation and Tokenization has been applied in first step. Sentence segmentation has been applied to form sentence chunks. Tokenization is has been applied to break stream of text up into multiple forms such as phrases, symbols, words, or other meaningful elements called tokens..In second step stop words removal process has been applied to filter out unnecessary words. Stemming is done on tokenized input.

Classification Module

Automatic classification of emotional sentences need a feature sets. The most appropriate features that distinguish emotional expressions from non-emotional ones are the emotion words. To recognize these words in the sentences or customer reviews, publicly available lexical resources are available such as – WorldNet-Affect (Strapparava and Valitutti, et al., 2006).Researchers have identified various categories of basic and secondary emotions. Ekman six basic emotions is one of the most popular models in various domains (Ekamn, 1992). Detection of emotional orientation of text has been performed naïve approach by determining emotion words, such as "happy", "terrified", and "amazed" in customer reviews. The clear advantage of this approach is that it requires no training data.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Description of emotion word list that has been extracted from WordNet affect

Emotion Category	#synsets	Sample Words
Happiness	227	joy, love, rejoicing, glee, happiness, euphoria, enthusiasm, admiration, cheerful, content, happy, merrily
Sadness	123	sorrow, misery, woe, gloom, grief, depression, heartbreak, mournful, hapless, guilty, sadly, remorsefully, repent, bored
Anger	127	wrath, fury, rage, angry, annoyed, pissed, mad, sore, livid, displeasingly, aggressive, hateful, hostile, malice, spite, resentfully
Disgust	19	repulsion, nauseous, foul, abhorrent, fed_up, abominably, revolt, sicken
Surprise	28	wonder, fantastic, marvelous, baffle, bewilder, astonishing, awestruck, stupefy, dumbfound, staggering
Fear	82	scary, fright, panic, terror, chilling, frightful, terrible, intimidate, dread, anxiously, apprehension

Emotion Detection Module

After the classification of emotional words from the customer reviews a baseline system called as emotion detector has been used to counts the number of emotion words of each category in a sentence, and then assigns the category with the largest number of words to the sentence. Ties were resolved by choosing the emotion label based on a random predefined ordering of emotion classes. A sentence containing no emotion word of any type was assigned the *no-emotion* category. Emotion detector gives an insight of overall category present in Table 2 of a particular customer review of a product.

Sentence-Level Analysis

Sentence-level analysis provides aggregate score of emotion present in all the product categories with emotion levels of customer review. This provides emotion related to specific brand and categories.

IMPORTANCE OF EMOTIONS DETECTION

Emotion Detection in customer reviews is important for e-commerce industries because it identifies meaningful ways to improve their operations and marketing strategy for their products and services. This can be accounted by measuring the public's emotions on products, services, and advertisements. E-commerce companies can use emotion detection framework for prediction of their services based on diffusion of innovation theories. This framework is also helpful in new product launch in the market by examine public emotion towards the product or services. According to marketing survey agencies, customer reviews provides new battleground for brands. E-commerce companies focused to build loyalty with product buyers, which help to grow and expand their business. New emerging platforms based on social media provide rich source of information about the product or services. Emotion detection framework can help companies in measuring public feeling related to products and services.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ADVANTAGES OF EMOTION DETECTION

E-commerce companies always focused to build long term relation with their customers. Customer services such as product delivery timing, response to their query etc are mainly focused by e-commerce companies. The e-commerce should always aware their customer response about product or services they sale on their portal. Emotion detection framework can easily give real-time information about customer's mood and feelings. Emotion detection play an important role to gain advantages over other techniques or policies and helps companies in improving their market strategy. Following are the key benefits provided by emotion detection framework:

1. **Accessibility:** Emotion detection framework provides real-time emotions 24X7.
2. **Gain Business Intelligence:** This framework provides meaningful inference which help business to take rapid decisions.
3. **Quickly Identify the Tonality:** Framework provides real-time trends related to products and services with emotion spikes based on emotion models.
4. **Easily Built Engagement Platform:** Measurement of effectiveness of promotional offers and viral advertisements can easily be understood with the help of this emotion detection. This will help in better engagement with customer to build up long relationship.
5. **Beat Competition:** Track competitor responses help in improvement of action plan over time.
6. **Alerts Service:** Alerts can be set up for the brands to be notified of specific emotion labels and rapid decision cans be easily taken out.
7. **Improve Customer Service:** The major advantages of this framework help in improving customer services which is the major challenge for e-commerce companies in current scenario.

CONCLUSION

This chapter describes the potential of real –time emotion detection system based on customer reviews present in e-commerce portals. This framework can help in providing solutions to number of problems related to businesses and marketing. This chapter present an emotion detection framework based on Ekman emotion model. Methods and technique are presented help in knowing customers emotions with help of product reviews. This framework gives important and valuable knowledge towards improvement in departments such as customer management, customer relations & marketing, product management and corporate management.

The emotion detection framework presented in this chapter can be pursued further in several domains. One of the tasks is to consider emotion intensity for classification. Explore the relation between emotion classes and emotion intensity.

REFERENCES

Chesley, P., Vincent, B., Xu, L., & Srihari, R. (2006). Using Verbs and Adjectives to Automatically Classify Blog Sentiment. *Proceedings of the AAAI-2006 Spring Symposium on Computational Approaches to Analyzing Weblogs*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Clore, G. L., Ortony, A., & Foss, M. A. (1987). The psychological foundations of the affective lexicon. *Journal of Personality and Social Psychology*, *53*(4), 751–766. doi:10.1037/0022-3514.53.4.751

Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., & Taylor, J. (2011). Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine*, *18*(1), 32–80. doi:10.1109/79.911197

Dey, L., Afroz, N., & Nath, R. P. D. (2014) Emotion extraction from real time chat messenger. *Proceedings of 3rd International Conference on Informatics, Electronics & Vision*, 1-5. 10.1109/ICIEV.2014.6850785

Dung & Cao. (2012). *A high-order hidden Markov model for emotion detection from textual data*. Springer Berlin Heidelberg.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, *6*(1), 169–200. doi:10.1080/02699939208411068

Ghazi, D., Inkpen, D., & Szpakowicz, S. (2010). Hierarchical versus flat classification of emotions in text. *Proceedings of the NAACL HLT 2010 workshop on computational approaches to analysis and generation of emotion in text*, 140-146.

Izard, C. E. (1977). *Human emotions*. New York: Plenum Press. doi:10.1007/978-1-4899-2209-0

Jain, V. K., & Kumar, S. (2015). An Effective Approach to Track Levels of Influenza-A (H1N1) Pandemic in India Using Twitter. *Procedia Computer Science*, *70*(1), 801–807. doi:10.1016/j.procs.2015.10.120

Jain, V. K., Kumar, S., Jain, N., & Verma, P. (2016). A novel Approach to Track Levels public emotions related to epidemics in multilingual data. *2nd International conference and Youth School Information technology and Nanotechnology,* 883-889.

Mihalcea, R., & Strapparava, C. (2005). Making Computers Laugh: Investigations in Automatic Humor Recognition. *Proceedings of the Joint Conference on Human Language Technology/Empirical Methods in Natural Language Processing*, 531-538.

Neviarouskaya, A., Prendinger, H., & Ishizuka, M. (2007). Analysis of affect expressed through the evolving language of online communication. *Proceedings of the 12th International Conference on Intelligent User Interfaces,* 278-281. 10.1145/1216295.1216346

Ortony, A., Clore, G. L., & Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge University Press. doi:10.1017/CBO9780511571299

Osgood, C. E., Succi, G. J., & Tannenbaum, P. H. (1957). *The Measurement of Meaning*. Urbana, IL: University of Illinois Press.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, *2*(1), 1–135. doi:10.1561/1500000011

Perikos, I., & Hatzilygeroudis, I. (2013). *Recognizing emotions present in Natural Language Sentences*. University of Patras.

Plutchik, R. (1980). A General Psycho evolutionary Theory of Emotion. In Emotion: Theory, Research and Experience: Theories of Emotions (vol. 1). New York: Academic.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Raghavan, V. (1967). *The Number of Rasa-S*. Theosophical Publishing House.

Read, J., Hope, D., & Carroll, J. (2007). Annotating expressions of Appraisal in English. *Proceedings of the ACL-2007 Linguistic Annotation Workshop*, 93-100.

Shaheen, S., El-Hajj, W., Hajj, H., & Elbassuoni, S. (2014). Emotion Recognition from Text Based on Automatically Generated Rules. *2014 IEEE International Conference on Data Mining Workshop*, 383-382. 10.1109/ICDMW.2014.80

Strapparava, C., Valitutti, A., & Stock, O. (2006). The affective weight of lexicon. *Proceedings of the Fifth International Conference on Language Resources and Evaluation,* 423–426.

Tomkins, S. S. (1962). *Affect, imagery, consciousness. In The positive affects* (pp. 20–31). New York: Springer.

Watson, D., & Tellegen, A. (1985). Towards a consensual structure of mood. *Psychological Bulletin*, *98*(2), 219–235. doi:10.1037/0033-2909.98.2.219 PMID:3901060

KEY TERMS AND DEFINITIONS

Machine Learning: It is a science of getting computers to act without being explicitly programmed.

Natural Language Processing: It uses methods of computer science, artificial intelligence, and computational linguistics to understand interactions between computers and human languages.

Opinion Mining: Categorization of opinions.

Text Classification: Classification of documents into a fixed number of predefined categories.

This research was previously published in Applying Predictive Analytics Within the Service Sector edited by Rajendra Sahu, Manoj Dash, and Anil Kumar; pages 125-134; copyright year 2017 by Business Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 40
A Fuzzy Logic Based Synonym Resolution Approach for Automated Information Retrieval

Mamta Kathuria
YMCA University of Science and Technology, India

Chander Kumar Nagpal
YMCA University of Science and Technology, India

Neelam Duhan
YMCA University of Science and Technology, India

ABSTRACT

Precise semantic similarity measurement between words is vital from the viewpoint of many automated applications in the areas of word sense disambiguation, machine translation, information retrieval and data clustering, etc. Rapid growth of the automated resources and their diversified novel applications has further reinforced this requirement. However, accurate measurement of semantic similarity is a daunting task due to inherent ambiguities of the natural language, spread of web documents across various domains, localities and dialects. All these issues render to the inadequacy of the manually maintained semantic similarity resources (i.e. dictionaries). This article uses context sets of the words under consideration in multiple corpora to compute semantic similarity and provides credible and verifiable semantic similarity results directly usable for automated applications in the intelligent manner using fuzzy inference mechanism. It can also be used to strengthen the existing lexical resources by augmenting the context set and properly defined extent of semantic similarity.

DOI: 10.4018/978-1-7998-0951-7.ch040

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

With the passage of time, human dependency has increased manifold on the automated resources. These resources depend upon the underlying inference mechanism and stored data for their decisions/actions. Unlike humans, these resources are devoid of their own intelligence and analytical capability. Therefore, it is necessary that these systems be embedded with the resources which are totally free from ambiguity so that their decisions/actions are reliable and credible. As and when an automated system has to deal with an inherently ambiguous environment such as Natural Language Processing(NLP), there are two options available to the designer to increase the reliability and credibility of the system:

1. To deal directly with the ambiguity at the inference level by designing the knowledge base with the help of tools like fuzzy logic which have the ability to handle the ambiguity to an extent and provide a coarse (not that precise) automated solution;
2. To make the underlying data less ambiguous and more suitable for automated applications.

In an automated system, which works in an inherently ambiguous environment, both the approaches may be required but an effort in the direction of second approach shall reduce the requirement of first approach making the system more precise.

One of the most frequently used automated systems in the present-day environment is *web search engine,* commonly referred to as search engine. It tries to provide a set of documents semantically related to the query submitted by the user in the natural language. For getting better results, the search needs to be widened by using terms semantically similar to input query. This leads to the requirement of unambiguous, efficient lexical resources providing precise semantic similarity information directly usable by the automated systems. The heterogeneity in web literature owing to its global spread over geographical boundaries, languages, dialects combined with other inherent ambiguities in the natural language makes this task quite difficult which is beyond the grasp of manually maintained dictionaries. To overcome this drawback, the current literature provides several empirical methods based upon web page count and text snippets count as discussed in literature survey presented in this paper. These methods basically fail to take into account the exact context of the word pair usage due to following reasons:

- Occurrence of a pair of words within the same page does not ensure the similarity of their contexts;
- Sometimes the size of the snippets is increased to ensure the larger number of instances which may result in the loss of context.

To ensure the better efficiency and credibility of these web based automated systems such as search engines, it is necessary that the underlying lexical resources be automation friendly and as less ambiguous as possible. An information retrieval system based upon the crude semantic similarity taken from online lexical sources like WordNet (2005), etc., without considering the context, leads to fetching of large number of undesired pages which otherwise may not be that useful. Therefore, there is a need to tune these lexical resources in such a manner that the derived synonyms are also provided with the following:

- Their respective applicable context set;
- Properly defined extent of semantic similarity.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

It shall help the automated systems in choosing an appropriate synonym set.

The proposed work accomplishes the above-mentioned goals by defining a credible synonym resolution process. The work intends to pave the way for the design of next generation lexical resources and shall be helpful in the areas of natural language processing, information retrieval, web mining, word sense disambiguation and automated machine translation.

The paper has been organized as follows: The introduction section of the paper is followed by the literature survey section. After taking a look on the available literature, the objectives of the proposed work have been defined. Thereafter, the details of the proposed work have been discussed. After discussing the proposed work, results have been presented and analysed. Last section concludes the work with possible future extensions.

LITERATURE SURVEY

The inadequacy of the manually created lexical resources has been long felt due to their inability to cater various diversified domains of knowledge area such as engineering, medical, music and finance etc. Moreover, mostly their updation is also based upon the perception of very few people. With the growth of web over multifaceted heterogeneous environments spread over variety of domains, people, nationalities, dialects etc., the limitations of these resources are being strongly felt than ever. Moreover, the need for meaningful query expansion has gained more and more prominence wherein the lexical resources are unable to provide the requisite support to the web search engines. This issue can only be addressed through proper semantic resolution which in turn requires knowledge about exact estimate of the semantic similarity between the terms. The growth of the web has pushed the research in the area of precise semantic similarity measurement to help web users in getting the accurate results through proper semantic interpretation of their queries. In this section, a look at these contributions is being taken.

Rada, Mili, Bicknell, Blettner (1989) proposed a method of semantic similarity by taking into consideration the taxonomy of words. Shorter the path in taxonomy, more similar the words are.

Resnik (1995) proposed semantic similarity measurement between two concepts using Brown Corpus (1991) and WordNet(2005). The proposed method finds the similarity between the two concepts C1 and C2 on the basis of a third concept C that subsumes both the concepts and is the biggest of all such possible concepts.

Li, Bandar & Mclean (2003) proposed a semantic similarity measure based on various features such as path length, depth and local density of the taxonomy. They combined these structural features of the taxonomy along with the information model.

A double-checking model using text snippets was suggested by Chen, Lin & Wei (2006) for measuring semantic similarity between two words. For computing the similarity between words P and Q, occurrences of word P are counted in the snippets for word Q and similarly the occurrences of word Q are counted in the snippets for word P. Then, the corresponding values are used to calculate the similarity between P and Q by using Co-occurrence based Double-Checking (CODC) measure as defined below:

$$CODC\left(P,Q\right)=\begin{cases}0, & if\ f\left(P\,@\,Q\right)=0\\ exp\left(log\left[\dfrac{f\left(P\,@\,Q\right)}{H\left(P\right)}\times\dfrac{f\left(Q\,@\,P\right)}{H\left(Q\right)}\right]^{\alpha}\right) & otherwise\end{cases} \tag{1}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

where f(P@Q) is the number of occurrences of P in the top-ranking snippets of Google corresponding to query Q, H(P) denotes the page count for query P, and α is a constant which was experimentally set to the value 0.15 by the authors. The drawback of the method is its dependency on the ranking method adopted by a search engine in use. It happens because a search engine considers publication date and link structure along with semantic similarity while ranking the search results corresponding to the user query.

Initial efforts to make a decision about semantic similarity based upon web content were relying upon page count. In these cases, page count of the word pair (say P and Q) is taken into consideration for computation of semantic similarity. Here, H(P) and H(Q) indicate the page counts for word P and Q respectively and H(P∩Q) denotes the page count for conjunctive query 'P AND Q'. Based upon the page count, various indices have been proposed, which are as follows:

$$WebJaccard\left(P,Q\right)=\begin{cases}0, & if\,H\left(P\cap Q\right)\le c\\ \dfrac{H\left(P\cap Q\right)}{H\left(P\right)+H\left(Q\right)-\left(H\left(P\cap Q\right)\right.}, & otherwise\end{cases} \tag{2}$$

This coefficient is set to zero if page count H(P∩Q) for the conjunctive query is less than a predefined threshold c:

$$WebOverlap\left(P,Q\right)=\begin{cases}0, & if\,H\left(P\cap Q\right)\le c\\ \dfrac{H\left(P\cap Q\right)}{\min(H\left(P\right),\,H\left(Q\right))}, & otherwise\end{cases} \tag{3}$$

$$WebDice\left(P,Q\right)=\begin{cases}0, & if\,H\left(P\cap Q\right)\le c\\ \dfrac{2H\left(P\cap Q\right)}{H\left(P\right)+H\left(Q\right)}, & otherwise\end{cases} \tag{4}$$

The fourth measure WebPMI (Pointwise Mutual Information) reflects the independence between two probabilistic events and is defined as:

$$WebPMI\left(P,Q\right)=\begin{cases}0, & if\,H\left(P\cap Q\right)\le c\\ log_2\left(\dfrac{H\left(P\cap Q\right)\big/N}{H\left(P\right)\big/N\cdot H\left(Q\right)\big/N}\right), & otherwise\end{cases} \tag{5}$$

where N is the number of documents being indexed by the search engine. Probabilities are assessed based on maximum likelihood principle.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

With the advancement in search engine technology, new empirical methods were defined in the light of search results. Cilibrasi & Vitanyi (2007) defined a semantic similarity measure called Normalized Google Distance (NGD) based upon page count provided by the search engine. The NGD measure can be defined as given in Equation (6):

$$NGD\left(P,Q\right)=\frac{\max\left\{\log H\left(P\right),\log H\left(Q\right)\right\}-\log H\left(P,Q\right)}{\log N-\min\{\log H\left(P\right),\log H\left(Q\right)\}} \tag{6}$$

All these methods consider the co-occurrence of words without taking into account the context of the words in which they occur on the page.

Bollegala, Matsuo, and Ishizuka (2011) proposed an empirical method to measure semantic similarity between two words based on page counts and text snippets retrieved from a search engine in response to the said words. They identified several semantic relations between given words by the using pattern extraction and pattern clustering algorithms. They also used Support Vector Machine (SVM) based supervised learning model for deciding the appropriate boundaries of classification.

Specia, Jauhar, & Mihalcea(2012) provided a mechanism for Lexical Simplification which involves the replacement of words and phrases through their simpler variants using complexity analysis, substitute lookup and context-based ranking.

Jurgens, Mohammad, Turney & Holyoak(2012) have focussed on relational similarity such as entity:sound (dog:bark, cat:meow), cause:effect (virus:flu) etc. for the purpose of semantic similarity resolution.

Pilehvar, Jurgens, & Navigli, (2013) have presented a unified approach for finding semantic similarity that operates at multiple levels from comparing word senses to comparing text documents. The method leverages a common probabilistic representation over word senses in order to compare different types of linguistic data. This unified representation shows state-of-the-art performance on three tasks: semantic textual similarity, word similarity and word sense coarsening.

Severyn, Nicosia, & Mchitti (2013) have taken up the task of semantic textual similarity (STS) using a large number of pair-wise similarity features. Their model includes: encoding of input texts into relational syntactic structures, use of tree kernels to handle feature engineering, combining both structural and feature vector representations in a single scoring model using Support Vector Regression (SVR). The contribution of the work is quite significant in the area of semantic textual similarity.

Adhikesavan (2015) has proposed a method which takes into account association rule mining, classification, clustering and information available on the web for measuring the similarity between words/ sentences. The hallmark of the proposed scheme is that it can be adapted for different domains. The author evaluates the results using various datasets which shows its performance over other competing methods.

Recski, Iklódi, Pajkossy and Kornai (2016) have proposed a hybrid system that measures the semantic similarity between a set of word pairs using WordNet and concept dictionary. The concept dictionary with its ability to encode semantic relations between words using concept graphs improves the performance of the system which is not otherwise possible using standard tools like Wordnet. The drawback of the approach is inability to find the semantic relation between phrases.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ranjbar, Mashhadirajab and Shamsfard (2017) have proposed a method for evaluating semantic similarity between the words of Persian language on the basis of corpus and knowledge based methods. The corpus based method uses Word2Vec and knowledge based method uses FarsNet (lexical ontology). The drawback of the work is that it uses only one corpus and its application is limited to Persian Language only.

Hakami, Hayashi and Bollegala (2017) have proposed a bilinear operator called PairDiff to compute the vector offset between the corresponding word embeddings. The effectiveness of the operator has been checked empirically on several benchmark data sets. However, the work lacks a critical theoretical evaluation.

Elekes, Schäler and Böhm (2017) have analysed the statistical distribution of similarity values systematically using two series of experiments. First experiment examines the distribution of similarity values depending upon different embedding-model algorithms and parameters. In the second one, they have tried to show that intuitive similarity thresholds do not exist. They have been able to show that though word embedding models have yielded meaningful results in many cases yet it is necessary to examine when exactly similarity values in word embedding models are meaningful.

Falho, Rodrigues, Coheur and Quaresma (2017) in their work have studied Semantic Textual Similarity and Multilingual Word Similarity for English and Spanish languages using a rich set of features. Their goal is to move from lexical to semantic features on the basis of Abstract Meaning Representation and SMATCH measure by combining *WordNet* information with Word Embeddings. The hallmark/ drawback of the approach is its simplicity/ naiveness.

Bjerva and Östling (2017) in their work have assessed the semantic similarity between sentences which can be in same or different languages. They have employed multilingual word representations for similar words of different languages which are semantically close to each other. The languages used are English, Spanish, Arabic, and Turkish. The hallmark of the scheme is that the semantic similarity can be inferred even for languages for which no annotated data exists. The drawback of the proposal is that the system fails to deliver requisite performance on the shared task test sets.

After having a critical look on the literature, the following drawbacks have been identified:

- Mere page count based methods are quite crude;
- A text snippet may be quite long;
- Word taxonomy may depend upon various heterogeneous aspects;
- Processes like concept identification and relation identification are perception dependent;
- Word embedding models are not that consistent and their applicability depends upon case to case basis.

These challenges have been the motivating factors for the proposed work. The next section defines the objective of the proposed work.

OBJECTIVE OF THE PROPOSED WORK

The main objective of the work is to propose a synonym resolution method based upon the immediate context of the word under consideration in various corpora. The outcome of the work includes context set identification for a word and computation of an index indicating the extent of semantic similarity

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

between a pair of words based upon the similarity of their context set. The lists of the contexts for the word under consideration were obtained from various corpora by using their respective search utilities. The computed index was then fuzzified in order to design a Fuzzy Rule Base (FRB) for the purpose of automation in the web search engines and other such applications.

PROPOSED WORK

In the area of the information retrieval, the semantic aspect of the query is quite vital. As and when a user submits a query to the search engine, he/she expects to retrieve a set of documents which are semantically related to the query submitted. To accomplish this, it is required that users' query be expanded to cover more and more expressions which are semantically same. A normal keyword based information retrieval system is unable to use the concept of semantic similarity thereby missing the major chunk of otherwise relevant pages. To overcome this drawback, it is necessary that query be expanded through meaningful semantic extensions using appropriate choice of concepts/words. This appropriate choice requires the precise knowledge about semantic similarity between original word in the query and the word to be substituted. But the precise measurement of the semantic similarity is a challenging task as most of words in a natural language have context dependent meaning. A search engine, if it does not take into consideration the context of query and directly uses the online sources, may fetch large number of undesired pages which are otherwise not required. Keeping this need in view, a new approach is being presented in this paper which tunes the lexical resources wherein the derived synonyms of a word are also provided with their applicable context. The work will help the automated systems like search engines in extending the query through appropriate choice of context based synonyms. Moreover, the extent of semantic similarity has also been defined to help the automated tools in choosing the appropriate list of synonyms in the order of semantic similarity index till the requisite numbers of pages are provided to the user. Thus, the goal of the present work is synonym resolution wherein, for a given word, a lexical resource would be enriched with:

- Its applicable contexts set i.e. set of contexts to which the word is applicable;
- A set of synonyms of the words along with intersection context set i.e. the set of contexts which is applicable to both (the word and the corresponding synonyms);
- A similarity index indicating the extent of similarity which will help the automated tool in making the decision about the order of substitution. Larger the index, more is the priority for substitution. A cut off value may also be decided beyond which the substitution is not desirable.

To enable the automated tool in making the intelligent decision on the basis of the additional information created in the work, the similarity index has been fuzzified through the design of appropriate fuzzy sets. This will help the information retrieval system in following manner:

- Minor differences arising out of computation due to the choice of search engine, set of documents, dialects etc. are eliminated as two close values would normally be residing in the same fuzzy set;
- The fuzzy sets created in such a manner can be used for intelligent decision making by creating a Fuzzy Rule Base that can be run on an appropriate fuzzy inference engine.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Now, to accomplish the proposed work, the following Algorithm 1 has been designed.

Algorithm 1. Fuzzy_rule_base (w, n sets of corpus)

Input: The word w whose rule base is to be created, n sets of word corpus
Output: The fuzzy rule base of the input word

Step 1: *Choose a candidate word for finding semantic similarity, say w*
Step 2: *Choose a set of n corpora.*
Step 3: *For each corpus i, identify set of contexts for w,* C_i
Step 4: *Compute* $C_{pw} = C_1 \cup C_2 \cup \ldots \cup C_n$ *// Possible list of contexts for w*
Step 5: *Compute* $C_{cw} = C_1 \cup C_2 \cup \ldots \cup C_n$ *// List of most commonly used contexts for w*
Step 6: *Augment* C_{pw} *and* C_{cw} *in lexicon alongwith the word w.*
Step 7: *Let S be the set of synonyms for w obtained from a lexical resource, say WordNet.*
Step 8: *for each x∈S, Compute Cpx and Ccx using step 3-5 //List of possible and commonly used contexts for x respectively*
Step 9: *Compute similarity index between w and x using WebJaccard, WebOverlap and WebDice type co-occurrence measure. Here, the context set cardinality is used instead of page count.*

Modified_WebJaccard (w, x)= $|C_{pw} \cap C_{px}| / |(C_{pw} + C_{px} - C_{pw} \cap C_{px})|$
Modified_WebOverlap(w,x)= $C_{pw} \cap C_{px} / min(C_{pw}, C_{px})$
Modified_WebDice(w,x)= $2*C_{pw} \cap C_{px} / (C_{pw} + C_{px})$

Step10: *Normalize the computed co-occurrence to a range* $[r_1, r_2]$.
Step11: *Use* $[r_1, r_2]$ *as domain of discourse to create fuzzy sets expressing the extent of similarity between the words w and x. Let the fuzzy sets be: not_similar, poorly_similar, medium_similar, highly_similar, extremely_similar.*
Step12: *Create an appropriate Fuzzy Rule Base (FRB) to implement the computed similarity in automated manner.*
Step 13: *Use FRB for query expansion and information retrieval.*
Step 14: *Stop*

Explanation and Implementation of Proposed Algorithm

To identify the contexts related to a word, four corpora were chosen namely Coca Corpus (1990), BNC Corpus (1980), Wikipedia Corpus(2006) and GloWbE Corpus(2005). The reason for choosing the above corpora is their ability to allow search on the basis of word, phrase, part of speech and synonyms.

The details of these corpora have been provided in Table 1.

The proposed algorithm was implemented using a very commonly used word beautiful. The reason for choosing the word beautiful is because of its commonness and availability of its large number of synonyms applicable to different varieties of contexts. Its available synonyms were found using the lexical resource *WordNet* thereby creating the set *S*. For the purpose of context identification, the size of text window was taken as 2 (two) because a shorter window ensures that proper relevance of the context. Moreover, since the word is of adjective type, this window size is quite appropriate. For other types of words, the size of the window can be changed. The computation of possible set of contexts C_{pw} and commonly used contexts C_{cw} using multiple corpora ensures the elimination of bias that may occur with a particular corpus. To apply the algorithm, few synonyms of beautiful were taken namely: Pretty, Lovely, Gorgeous, Glorious, Good_looking and Stunning.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Details of Corpora used in this work

Corpus	Wide Range of Genres (e.g.)	No. of Words	Type of Documents
Coca Corpus	spoken, fiction, popular magazines, newspapers, and academic texts.	520 million words	Text
BNC Corpus	spoken, fiction, magazines, newspapers, and academic)	100 million words	Text
Wikipedia Corpus	4.4 million documents related to microbiology, economics, basketball, Buddhism, or thousands of other topics.	1.9 billion	Text
GloWbE Corpus	Any type of data related to newspaper, magazines and academic.	1.9 billion	Text

The extracted list of synonyms is checked for similarity with the input adjective word. The different corpora (four) are used to identify the contexts related to a given word and its synonyms. Then list of all possible contexts are taken out by considering the union of contexts taken from all the four corpora. The list of most commonly used contexts is taken out by considering the intersection of contexts using these corpora. In the literature, WebJaccard, WebOverlap and WebDice are different types of standard indices used to compute similarity index between input word and its synonymous on the basis of page-count. All these indices have been modified in our algorithm to use the context set cardinality instead of page count. It is for this reason that they have been named as Modified_WebJaccard, Modified_WebOverlap and Modified_WebDice respectively.

The computed index results have been normalized to the range [0,100] through the use of a normalization factor. The purpose of normalization is to create a relative standing between the various candidate synonyms through a standard level of parity like percentage. The so computed normalized index can also be used as the membership indicator in the corresponding fuzzy set. The range [0,100] is then used as a domain of discourse to create fuzzy sets expressing the extent of similarity between two words through linguistic expressions. The fuzzy set framework created in such a manner leads to the creation of a Fuzzy Rule Base (FRB) that can be used for automated query expansion and information retrieval through the usage of computed normalized similarity index.

The major issue in this context was the calculation of normalization factor. The empirical calculation of this factor may slightly vary with the perception of the designer. In our case, it is considered that two different words are almost similar (assuming that they cannot be exactly similar) at the level of 100 if their context set is at least 80% same as that of the smaller context set and sizes of their context sets don't differ by 25%. Of course, the designer of the automated system can choose different values as suited to their application and the precision requirement. To illustrate the normalization process, let us take two words w_1 and w_2 with cardinality of their context sets $|C_{w1}|$ and $|C_{w2}|$:

Let $|C_{w1}| = 100$, $|C_{w2}| = 125$ // Difference in cardinality of context set being 25%

$|C_{w1} \cap C_{w2}| = 80$ // Context set matching is 80% of the smaller context set

$|C_{w1} \cup C_{w2}| = |C_{w1}| + |C_{w2}| - |C_{w1} \cap C_{w2}| = 145$

Expected normalized similarity value=100

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

This data leads to values of different indices as follows:

Modified_ WebJaccard= 0.55

Modified_ WebOverlap as 80/100=0.80

Modified_ WebDice as 160/225=0.71

The normalization factor under the circumstances for these indices will be 100/0.55=181.81 (182), 100/0.80=125 and 100/0.71= 140.84(141) respectively subject to maximum of 100. The so computed normalization factor has been used in the Tables 3-9 of this paper to computer the normalized similarity index.

Given below are possible and commonly used context sets for the example word 'beautiful' and its synonyms.

- **List of possible contexts for word Beautiful C_{pw}** { area, art, baby, bay, beach, black, blonde, blue, body, book, boy, bride, building, children, church, city, cloth, color, country, countryside, creature, dark, daughter, day, dreamer, dress, evening, eyes, face, family, female, fish, flower, game, garden, girl, hair, head, home, house, image, island, job, lady, lake, landscape, liar, life, little, maiden, man, mind, morning, mother, mountain, movement, music, name, natural, new, night, object, old, park, part, person, photo, picture, piece, place, post, princess, red, river, room, scenery, setting, sight, singing, skin, smile, song, soul, sound, south, spring, story, stranger, summer, surrounding, temple, thing, town, valley, view, village, voice, way, weather, wedding, white, wife, woman, word, work, world, young}=107
- **List of Commonly used contexts for word beautiful C_{cw}**{baby, beach, city, country, daughter, day, face, flower, garden, girl, lady, person, place, scenery, thing, view, voice, wife, woman, young} = 20
- **List of Commonly used contexts for word Pretty C_{CX1}** {girl, woman, face, picture}=4
- **List of possible contexts for word Pretty C_{PX1}**{ baby, blonde, boy, busy, child, cloth, committee, dress, eyes, face, garden, girl, guardian, horse, lady, light, maid, picture, pink, poison, sight, soldier, solid, song, sweet, thing, village, visitor, woman}=29
- **List of Commonly used contexts for word Lovely C_{CX2}**{bone, children, city, couple, daughter, day, eyes, face, family, garden, girl, guy, lady, man, night, person, place, sight, song, story, thing, voice, wife, woman}=24
- **List of possible contexts for word Lovely C_{PX2}**{afternoon, angel, area, article, ballad, bank, blog, blue, body, bone, book, boy, bride, bunch, butcher, carrot, children, city, cloth, color, colour, country, countryside, county, couple, daughter, day, dinner, dog, dream, dress, egg, evening, eyes, face, family, feather, feeling, flower, food, friend, garden, gift, girl, green, guy, hair, head, holiday, horse, hotel, house, idea, idol, image, kid, lady, land, lane, light, little, lock, lunch, man, meal, moment, music, name, night, old, party, person, photo, picture, piece, place, princess, reader, red, rita, room, rose, seat, setting, shade, sight, smell, smile, song, sound, spot, story, stuff, summer, surprise, tea, thing, thought, time, town, valley, view, voice, war, way, weather, white, wife, woman, word, world, young}=112

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- **List of Commonly used contexts for word Magnificent C_{CX3}**{achievement, bird, building, city, collection, display, example, garden, house, job, old, performance, piece, scenery, sight, view, work}=17
- **List of possible contexts for word Magnificent C_{PX3}**{ achievement, amberson, animal, architecture, art, away, backdrop, baroque, bastard, beach, beast, beauty, bird, black, blue, body, book, bridge, bronze, building, but, butcher, career, castle, cathedral, century, church, city, cliff, collection, contribution, coral, country, countryside, courage, court, creation, creature, day, desolation, dining, display, dog, dress, effort, entrance, event, example, experience, figure, five, four-poster, fraud, frigatebird, funeral, game, garden, gift, goal, golden, gothic, grey, hall, head, history, home, horse, hotel, house, job, journey, lake, landscape, library, life, man, mansion, marble, medieval, mile, montague, monument, moody, mountain, muraco, natural, new, obsession, old, opportunity, palace, panorama, park, performance, physique, piece, place, player, record, red, renaissance, response, ring, room, scale, scenery, season, second, set, setting, seven, show, sight, site, six, sound, specimen, spectacle, stained, state, stone, structure, sunset, surrounding, temple, thing, tomb, tree, victory, view, villa, vista, voice, way, white, woman, wooden, work, world, yankee, young}=141
- **List of Commonly used contexts for word Good_Looking C_{CX4}**{ boy, couple, dog, face, fellow, female, girl, guy, man, person, woman}=11
- **List of possible contexts for word Good_Looking C_{PX4}**{ actor, african-american, american, animal, babe, baby, bastard, bird, black, blog, bloke, blonde, boat, body, boss, boy, breed, british, broad, brother, brunette, cabin, car, castored, chap, character, chestnut, chick, child, children, cloth, club, college, color, cook, cookware, corpse, couple, crowd, curve, daddy, danish, date, design, desk, detective, diagram, dish, doctor, document, dog, dreamer, drew, duck, dude, duke, european, face, family, feature, fellow, female, field, fighter, film, folk, football, foreigner, frame, friend, game, gentleman, girl, glass, graphics, group, guard, gun, guy, hair, handset, head, horse, husband, image, kid, lady, living, male, man, manuscript, mary, movie, musician, officer, one, pair, patient, person, phone, photo, pig, playboy, product, professional, series, shot, stranger, sunshine, surgery, teacher, teenager, victory, view, voice, waiter, weather, white, woman, year, young, youth}=122
- **List of Commonly used contexts for word Glorious C_{CX5}** { battle, career, celebration, city, day, death, era, food, future, history, land, life, light, moment, night, revolution, summer, sun, thing, tradition, victory, voice, year}=23
- **List of possible contexts for word Glorious $C_{PX5\{}$** abandon, achievement, adventure, afternoon, age, amateur, appearing, army, ascension, battle, beach, bloom, blue, body, book, burden, career, cause, celebration, century, chance, chapter, church, city, climax, color, colour, company, country, countryside, day, dead, death, deed, display, empire, end, era, experience, eyes, fire, first, food, freedom, future, garden, goal, god, gospel, green, hair, heritage, high, history, human, king, land, leader, life, light, lord, love, memory, mess, minute, mission, moment, month, morning, mother, mud, music, mystery, name, nation, night, noise, opportunity, order, past, path, peace, period, place, player, power, prospect, red, return, revolution, ring, role, run, scenery, season, sense, sight, song, sound, spring, success, summer, sun, sunrise, sunset, sunshine, thing, time, tradition, victory, view, virgin, voice, war, way, weather, week, work, world, year, youth}=121
- **List of Commonly used contexts for word Stunning C_{CX6}**{ array, beauty, collection, landscape, performance, photograph, piece, scenery}=8

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- **List of possible contexts for word Stunning C_{PX6}**{ about, accomplishment, ace, achievement, admission, album, amount, announcement, arcade, architecture, array, art, artwork, attack, beach, beauty, black, blonde, blue, book, cast, central, claim, clarity, coast, collapse, collection, color, conclusion, contrast, costume, countryside, creation, dark, day, debut, decision, defeat, design, detail, development, discovery, display, dot, double, dress, effect, election, end, evening, example, exhibition, fact, fall, fashion, film, fish, form, garden, girl, goal, good, graphics, guitar, hat, home, illusion, image, impact, lake, landscape, lap, light, line, location, look, loss, military, moment, mountain, move, natural, news, number, opening, panoramic, performance, photo, photograph, picture, piece, place, portrait, presentation, record, result, reversal, scenery, sea, season, series, set, sight, solo, souvenir, speed, start, statement, story, strike, success, surrounding, thing, turn, variety, victory, video, view, visual, voice, volley, white, woman, work, young}=125

RESULTS

Table 2 represents the computation of various indices based upon the extracted information. Results have been normalized using the normalization factors described earlier in this paper. The computed results have been compared with *UMBC* toolkit (2004), [A Similarity ToolKit] in Tables 3 to Table 9. The toolkit uses statistical method that is based on Latent Semantic Analysis (LSA) and distributional similarity. The whole process is automated and can be trained using different corpora. The technique used in this toolkit assumes that the semantics of a phrase is compositional on its component words. The concept and relation similarity can be found for noun, verb, adjective and adverb using either refined Stanford Webbase corpus (2007) or LDC English Gigaword corpus (2003).

The computed results show quite an agreement with the toolkit and the errors due to various factors owing to the usual ambiguity of the natural language justify the need for fuzzification. The domain of discourse for the fuzzy sets is [0,100] and the chosen fuzzy sets are:

NS: Not_Similar
PrS: Poorly_Similar
SS: Somewhat_Similar
QS: Quite_Similar
PfS: Perfectly Similar

Now, one can choose an appropriate index and use it for fuzzification. Our preferred choice is Modified WebJaccard as it into account both union and intersection aspect. The membership graph of various fuzzy sets is shown in Figure 1.

The design of fuzzy sets follows the fuzzy rule base (FRB) needed for automated decision making. The design of fuzzy sets and the FRB is illustrative and designer of the system can chose modify them as per the requirement.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Computation of different Indices

	Corpus	C_{pw}	C_{px}	$C_{pw} \cap C_{px}$	$C_{pw} \cup C_{px}$	Modified WebJaccard (*w,x*)	Modified WebOverlap (*w,x*)	Modified WebDice (*w,x*)	UMBC Toolkit(%age)
Beautiful									
Pretty	*COCA* +BNC+ *WIKI*+ GloWbE	107	29	16	120	0.133	0.552	0.235	40
Lovely			112	60	159	0.377	0.560	0.548	80
Magnificent			141	47	202	0.232	0.439	0.377	60
Glorious			121	34	194	0.175	0.318	0.298	30
Good_ Looking			122	28	201	0.139	0.262	0.245	Not Available
Stunning			125	35	197	0.177	0.327	0.302	50
Pretty									
Beautiful	*COCA* +BNC +*WIKI* + GloWbE	29	107	16	120	0.133	0.552	0.235	40
Lovely			112	15	126	0.119	0.517	0.213	50
Magnificent			141	6	165	0.036	0.207	0.070	Not Available
Good_ Looking			122	10	141	0.071	0.345	0.132	30
Glorious			121	6	144	0.042	0.207	0.080	Not Available
Stunning			125	9	145	0.062	0.310	0.117	Not Available
Lovely									
Beautiful	*COCA* +BNC+*WIKI*+ GloWbE	112	107	60	159	0.377	0.560	0.548	80
Pretty			29	15	126	0.119	0.517	0.213	50
Magnificent			141	33	221	0.149	0.295	0.260	50
Good_ Looking			122	28	206	0.136	0.250	0.239	Not Available
Glorious			121	35	198	0.177	0.312	0.300	40
Stunning			125	24	213	0.113	0.214	0.203	40
Magnificent									
Beautiful	*COCA* +BNC+*WIKI*+ GloWbE	141	107	47	202	0.233	0.440	0.377	60
Pretty			29	6	165	0.036	0.207	0.070	50
Lovely			112	33	221	0.149	0.295	0.260	50
Good_ Looking			122	23	241	0.095	0.188	0.174	Not Available
Glorious			121	35	228	0.153	0.290	0.266	50
Stunning			125	39	228	0.171	0.312	0.292	60
Good_Looking									
Beautiful	*COCA* +BNC+*WIKI*+ GloWbE	122	107	28	201	0.139	0.262	0.245	Not Available
Pretty			29	10	141	0.071	0.345	0.132	Not Available
Lovely			112	28	206	0.136	0.250	0.239	Not Available
Magnificent			141	23	241	0.095	0.188	0.174	Not Available
Stunning			121	10	233	0.043	0.083	0.082	Not Available

continues on following page

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Continued

	Corpus	C_{pw}	C_{px}	$C_{pw} \cap C_{px}$	$C_{pw} \cup C_{px}$	Modified WebJaccard (*w,x*)	Modified WebOverlap (*w,x*)	Modified WebDice (*w,x*)	UMBC Toolkit(%age)
Glorious									
Beautiful	*COCA* +BNC+*WIKI*+ GloWbE	121	107	34	194	0.175	0.318	0.302	30
Pretty			29	6	144	0.042	0.207	0.080	Not Available
Lovely			112	35	198	0.177	0.312	0.300	40
Magnificent			141	35	228	0.153	0.290	0.266	50
Good_ Looking			122	10	233	0.043	0.083	0.082	Not Available
Stunning			125	23	223	0.103	0.190	0.187	30
Stunning									
Beautiful	*COCA* +BNC+*WIKI*+ GloWbE	125	107	35	197	0.177	0.327	0.302	50
Pretty			29	9	145	0.062	0.310	0.117	Not Available
Lovely			112	24	213	0.113	0.214	0.203	40
Magnificent			141	39	228	0.171	0.312	0.292	60
Good_ Looking			122	16	231	0.069	0.131	0.129	Not Available
Glorious			121	23	227	0.103	0.190	0.187	30

Table 3. Comparison with UMBC Toolkit for 'beautiful'

Word Chosen: Beautiful (w)				
Synonym(x)	Modified WebJaccard (*w,x*)	Modified WebOverlap (*w,x*)	Modified WebDice (*w,x*)	Toolkit (Value in %age)
Pretty	24.21	69	33.13	40
Lovely	68.61	70	77.27	80
Magnificent	42.22	54.9	53.21	60
Glorious	31.85	39.75	42.02	30
Good_Looking	25.29	32.75	34.54	X*
Stunning	32.21	40.87	42.58	50

*X means not available.

Fuzzy Rule Base

If the similarity index is *perfectly similar* use W2 in place W1 for word sense disambiguation, web page classification and query expansion.
If the similarity index is *quite similar* use W2 in place W1 for query expansion and web page classification.
If the similarity index is *somewhat similar* use W2 in place W1 in query expansion.
If the similarity index is *poorly similar* use W2 in place W1 for query expansion after checking the context.
If the similarity index is *not similar* don't use W2 in place W1.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Comparison with UMBC Toolkit for 'pretty'

Word Chosen: Pretty (w)				
Synonym(x)	**Modified WebJaccard (*w,x*)**	**Modified WebOverlap (*w,x*)**	**Modified WebDice (*w,x*)**	**Toolkit**
Beautiful	24.21	69	33.13	40
Lovely	21.66	64.62	30.03	50
Magnificent	6.55	25.87	9.87	X
Good_Looking	12.92	43.13	9.21	30
Glorious	7.64	25.87	11.28	X
Stunning	11.28	38.75	16.50	X

Table 5. Comparison with UMBC Toolkit for 'lovely'

Word Chosen: Lovely (w)				
Synonym(x)	**Modified WebJaccard (*w,x*)**	**Modified WebOverlap (*w,x*)**	**Modified WebDice (*w,x*)**	**Toolkit**
Beautiful	68.61	70	77.27	80
Pretty	21.66	64.62	30.03	50
Magnificent	27.12	36.87	36.66	50
Good_Looking	24.75	31.25	33.70	X
Glorious	32.21	39	42.30	40
Stunning	20.57	26.75	28.62	40

Table 6. Comparison with UMBC Toolkit for 'magnificent'

Word Chosen: Magnificent (w)				
Synonym(x)	**Modified WebJaccard (*w,x*)**	**Modified WebOverlap (*w,x*)**	**Modified WebDice (*w,x*)**	**Toolkit**
Beautiful	42.41	55	53.20	60
Pretty	6.55	25.87	9.87	50
Lovely	27.12	36.87	36.66	50
Good_Looking	17.29	23.5	24.53	X
Glorious	27.85	36.25	37.51	50
Stunning	31.12	39	41.17	60

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 7. Comparison with UMBC Toolkit for 'good looking'

Word Chosen: Good_looking (w)				
Synonym(x)	**Modified WebJaccard (*w,x*)**	**Modified WebOverlap (*w,x*)**	**Modified WebDice (*w,x*)**	**Toolkit**
Beautiful	25.30	32.75	34.54	X
Pretty	12.92	43.13	18.61	X
Lovely	24.75	31.25	33.70	X
Magnificent	17.29	23.50	24.53	X
Stunning	7.83	10.3)	11.56	X

Table 8. Comparison with UMBC Toolkit for 'glorious'

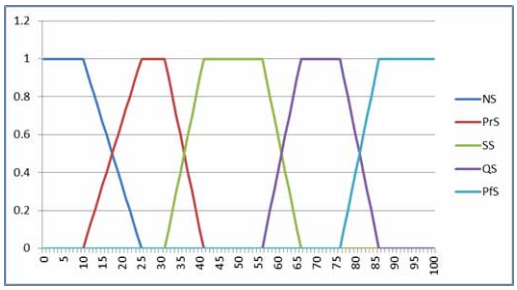
Word Chosen: Glorious (w)				
Synonym(x)	**Modified WebJaccard (*w,x*)**	**Modified WebOverlap (*w,x*)**	**Modified WebDice (*w,x*)**	**Toolkit**
Beautiful	31.85	39.75	42.58	30
Pretty	7.64	25.87	11.28	X
Lovely	32.21	39	42.30	40
Magnificent	27.85	36.25	37.51	50
Good_Looking	7.83	10.37	11.56	X
Stunning	18.75	23.75	26.79	30

Table 9. Comparison with UMBC Toolkit for 'Stunning'

Word Chosen: Stunning(w)				
Synonym(x)	**Modified WebJaccard (*w,x*)**	**Modified WebOverlap (*w,x*)**	**Modified WebDice (*w,x*)**	**Toolkit**
Beautiful	32.21	40.87	42.58	50
Pretty	11.28	38.75	16.54	X
Lovely	20.57	26.75	28.62	40
Magnificent	31.12	39	41.17	60
Good_Looking	12.56	16.37	18.19	X
Glorious	18.75	23.75	26.79	30

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Membership graph of various Fuzzy sets



Format of Next Generation Lexical Resources

In the last, the desired format of next generation lexical resources has been shown.

The lexical resource will look as follows.

Modified Lexical Resource

<chosen Word>: <Set of possible contexts>
<Set of most commonly used contexts>
<Synonym-1>:<similarity index>
< common context set>
<Synonym-2>:<similarity index>
< common context set>
……..
……..
……...
<Synonym-k>:<similarity index>
< common context set>

CONCLUSION AND FUTURE SCOPE

The paper presents an efficient and effective approach for finding semantic similarity between words depending upon their contexts. To ensure the neutrality of the approach, multiple corpora have been taken. The similarity index has been computed on the basis of commonality of the contexts. Various benchmark indices have been used to find the similarity index and the results have been normalized. The obtained results can be used for enrichment of lexical resources. The results obtained have been compared with a standard toolkit for the purpose of authentication.

The proposed work provides a basic framework that can be carried out on various options at different levels such as with the choice of larger number of corpora, making better choices for corpora, changing the benchmark for basic index computation, calculation of normalization factor, choice of fuzzy sets and their domain of discourse, redefining of fuzzy rule base etc. paving the way for better synonym resolution for automated systems and enrichment of lexical resources.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Adhikesavan, K. (2015). An Integrated Approach for Measuring Semantic Similarity between Words and Sentences using Web Search Engine. *The International Arab Journal of Information Technology*, *12*(6).

Bjerva, J., & Östling, R. (2017). ResSim at SemEval-2017 Task 1: Multilingual word representations for semantic textual similarity. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)* (pp. 154-158). 10.18653/v1/S17-2021

Bollegala, D., Matsuo, Y., & Ishizuka, M. (2011). A web search engine-based approach to measure semantic similarity between words. *IEEE Transactions on Knowledge and Data Engineering*, *23*(7), 977–990. doi:10.1109/TKDE.2010.172

Brown Corpus. (1991). documents. Retrieved from http://www.essex.ac.uk/linguistics/external/clmt/w3c/corpus_ling/content/corpora/list/private/brown/brown.html

Chen, H. H., Lin, M. S., & Wei, Y. C. (2006, July). Novel association measures using web search with double checking. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics* (pp. 1009-1016). Association for Computational Linguistics. 10.3115/1220175.1220302

Cilibrasi, R. L., & Vitanyi, P. M. (2007). The google similarity distance. *IEEE Transactions on Knowledge and Data Engineering*, *19*(3), 370–383. doi:10.1109/TKDE.2007.48

Coca Corpus. (1990). Document set. Retrieved from http://corpus.byu.edu/coca/

CorpusB. N. C. (1980). Document set. Retrieved from http://corpus.byu.edu/bnc/

Elekes, Á., Schäler, M., & Böhm, K. (2017). *On the Various Semantics of Similarity in Word Embedding Models*. KIT – University of the State of Baden-Wuerttemberg and National Research Center of the Helmholtz Association. doi:10.1109/JCDL.2017.7991568

Fialho, P., Rodrigues, H. P., Coheur, L., & Quaresma, P. (2017). L2f/inesc-id at semeval-2017 tasks 1 and 2: Lexical and semantic features in word and textual similarity. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)* (pp. 213-219). 10.18653/v1/S17-2032

GloWbE Corpus. (2005). Document set. Retrieved from http://corpus.byu.edu/glowbe/

Hakami, H., Hayashi, K., & Bollegala, D. (2017). An Optimality Proof for the PairDiff operator for Representing Relations between Words. arXiv:1709.06673

Jurgens, D. A., Turney, P. D., Mohammad, S. M., & Holyoak, K. J. (2012, June). Semeval-2012 task 2: Measuring degrees of relational similarity. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics* (pp. 356-364). Association for Computational Linguistics.

LDC. (2003). English Gigaword corpus Retrieved from https://catalog.ldc.upenn.edu/ldc2003t05

Li, Y., Bandar, Z. A., & McLean, D. (2003). An approach for measuring semantic similarity between words using multiple information sources. *IEEE Transactions on Knowledge and Data Engineering*, *15*(4), 871–882. doi:10.1109/TKDE.2003.1209005

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Losif, E., & Potamianos, A. (2010). Unsupervised semantic similarity computation between terms using web documents. *IEEE Transactions on knowledge and data engineering, 22*(11), 1637-1647.

Pilehvar, M. T., Jurgens, D., & Navigli, R. (2013). Align, Disambiguate and Walk: A Unified Approach for Measuring Semantic Similarity. *ACL,* (1), 1341-1351.

Rada, R., Mili, H., Bicknell, E., & Blettner, M. (1989). Development and application of a metric on semantic nets. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*(1), 17–30. doi:10.1109/21.24528

Ranjbar, N., Mashhadirajab, F., & Shamsfard, M. (2017). Mahtab at SemEval-2017 Task 2: Combination of Corpus-based and Knowledge-based Methods to Measure Semantic Word Similarity. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)* (pp. 256-260). 10.18653/v1/S17-2040

Recski, G. A., Iklódi, E., Pajkossy, K. A., & Kornai, A. (2016). *Measuring semantic similarity of words using concept networks*. Association for Computational Linguistics. doi:10.18653/v1/W16-1622

Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxonomy. arXiv: cmp-lg/9511007

Semantic Similarity Toolkit. (2004). Similarity ToolKit. Retrieved from http://swoogle.umbc.edu/SimService/

Severyn, A., Nicosia, M., & Moschitti, A. (2013). Learning Semantic Textual Similarity with Structural Representations. In *Proc. 51st Annu. Meeting Assoc. for Comput. Linguist., ACL (2)*, 714-718.

Specia, L., Jauhar, S. K., & Mihalcea, R. (2012, June). Semeval-2012 task 1: English lexical simplification. In *Proceedings of the First Joint Conference on Lexical and Computational Semantics* (pp. 347-355). Association for Computational Linguistics.

Stanford Webbase corpus (2007). Retrieved from http://ebiquity.umbc.edu/resource/html/id/351

Wikipedia Corpus. (2006). Document set. Retrieved from http://corpus.byu.edu/wiki/

WordNet. (2005). Lexical Resource. Retrieved from http://wordnetweb.princeton.edu/perl/webwn

This research was previously published in the International Journal on Semantic Web and Information Systems (IJSWIS), 14(4); edited by Miltiadis D. Lytras; pages 92-109; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Section 4
Utilization and Applications

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 41
Finding the Semantic Relationship Between Wikipedia Articles Based on a Useful Entry Relationship

Lin-Chih Chen
National Dong Hwa University, Tawian

ABSTRACT

Wikipedia is the largest online Internet encyclopedia, and everyone can create and edit different articles. On the one hand, because it contains huge amounts of articles and there are many different language versions, it often faces synonymous and polysemy problems. On the other hand, since some of the similar Wikipedia articles may have the same topic of discussion, it needs a suitable way to identify effectively the semantic relationships between articles. This paper first uses three well-known semantic analysis models LSA, PLSA, and LDA as evaluation benchmarks. Then, it uses the entry relationship between Wikipedia articles to design its model. According to the experimental results and analysis, its model has high performance and low cost characteristics compared with other models. The advantages of its model are as follows: (1) it is a good model for finding the semantic relationships between Wikipedia articles; (2) it is suitable for dealing with huge amounts of documentation.

1. INTRODUCTION

World Wide Web (hereinafter referred to as the Web) is a distributed information-sharing platform that allows a wide range of users to share their information via the Internet. Starting with Web 1.0, users only need to use the relevant HTML syntax to create a static web page and publish the page to the web server host. Web 1.0 only provides limited one-way interaction between hosts and users (Singh, Bebi, & Gulati, 2011). That is, users cannot share their ideas directly and immediately with other users by using Web 1.0.

DOI: 10.4018/978-1-7998-0951-7.ch041

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

From traditional Web 1.0 to modern Web 2.0, users can use various information and communication platforms to achieve the two-way interaction between users. Until today, there are many well-known Web 2.0 platforms, such as Facebook, Twitter, YouTube, and Wikipedia, which provide users with an environment to comment, review and share ideas with other users. According to Best (2006), Web 2.0's main features compared to Web 1.0 include a rich user experience, user participation, dynamic content, metadata, Web standards, and scalability.

A widely-used Web 2.0 platform is Wikipedia, which is a free access to the collaborative Internet encyclopedia, where everyone can create and edit their articles or entries. The advantages of Wikipedia compared to other traditional encyclopedias include the following: (1) it contains almost all the possible topics in different subjects (Garcia & Ng, 2006), (2) it can respond quickly to any new event (Jokinen & Wilcock, 2012), (3) it offers various different language versions of the encyclopedia for different users (Hale, 2014).

When users search in Wikipedia, they often need to find semantic relationships that may occur between search terms. In general, users often need to consider the synonymy and polysemous relationships between search terms to help them select the most suitable Wikipedia entry. This is important because most users are difficult to distinguish between synonyms and polysemy entries in Wikipedia's 37 million entries (CBS, 2015).

Semantic analysis models are widely used to identify semantic relationships between terms (Egozi, Markovitch, & Gabrilovich, 2011; Ji, Jing, Wang, & Su, 2012; Liu, Zhang, Chang, & Sun, 2011; Lu, Mei, & Zhai, 2011; Lu, Peng, & Ip, 2010). In recent years, some well-known semantic analysis models have emerged, such as Latent Semantic Analysis (LSA) (Landauer, Foltz, & Laham, 1998), Probabilistic LSA (PLSA) (Hofmann, 2001), and Latent Dirichlet Allocation (LDA) (Blei, Ng, & Jordan, 2003) to find possible semantic relationships between terms. However, these models lack the mechanism to find efficiently the semantic relationships between Wikipedia entries. This is also important because similar entries in Wikipedia may have the same discussion topic. Thus, in this paper, we use a new entry relationship to identify hidden semantic relationships that exist between Wikipedia articles.

The rest of this paper is organized as follows. First, in section 2, we briefly review some of the literature related to this paper. Next, in section 3, we introduce all the semantic analysis models used in this paper. Then, in section 4, we analyze and discuss the relevant experimental results. Finally, in section 5, we conclude this paper and discuss future research directions.

2. LITERATURE REVIEW

In this section, we briefly review some of the research literature related to this paper, including Wikipedia applications and semantic analysis models. In this section, we provide two tables for readers to read and compare related literature.

2.1. Wikipedia Applications

Many researchers have tried to use Wikipedia as the main source of research to solve many different information retrieval problems. Table 1 shows some recent studies on Wikipedia applications. Milne and Witten (2013) created a Wikipedia miner toolkit that allows users to combine the semantics of Wikipedia into their applications. The benefit of the toolkit is that it can efficiently classify different elements of

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Wikipedia, such as articles, labels, and links. The classification methods used in the toolkit include articles and labels comparators, labels and links disambiguators, and links detectors. Wu and Weld (2010) proposed an open information extraction system to build training data. They deal with the unbounded number of relationships found on the web by using the heuristic match between the Wikipedia infobox attribute values and the corresponding sentences. Through this training data, the system can find the possible semantic relationships from the natural language text. Lehmann et al. (2015) presented a DBpedia project that provides query and search capabilities to a wide range of communities. The project extracts some important structured data from Wikipedia through the following extractors: mapping-based infobox, raw infobox, a feature, and statistics. Hahn et al. (2010) proposed a Wikipedia search mechanism that enables users to ask complex questions about Wikipedia knowledge. The mechanism first sends the query to DBpedia to produce the facet value for each potential Wikipedia article. It then uses the sparse tree to store the facet value, which is used to search for related questions. Ciglan and Nørvåg (2010) proposed a WikiPop system to detect trends that are popular topics related to the user's interest. The system first uses a Wikipedia link graph and a graph-based recommendation algorithm to identify some related topics. It then uses the page view statistics to filter out the most popular related topics.

2.2. Semantic Analysis Models

Currently, the most famous semantic analysis models are LSA (Landauer et al., 1998), PLSA (Hofmann, 2001) and LDA (Blei et al., 2003), respectively. We compare and analyze these models based on the methods, advantages and disadvantages, and the concept of problem solving, as shown in Table 2.

LSA first uses the Singular Value Decomposition (SVD) to separate the topic matrix from the original term-document matrix. Then, it uses a dimension reduction technique to filter out some noisy topics from the topic matrix. Its advantage is that it can effectively deal with synonyms by using the dimension reduction technique (Dumais, 2004). On the contrary, its disadvantage is that it cannot effectively deal with the problem of polysemy because SVD is a one-to-one mapping technique from a specific term to a specific document (Landauer, McNamara, Dennis, & Kintsch, 2013). Many researchers have used LSA to solve different Information Retrieval (IR) problems. For the evaluation of short free text answers, some researchers (Klein, Kyrilov, & Tokman, 2011; Lintean, Moldovan, Rus, & McNamara, 2010) attempted

Table 1. Summary of Wikipedia applications

Researcher	Application	Approach	Advantage
Milne and Witten (2013)	Wikipedia miner toolkit	Articles and labels comparators, labels and links disambiguators, and links detectors	It can efficiently classify different elements of Wikipedia
Wu and Weld (2010)	Open information extraction system	The heuristic match between the Wikipedia infobox attribute values and the corresponding sentences	It can deal with the unbounded number of relationships
Lehmann et al. (2015)	DBpedia project	Mapping-based infobox, raw infobox, feature, and statistics	It provides query and search capabilities to a wide range of communities
Hahn et al. (2010)	Wikipedia search mechanism	DBpedia & sparse tree	It allows complex questions in Wikipedia
Ciglan and Nørvåg (2010)	WikiPop system	Link graph, recommended algorithm, page view statistics	It can detect popular topics related to the user

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Discussion on different semantic analysis models

Model	Method	Advantage (A) and Disadvantage (D)	Problem Solving
LSA	SVD & Dimension reduction	A: Synonymy D: Polysemy	Evaluate short free text answers, Document summarization, Plagiarism detection in source code
PLSA	EM algorithm	A: Synonymy and Polysemy D: Long estimated time	Location prediction service, sports video summaries, Detect sound events
LDA	Dirichlet distribution & Gibbs algorithm	A: Unsupervised learning model D: Small or normal distribution of data	Automatic classification bug reporting, Search for multimedia content, Image annotation

to use LSA to automatically assess whether the answer was correct. For the summary of user-generated documents, the researchers (Ozsoy, Alpaslan, & Cicekli, 2011; Yeh, Keb, Yang, & Meng, 2005) used LSA to sort all the documents to achieve the purpose of document summarization. For the problem of plagiarism detection in source code, Cosma and Joy (2012) tried to use LSA to determine whether the source code was plagiarized.

PLSA uses the iterative Expectation Maximization (EM) algorithm to estimate the probability values for all cells in the term-document matrix. The EM algorithm first estimates the expected value of all latent topics based on the probabilities of the observed parameters. It then updates the probabilities of the observed parameters by maximizing the objective function. Compared with LSA, its advantage is that it can further deal with the problem of polysemy, because the EM algorithm is a statistical estimation technique that can simultaneously estimate multiple parameters (Fernandez-Beltran & Pla, 2015; Siddiqui, Mishra, & Verma, 2015). On the contrary, its disadvantage is that it takes a long-estimated time, because the EM algorithm is a time-consuming algorithm (Brahmane & Amune, 2014; Hsieh et al., 2015; Li, Li, & Li, 2013). In the past, many researchers have succeeded in using PLSA to solve various IR problems. To improve the location prediction service, McInerney et al. (2012) used PLSA to predict the mobility patterns of some new users. For sports video summaries and searches, Xu et al. (2008) first analyzed the broadcast video to get the relevant text, and then used PLSA to detect possible sporting events. Mesaros et al. (2011) used two stages to detect the sound events of human activity. In the first stage, PLSA finds the relationship between events. In the second stage, PLSA updates the probability of the event based on the history of the event.

LDA first uses the Dirichlet probability distribution to establish the probability distribution of the relevant parameters. Then it uses the Gibbs sampling algorithm to estimate the final probability of the relevant parameters. Its advantage is that it can easily find the relevant term for each topic, because it is a generation model based on unsupervised learning (Blei, 2012). On the contrary, its disadvantage is that it is not suitable for dealing with a small amount or normal distribution of data (Wang & Blei, 2013). In the past, many researchers used LDA to solve different IR problems. To automatically classify many bug reports, Somasundaram and Murphy (2012) used LDA to help users to streamline the process of resolving bugs. To help users search for multimedia content, Krestel et al. (2009) used LDA to recommend related multimedia tags to users. To achieve effective image annotation, Liénou et al. (2010) first searched for all pattern relationships between different image frames. It then applied the LDA to annotate automatically the image according to the pattern relationship.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

3. RESEARCH METHOD

In this section, we briefly introduce our system, as shown in Figure 1. It contains two major stages to produce relevant experimental data. In the Data Preprocessing stage, we first use the Web Spider to crawl the Search Engine Results Page (SERP) from Wikipedia to get the data source for this study. Since SERP is an unstructured document, we then use Perl Compatible Regular Expressions (PCRE) and Natural Language Processing (NLP) to convert unstructured SERP documents into structured documents. Since the input of the semantic analysis model is the term-document matrix, we finally use the Term Frequency Inverse Document Frequency (TFIDF) to convert all structured documents into the term-document matrix. In the Data Calculating stage, we first use the LSA, PLSA, and LDA models as our benchmark. We then develop a Latent Entry Analysis (LEA) model to find the latent semantic relationship between Wikipedia articles.

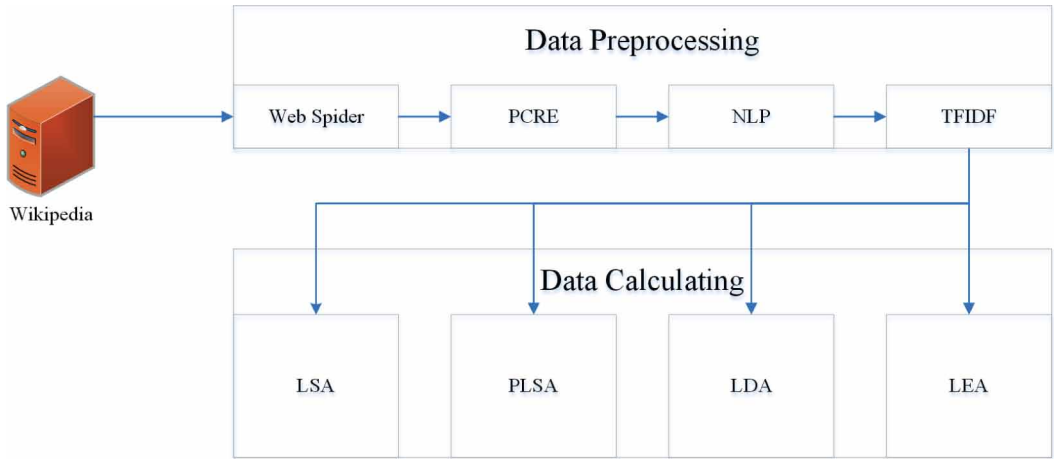
3.1. Data Preprocessing Stage

At this stage, we first developed a Web Spider technique for crawling multiple SERPs from Wikipedia at the same time. In theory, through our technique, the crawling time of multiple SERPs is the same as a single SERP when network latency and bandwidth are not considered.

Since SERP is an unstructured document (Gärtner, Rauber, & Berger, 2014; Naughton, Stokes, & Carthy, 2010), we next use PCRE technique (Hazel, 2017) to do pattern matching for all collected unstructured documents to get important analysis elements such as page title, URL, and description.

Then, we use some NLP techniques to avoid unnecessary characters or words that appear in structured documents. In this study, we use the following NLP techniques: stemming, removing stop words, removing non-word symbols. To avoid the problem that different words have the same root to reduce the performance of IR, we use the snowball stemming algorithm (Porter, 2015) to convert the word to its root. To avoid common and meaningless words in structured documents, we use 671 stop words suggested by Google (2015) to filter out stop words that may degrade IR performance. For the same reason, we also remove all non-word symbols, such as punctuation, special characters, spaces, and HTML tags.

Figure 1. The flow chart of this study



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

We finally use the TFIDF method to convert all structured documents into the term-document matrix, which is the input for different semantic analysis models. The conversion process is shown in Equation 1 and subsequent text:

$$TD\left(d,t\right)=\frac{N\left(d,t\right)}{N\left(d\right)}\log\left(\frac{\left|D\right|}{df\left(t\right)}\right) \tag{1}$$

where $TD(d,t)$ represents the weight that the term t appears in document d; $N(d,t)$ indicates the number of occurrences of t in d; $N(d)$ represents the total number of terms in d; $|D|$ represents the total number of structured documents; $df(t)$ is the number of documents where t appears.

3.2. Data Calculating Stage

At this stage, we first discuss the concepts of LSA, PLSA, and LDA, and then describe our LEA model.

3.2.1. LSA

LSA first uses the SVD technique to decompose the original term-document matrix into three new matrices, as shown in Equation 2 and subsequent text; where S is the topic semantic space, T is the term space corresponding to S, D is the document space corresponding to S, and D^{tr} is the transpose of matrix D:

$$TD(d,t) = TSD^{tr} \tag{2}$$

It then uses the dimension reduction technique to filter out the noise information in S. That is, to obtain the noise-free semantic space S', we preserve the K largest eigenvalues in S and treat the rest as noise. It finally produces the term-document matrix with noise-free topics by the product of the T, S' and D^{tr} matrices.

3.2.2. PLSA

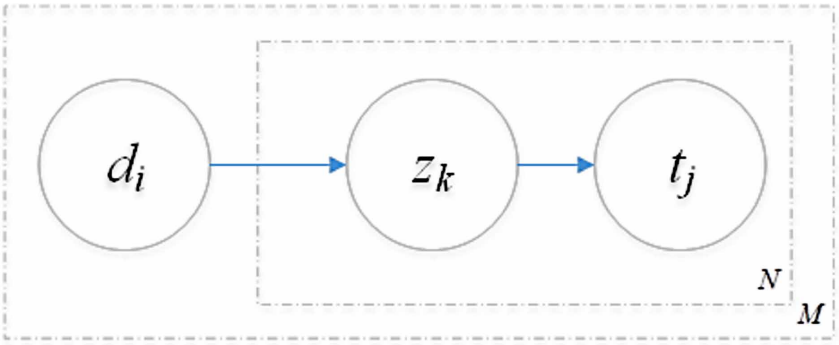
The PLSA uses the aspect model (Hofmann, 1999), which is a latent variable model for estimating unknown variables based on some known variables, to estimate the latent probabilities between terms and documents. The aspect model is shown in Figure 2. In this model, PLSA uses the iterative EM algorithm and the Maximum Likelihood Estimation (MLE) method to estimate the probability that the term appears in the document.

PLSA first selects a specific topic z_k on a given document d_i and then selects a specific term t_j on a given topic z_k. The symbols for PLSA are defined as follows: M is the number of documents, N is the number of terms, $d_i \in \{d_1,\ldots,d_M\}$ is a specific document, $t_j \in \{t_1,\ldots,t_N\}$ is a specific term, $z_k \in \{z_1,\ldots,z_K\}$ represents an observed topic variable.

Based on the aspect model, it uses the following equation to estimate the probability that the term t_j appears in document d_i, $p(d_i,t_j)$, where $p(d_i)$ is the probability of document d_i, $P(X|Y)$ is the conditional probability of X for a given Y occurrence:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. PLSA's aspect model



$$p\left(d_i,t_j\right)=p\left(d_i\right)\sum_k p\left(t_j|z_k\right)p\left(z_k|d_i\right)=\sum_k p\left(t_j|z_k\right)p\left(z_k\right)p(d_i\mid z_k) \tag{3}$$

To estimate the relevant parameters, PLSA uses the iterative EM algorithm and the MLE method to compute the result of $p(d_i,t_j)$. In the MLE method, it uses the following likelihood function as the objective function, where $n(d_i,t_j)$ represents the weight that the term t_j appears in document d_i:

$$L\left(d_i,t_j\right)=\sum_i\sum_j n\left(d_i,t_j\right)\log p\left(d_i,t_j\right) \tag{4}$$

To maximize the objective function, it uses the EM algorithm to estimate the probability of the relevant parameters. In the E step, the algorithm uses the probability of the observed parameter to estimate the expected value of the hidden topic parameter, as shown in the following equation:

$$p\left(z_k|d_i,t_j\right)=\frac{p\left(d_i|z_k\right)p\left(z_k\right)p(t_j\mid z_k)}{\sum_k p\left(d_i|z_k\right)p\left(z_k\right)p(t_j\mid z_k)} \tag{5}$$

In the M step, the algorithm first uses the result of the E step to maximize the value of the objective function and then updates the probability of the observed parameter in Equation 3, as shown in the following three equations:

$$p\left(t_j|z_k\right)=\frac{\sum_j n\left(d_i,t_j\right)p(z_k\mid d_i,t_j)}{\sum_j\sum_i n\left(d_i,t_j\right)p(z_k\mid d_i,t_j)} \tag{6}$$

$$p\left(z_k\right)=\frac{\sum_j\sum_i n\left(d_i,t_j\right)p(z_k\mid d_i,t_j)}{\sum_j\sum_i n\left(d_i,t_j\right)} \tag{7}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$p\left(d_i|z_k\right)=\frac{\sum_i n\left(d_i,t_j\right)p(z_k\mid d_i,t_j)}{\sum_j\sum_i n\left(d_i,t_j\right)p(z_k\mid d_i,t_j)} \tag{8}$$

3.2.3. LDA

The LDA uses an aspect model as shown in Figure 3 to represent the probability distribution between the term and the document. Compared with LSA, its advantage is that it is a probability model with explanatory topics. The relevant parameters for LDA are shown below:

- α denotes the parameters of the Dirichlet distribution prior on the topic-document distributions;
- β denotes the parameters of the Dirichlet distribution prior on the term-topic distributions;
- *K* is the number of topics;
- *M* is the number of documents;
- *N* is the number of terms;
- θ_i denotes the topic distribution of document *i*, which is corresponding to the conjugate distribution of parameter α in the Dirichlet distribution;
- φ_z denotes the terms distribution of topic *z*, which is corresponding to the conjugate distribution of parameter β in the Dirichlet distribution;
- $z_{i,j}$ denotes the topic of document *i* containing term *j*, which is corresponding to the conjugate distribution of parameter θ_i in the multinomial distribution;
- t_i denotes the specific term in document *i*, which is corresponding to the conjugate distribution of parameter φ_z in the multinomial distribution.

The LDA first estimates the conditional probability of the *K*-dimensional Dirichlet variable θ when given the parameter α($\theta_k\geq 0$ and $\alpha_k\geq 0$), as shown in the following equation; where α is the *K*-vector with element α_k, and (*x*) is the Gamma function:

$$p\left(\theta|\alpha\right)=\frac{\left(\sum_k \alpha_k\right)}{\prod_k\left(\alpha_k\right)}\theta_1^{\alpha_1-1}\ldots\theta_K^{\alpha_K-1} \tag{9}$$

It next estimates the joint distribution probability for all observed and unobserved parameters when given the observed parameters α and β, as shown in the following equation:

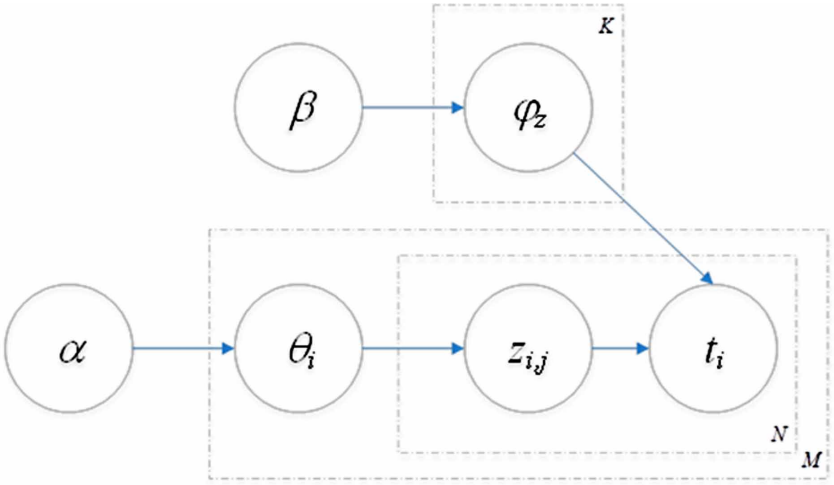
$$p\left(t_i,z_{i,j},\theta_i,\varphi_z|\alpha,\beta\right)=\prod_j p(\theta_i\mid\alpha)p\left(z_{i,j}|\theta_i\right)p\left(t_i|\varphi_z\right)p(\varphi_z\mid\beta) \tag{10}$$

It then estimates the MLE of the term distribution of the document by integrating over θ_i and φ_z, as shown in the following equation:

$$p\left(t_i|\alpha,\beta\right)=\int_{\theta_i}\int_{\varphi_z} p\left(t_i,z_{i,j},\theta_i,\varphi_z|\alpha,\beta\right)d\varphi_z d\theta_i \tag{11}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. LDA's aspect model



It finally uses the Gibbs sampling algorithm (Speh, Muhic, & Rupnik, 2013) to estimate the probability of α and β from the collected documents. Compared with PLSA, its advantage is to reduce significantly the required calculation time, because it uses Gibbs sampling algorithm rather than EM algorithm to estimate the probability of different parameters (Speh et al., 2013).

3.2.4. LEA

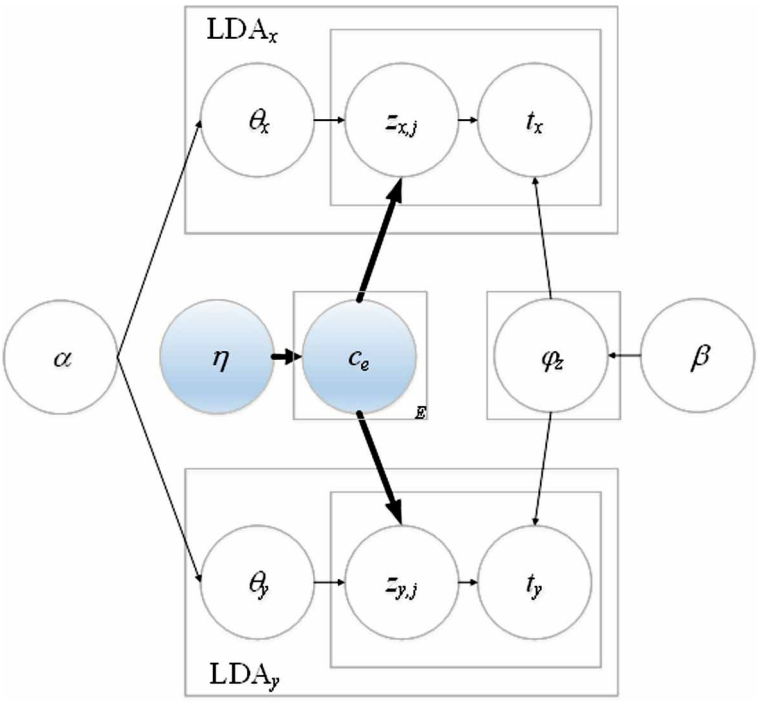
According to section 2.2, LSA can effectively deal with synonyms, but its lack of processing polysemy. PLSA can properly handle synonymous and polysemy problems, but its computation time is time-consuming. LDA can handle synonymous and polysemy problems, and the computation time is relatively fast. However, it is not suitable for dealing with documents of normal distribution, such as Wikipedia, because Wikipedia often has a hidden semantic relationship between different articles. In this paper, we propose the LEA model to present the entry relationship between Wikipedia articles.

LEA first needs to find a model to describe the relationship between terms and topics on a particular document. According to the LDA description, it is a suitable model to estimate the probability of term-topic and topic-document when given a specific document. Therefore, to establish a document link relationship between documents, we need two LDA models, LDA_x and LDA_y, to represent the LDA model of the x and y documents. To connect the two LDA models, in our LEA model, we additionally include an entry relationship parameter η and generate a binary link variable c_e. Figure 4 is the LEA model with the entry relationship between documents x and y. The relevant parameters for LEA are shown below:

- η is the parameters of the Dirichlet distribution prior on the topic-link distributions;
- c_e denotes the topic distribution of entry e, which is corresponding to the conjugate distribution of parameter η in the Dirichlet distribution;
- E denotes the number of entry relationships.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. LEA's aspect model



For any pair of documents, the LEA uses the topic-link relationship η between two documents to estimate how a topic affects the entry relationship. For all collected documents, LEA uses the weighted average of all ηs to estimate how different topics affect the entry relationship.

We now discuss the LEA model's estimation equation in detail. It first estimates the joint distribution probability of all observed and unobserved parameters when given the observed parameters α, β and η, as shown in the following equation:

$$p\left(t_i, z_{i,j}, \theta_i, \varphi_z, c_e | \alpha, \beta,\right) = \prod_j \left(\prod_{i \epsilon \{x,y\}} p(\theta_i \mid \alpha \right) p\left(z_{i,j} | \theta_i\right) p\left(t_i | \varphi_z\right) p\left(z_{i,j} \mid c_e\right) p\left(\varphi_z \mid \beta\right) p\left(c_l \mid\right) \tag{12}$$

It then estimates the MLE of the term distribution of all documents by integrating over θ_i, φ_z and c_e, as shown in the following equation:

$$p\left(t_i | \alpha, \beta,\right) = \int_{\theta_i} \int_{\varphi_z} \int_{c_e} p\left(t_i, z_{i,j}, \theta_i, \varphi_z, c_e | \alpha, \beta,\right) dc_e d\varphi_z d\theta_i \tag{13}$$

Table 3 is a comparison of different semantic analysis models. LSA uses SVD and dimension reduction to filter out noisy topics from raw data. Through dimension reduction, it can effectively identify synonyms between terms. PLSA uses the EM algorithm and the MLE method to estimate the relevant parameters. It can further identify the polysemy between terms because the EM algorithm can simultane-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ously estimate multiple parameters. LDA and LEA use Gibbs sampling algorithm and MLE method to estimate the relevant parameters. The difference between these two models is that they have a different focus. In LDA, it focuses on the topic between terms because of its two important parameters, α (topic-document) and β (term-topic), all focused on topic processing. In LEA, it focuses on how to identify effectively the entry relationship between documents. Through our model, we can easily cluster similar documents with the same discussion topic into a cluster. For speed considerations, LSA is the fastest because it does not use any iterative method to estimate the relevant parameters. In contrast, PLSA is the slowest because the calculation time of the EM algorithm is exponentially increasing as the number of terms and documents increases (Chen, 2012). LDA and LEA are significantly faster than PLSA because the two models use the Gibbs sampling algorithm instead of the EM algorithm to estimate the relevant parameters. In the next section, we will perform some experiments to compare the performance and cost of different semantic analysis models.

4. EXPERIMENTAL ANALYSIS AND RESULTS

In this experiment, the experimental data was selected from the top 100 visited articles in Wikipedia from 2012 to 2014. By looking at all the entries, we only keep one when there are duplicate entries in different years. The number of views for each entry is the sum of the number of visits in different years. Table 4 is the most popular 20 entries on Wikipedia. In this study, we selected 206 entries as experimental data for this study. All experimental data are shown at http://goo.gl/wn5aHZ.

For the Data Preprocessing stage in Figure 1, we use the PHP scripting language to produce the term-document matrix from Wikipedia. For the Data Calculating stage, we use the MATLAB programming language to simulate different models. To compare the performance of different models, we need some

Table 3. Comparison of different semantic analysis models

Model	Method	Speed	Focus
	Estimation Equation		
LSA	SVD & dimension reduction	Fastest	Synonyms between terms
	$TD(d,t) = TSD^{tr}$		
PLSA	EM & MLE	Slowest	Polysemy between terms
	$p(d_i,t_j) = \sum_k p(t_j \mid z_k)p(z_k)p(d_i \mid z_k)$		
LDA	Gibbs & MLE	Fast	Topics between terms
	$p(t_i,z_{i,j},\theta_i,\phi_z \mid \alpha,\beta) = \prod_j p(\theta_i \mid \alpha)p(z_{i,j} \mid \theta_i)p(t_i \mid \phi_z)p(\phi_z \mid \beta)$ $p(t_i \mid \alpha,\beta) = \int_{\theta_i}\int_{\phi_z} p(t_i,z_{i,j},\theta_i,\phi_z \mid \alpha,\beta)d\phi_z d\theta_i$		
LEA	Gibbs & MLE	Fast	Entry relationship between documents
	$p(t_i,z_{i,j},\theta_i,\phi_z,c_e \mid \alpha,\beta,\eta) = \prod_j(\prod_{i\in\{x,y\}} p(\theta_i \mid \alpha)p(z_{i,j} \mid \theta_i)p(t_i \mid \phi_z)p(z_{i,j} \mid c_e))p(\phi_z \mid \beta)p(c_l \mid \eta)$ $p(t_i \mid \alpha,\beta,\eta) = \int_{\theta_i}\int_{\phi_z}\int_{c_e} p(t_i,z_{i,j},\theta_i,\phi_z,c_e \mid \alpha,\beta,\eta)dc_e d\phi_z d\theta_i$		

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. The list of the 20 most popular entries in English Wikipedia from 2012 to 2014

Entry	Year of Appearance	Number of Views
Facebook	2012, 2013, 2014	83295511
Wiki	2012, 2013, 2014	56352325
United States	2012, 2013, 2014	45700980
YouTube	2012, 2013, 2014	44482179
Java	2013, 2014	43172982
The Walking Dead (TV series)	2012, 2013, 2014	39762769
Wikipedia	2012, 2013, 2014	36890189
The Big Bang Theory	2012, 2013, 2014	36887175
Google	2012, 2013, 2014	35439215
Breaking Bad	2012, 2013, 2014	35360135
World War II	2012, 2013, 2014	35229756
Online shopping	2014	34897548
Climatic Research Unit email controversy	2013, 2014	34025774
Alive	2013, 2014	33257949
How I Met Your Mother	2012, 2013, 2014	32741176
India	2012, 2013, 2014	32627605
One Direction	2012, 2013	32152677
Sex	2012, 2013, 2014	30113283
Game of Thrones	2013, 2014	28717418
Eminem	2012, 2013, 2014	28531162

similarity functions to evaluate the similarity between two terms or document vectors in the matrix. The similarity function can be any vector similarity function. The most common similarity functions are COSine similarity (COS) and CORrelation coefficient (COR) (Hofmann, 2003; Nguyen & Bai, 2011; Tan, Steinbach, & Kumar, 2005). Therefore, in this study, we use these two similarity functions to evaluate the performance of different models.

The number of topics will affect the performance of different models. On the one hand, more topics may lead to degradation of performance; on the other hand, fewer topics may lead to reduced discrimination between documents. Therefore, how to set properly the appropriate number of topics becomes an important job. According to Hoffman et al. (2008), we set the number of topics from 5 to 50 to evaluate the performance of different models.

We first enter the query "mobile phone jb" to Wikipedia to explain carefully synonyms and polysemy how to affect the performance of different models. Table 5 shows an example of COS and COR produced by the results of the different models. According to the table, we found that the terms "mobile phone", "cell phone," and "hand phone" have synonymous features. Similarly, the terms "jb" and "jailbreak" (or "jelly bean") also have synonymous features because "jailbreak" or "jelly bean" are abbreviated "jb". In addition, the term "jb" is also a polysemy because it has at least two different meanings "jailbreak" and "jelly bean". Based on the results of the table, we found that the COS and COR of PLSA, LDA and LEA were significantly better than LSA in terms of the evaluation of terms or document vectors. This means

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

that all models can find the semantic relationships that exist between "mobile phone", "cell phone", and "hand phone". However, LSA lacks the ability to find the semantic relationship between "jb" and the two polysemous terms. That is, all models can handle synonyms properly, but LSA's ability to handle polysemy is poor. The reason is that after SVD processing, the row vector in the matrix can only represent a term, so LSA cannot effectively deal with the polysemy between terms (Ishida & Ohta, 2002). Other models use the aspect model to estimate simultaneously the joint probability of terms and documents. By using the probabilistic approach, it can clearly classify the types and meanings of different terms, so these probabilistic models can properly handle the problem of polysemy.

Next, we perform an extended experiment on a large-scale dataset. For the 206 entries above, we crawl these entries and perform the Data Preprocessing stage to produce the term-document matrix needed for the different models. In this extended experiment, we also use COS and COR to measure the performance of different models. For this data set, we use a computer with Intel Core 2 Duo T9600 and 2GB memory to simulate the performance and computation time needed for different models.

The relevant experimental parameters in this study are described below. First, we use all the Wikipedia articles that match to the selected 206 entries as all the documents in the term-document matrix. Next, we analyze the links in all the documents and use the text matching to all the links as all the terms in the term-document matrix. The number of terms contained in all documents is between about 100

Table 5. An example of COS and COR results for different models

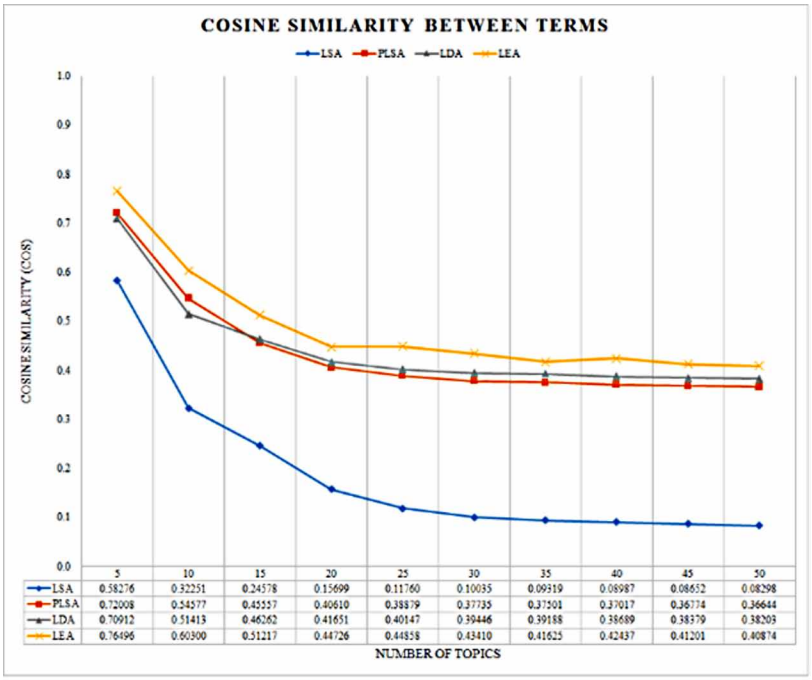
LSA				**PLSA**			
Document Term	**d_1**	**d_2**	**d_3**	**Document Term**	**d_1**	**d_2**	**d_3**
mobile phone	0.16	0.56	0.62	mobile phone	0.13	0.71	0.77
cell phone	0.68	0.12	0.46	cell phone	0.73	0.22	0.23
hand phone	0.13	0.95	0.12	hand phone	0.08	0.88	0.03
jb	0.02	0.12	0.56	jb	0.05	0.62	0.16
jailbreak	0.05	0.82	0.13	jailbreak	0.22	0.42	0.08
jelly bean	0.19	0.97	0.12	jelly bean	0.07	0.87	0.10
	COS	**COR**			**COS**	**COR**	
Between Terms	0.62563	0.02316		Between Terms	0.74445	0.28356	
Between Documents	0.51884	0.65071		Between Documents	0.49691	0.72293	
LDA				**LEA**			
Document Term	**d_1**	**d_2**	**d_3**	**Document Term**	**d_1**	**d_2**	**d_3**
mobile phone	0.58	0.05	0.37	mobile phone	0.50	0.74	0.45
cell phone	0.22	0.77	0.23	cell phone	0.31	0.95	0.21
hand phone	0.95	0.13	0.13	hand phone	0.35	0.43	0.42
jb	0.32	0.02	0.16	jb	0.04	0.16	0.63
jailbreak	0.82	0.05	0.08	jailbreak	0.03	0.16	0.52
jelly bean	0.57	0.08	0.68	jelly bean	0.50	0.62	0.53
	COS	**COR**			**COS**	**COR**	
Between Terms	0.73819	0.25816		Between Terms	0.79399	0.32949	
Between Documents	0.48206	0.72775		Between Documents	0.75181	0.86766	

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

and 200. The reason why there are so many terms in each document is that we selected the popular entries in Wikipedia. Then, we use the TFIDF method to calculate the weight of each cell in the original term-document matrix. Finally, as shown in Figure 1, the original term-document matrix is used for the source of different models and the experimental results are evaluated by calculation of different models.

The result of the different models consists of a term-document matrix with term (row) vectors and document (column) vectors. To verify the effect of different models on terms and documents, we need to calculate the similarity between row and column vectors separately. First, we discuss the effect of different models on terms (row vectors). Figures 5 and 6 show the average similarity between the different term vectors for different models. Each dot in the figures is the average of the similarity between the different term vectors under a given number of topics. According to the results in the figures, COS and COR will decrease as the number of topics increases. The best and worst models are LEA and LSA, respectively. When the number of topics is greater than 5, the cosine similarity of LSA has a significant downward trend, but the trend of other models is not as obvious as it is. This means that LSA is not suitable for handling multiple topics at the same time. According to the results shown in the figures, the performance of PLSA is close to LDA, but it needs much computation time to achieve this solution as described below. This means that PLSA is not suitable for handling huge amounts of documentation. Since the underlying model of the LEA is LDA, the LEA can significantly reduce the computational time as LDA. This means that LEA and LDA are suitable models for handling huge amounts of documentation. Compared with LDA, COS and COR of our model are superior to LDA in Wikipedia article processing. This means that when we use the LDA model in Wikipedia articles and consider the entry relationship between articles, it can further improve performance.

Figure 5. The average similarity between the different term vectors for different models

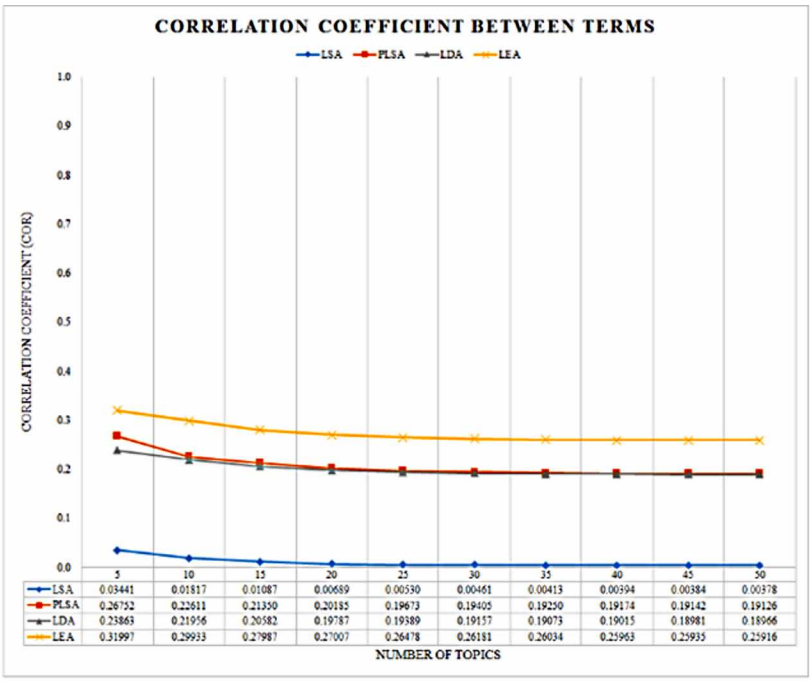


	5	10	15	20	25	30	35	40	45	50
LSA	0.58276	0.32251	0.24578	0.15699	0.11760	0.10035	0.09319	0.08987	0.08652	0.08298
PLSA	0.72008	0.54577	0.45557	0.40610	0.38879	0.37735	0.37501	0.37017	0.36774	0.36644
LDA	0.70912	0.51413	0.46262	0.41651	0.40147	0.39446	0.39188	0.38689	0.38379	0.38203
LEA	0.76496	0.60300	0.51217	0.44726	0.44858	0.43410	0.41625	0.42437	0.41201	0.40874

(a) Cosine Similarity between Terms for different models

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. The average similarity between the different term vectors for different models



	5	10	15	20	25	30	35	40	45	50
LSA	0.03441	0.01817	0.01087	0.00689	0.00530	0.00461	0.00413	0.00394	0.00384	0.00378
PLSA	0.26752	0.22611	0.21350	0.20185	0.19673	0.19405	0.19250	0.19174	0.19142	0.19126
LDA	0.23863	0.21956	0.20582	0.19787	0.19389	0.19157	0.19073	0.19015	0.18981	0.18966
LEA	0.31997	0.29933	0.27987	0.27007	0.26478	0.26181	0.26034	0.25963	0.25935	0.25916

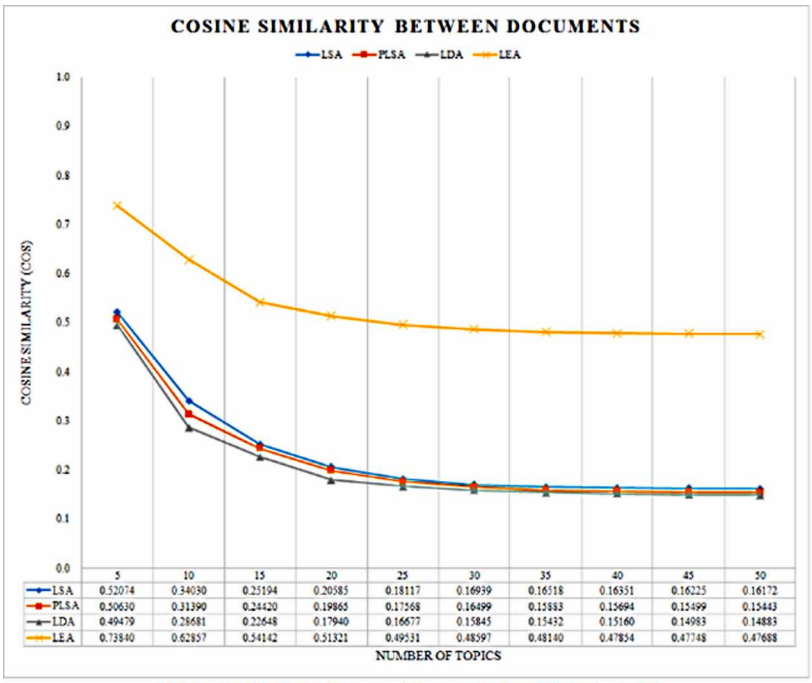
(b) Correlation Coefficient between Terms for different models

Next, we discuss the effect of different models on documents (column vectors). Figures 7 and 8 show the average similarity between the different document vectors for different models. Each dot in the figures is the average of the similarity between the different document vectors under a given number of topics. Based on the results in the figures, we found the average similarity of LSA, PLSA, and LDA has lower performance compared to LEA. This means that the ability of these models to understand document similarity in Wikipedia articles is consistent low. The reason for this is that these models do not consider the entry relationships between documents that often appear in Wikipedia articles. LEA uses the entry relationships between documents to find the semantic relationships that exist between Wikipedia articles. This means that we can easily cluster similar documents with the same discussion topic to the same cluster through the entry relationship. For example, in Table 5, documents d_1, d_2, and d_3 have the same discussion topic ("mobile phone"). With LEA's entry relationship, we can cluster d_1, d_2, and d_3 documents to the same cluster and further improve the average similarity between document vectors.

To understand the source of the differences in experimental data, we then use SPSS 14.0 for Windows to analyze the results of the two experiments. Considering COS and COR of LSA, PLSA, LDA and LEA, we use a statistical method, Analysis of Variance (ANOVA), to show that $F_{(K=5)} = 112.871$, $F_{(K=10)} = 430.771$, $F_{(K=15)} = 555.322$, $F_{(K=20)} = 723.694$, $F_{(K=25)} = 884.042$, $F_{(K=30)} = 903.743$, $F_{(K=35)} = 901.851$, $F_{(K=40)} = 929.021$, $F_{(K=45)} = 928.332$, $F_{(K=50)} = 932.991$ (Table 6) are greater than $F_{0.001}(3,3292) = 5.435$ (F-distribution). This provides a strong evidence to reject the null hypothesis, that is, there is a significant difference in the performance of COS and COR for different models. To compare the differences between the different models, we performed the pairwise comparisons of Fisher's Least Significant

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

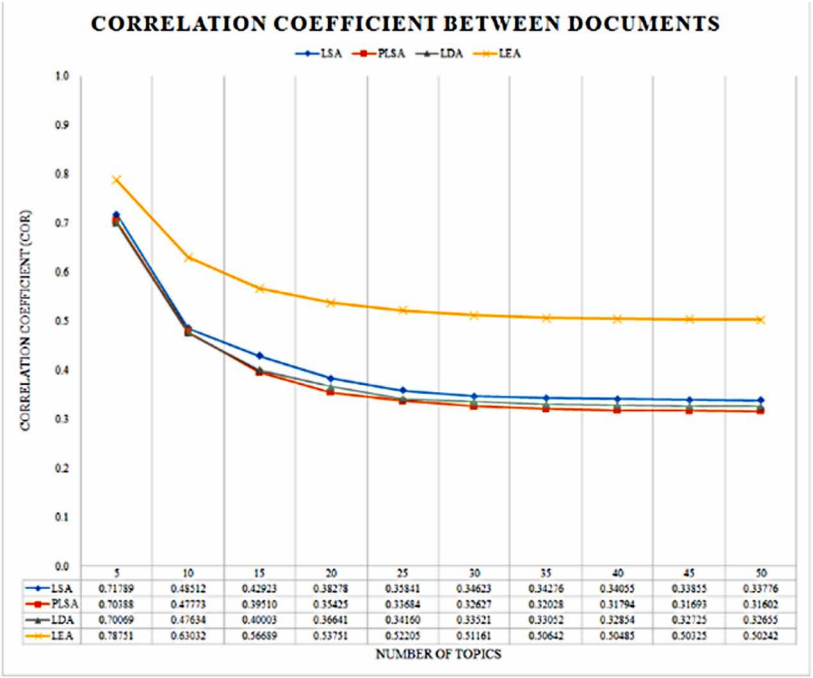
Figure 7. The average similarity between the different document vectors for different models



	5	10	15	20	25	30	35	40	45	50
LSA	0.52074	0.34030	0.25194	0.20585	0.18117	0.16939	0.16518	0.16351	0.16225	0.16172
PLSA	0.50630	0.31390	0.24420	0.19865	0.17568	0.16499	0.15883	0.15694	0.15499	0.15443
LDA	0.49479	0.28681	0.22648	0.17940	0.16677	0.15845	0.15432	0.15160	0.14983	0.14883
LEA	0.73840	0.62857	0.54142	0.51321	0.49531	0.48597	0.48140	0.47854	0.47748	0.47688

(a) Cosine Similarity between Documents for different models

Figure 8. The average similarity between the different document vectors for different models



	5	10	15	20	25	30	35	40	45	50
LSA	0.71789	0.48512	0.42923	0.38278	0.35841	0.34623	0.34276	0.34055	0.33855	0.33776
PLSA	0.70388	0.47773	0.39510	0.35425	0.33684	0.32627	0.32028	0.31794	0.31693	0.31602
LDA	0.70069	0.47634	0.40003	0.36641	0.34160	0.33521	0.33052	0.32854	0.32725	0.32655
LEA	0.78751	0.63032	0.56689	0.53751	0.52205	0.51161	0.50642	0.50485	0.50325	0.50242

(b) Correlation Coefficient between Documents for different models

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Difference (LSD) at 1% significance level. Since the result of the LSD is lengthy and tedious, we put the complete statistical results at http://goo.gl/EGjJfC. As shown in the LSD results, we found that LEA (LSA) is better (worse) than other models.

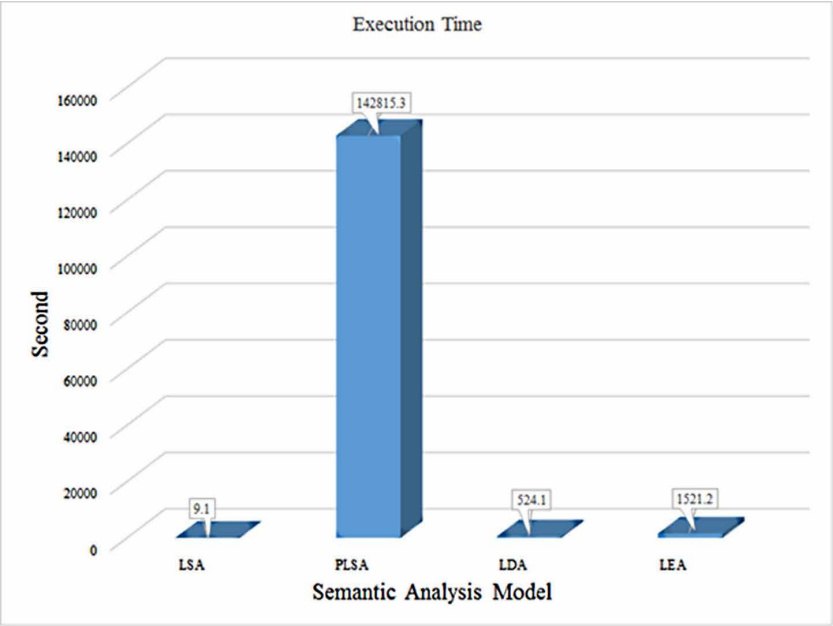
Finally, we analyze the computation time (cost) needed by the different models to verify that PLSA is not suitable for handling huge amounts of documentation. Figure 9 is the computation time required for the different models. Compared with other models, LSA has the fastest calculation time because it does not use any iterative method to estimate the relevant parameters. However, its performance is

Table 6. ANOVA results for different models

		Sum of Squares	df	Mean Square	F	Sig.
K = 5	**Between Groups**	15.025	3	5.008	112.871	.000
	Within Groups	146.070	3292	.044		
	Total	161.095	3295			
K = 10	**Between Groups**	26.561	3	8.854	430.771	.000
	Within Groups	67.661	3292	.021		
	Total	94.222	3295			
K = 15	**Between Groups**	24.790	3	8.263	555.322	.000
	Within Groups	48.985	3292	.015		
	Total	73.774	3295			
K = 20	**Between Groups**	26.917	3	8.972	723.694	.000
	Within Groups	40.814	3292	.012		
	Total	67.731	3295			
K = 25	**Between Groups**	30.056	3	10.019	884.042	.000
	Within Groups	37.308	3292	.011		
	Total	67.364	3295			
K = 30	**Between Groups**	29.383	3	9.794	903.743	.000
	Within Groups	35.677	3292	.011		
	Total	65.059	3295			
K = 35	**Between Groups**	29.377	3	9.792	901.851	.000
	Within Groups	35.745	3292	.011		
	Total	65.122	3295			
K = 40	**Between Groups**	29.669	3	9.890	929.021	.000
	Within Groups	35.044	3292	.011		
	Total	64.713	3295			
K = 45	**Between Groups**	29.323	3	9.774	928.332	.000
	Within Groups	34.662	3292	.011		
	Total	63.985	3295			
K = 50	**Between Groups**	29.428	3	9.809	932.991	.000
	Within Groups	34.611	3292	.011		
	Total	64.039	3295			

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Computation time for different models



the worst because it lacks the ability to handle polysemy. According to the results of the figure, PLSA spent 142815.3 seconds to run the EM algorithm. This means that PLSA needs much computation time to estimate the relevant parameters compared to other models. This is because its computation time exponentially increases as the number of terms and documents increases (Chen, 2012). The advantage of LDA and LEA compared to PLSA is that they can significantly reduce computational time because they use the Gibbs sampling algorithm instead of the EM algorithm to estimate the relevant parameters.

5. CONCLUSION AND FUTURE WORK

In this paper, we use the entry relationship between Wikipedia articles to design our semantic analysis model. The first advantage of our model is that it can effectively identify the semantic relationship between Wikipedia entries. According to the relevant experimental results, the performance of our model is significantly better than other models. The second advantage of our model is that it can significantly reduce the computational time so that our model is suitable for handling huge amounts of documents.

In the future, we plan to add a time series relationship to our model so that we can perform further clustering for documents that have similar update times. Through this relationship, we can clearly distinguish between iPhone 6 and iPhone 6S, because these two types of iPhone phones have different release date.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ACKNOWLEDGMENT

We would like to thank anonymous reviewers of the paper for their constructive comments which help us to improve the paper in several ways. This work was supported in part by Ministry of Science and Technology, Taiwan under Grant Most 105-2221-E-259-030 & 104-2221-E-259-038.

REFERENCES

Best, D. (2006). Web 2.0 Next Big Thing or Next Big Internet Bubble? Retrieved from https://goo.gl/u8mBla

Blei, D. M. (2012). Probabilistic Topic Models. *Communications of the ACM*, *55*(4), 77–84. doi:10.1145/2133806.2133826

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*(1), 993–1022.

Brahmane, A. V., & Amune, A. (2014). A Survey of Dynamic Distributed Network Intrusion Detection using Online Adaboost-Based Parameterized Methods. *International Journal of Innovative Research in Advanced Engineering*, *1*(9), 256–262.

CBS. (2015). Wikipedia cofounder Jimmy Wales on 60 Minutes. Retrieved from https://goo.gl/lGoq7k

Chen, L.-C. (2012). Building a Term Suggestion and Ranking System based on a Probabilistic Analysis Model and a Semantic Analysis Graph. *Decision Support Systems*, *53*(1), 257–266. doi:10.1016/j.dss.2012.02.001

Ciglan, M., & Nørvåg, K. (2010, October 26-30). WikiPop - Personalized Event Detection System Based on Wikipedia Page View Statistics. *Paper presented at the Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, Toronto, Canada.

Cosma, G., & Joy, M. (2012). An Approach to Source-Code Plagiarism Detection and Investigation Using Latent Semantic Analysis. *IEEE Transactions on Computers*, *61*(3), 379–394. doi:10.1109/TC.2011.223

Dumais, S. T. (2004). Latent Semantic Analysis. *Annual Review of Information Science & Technology*, *38*(1), 189–230.

Egozi, O., Markovitch, S., & Gabrilovich, E. (2011). Concept-Based Information Retrieval Using Explicit Semantic Analysis. *ACM Transactions on Information Systems, 29*(2), 8:1-8:34.

Fernandez-Beltran, R., & Pla, F. (2015). Incremental Probabilistic Latent Semantic Analysis for Video Retrieval. *Image and Vision Computing*, *38*(C), 1–12. doi:10.1016/j.imavis.2015.02.003

Garcia, I., & Ng, Y.-K. (2006, November). Eliminating Redundant and Less-Informative RSS News Articles Based on Word Similarity and a Fuzzy Equivalence Relation. *Paper presented at the Proceedings of the 18th IEEE International Conference on Tools with Artificial Intelligence*, Arlington, VA. 10.1109/ICTAI.2006.54

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Gärtner, M., Rauber, A., & Berger, H. (2014). Bridging Structured and Unstructured Data via Hybrid Semantic Search and Interactive Ontology-Enhanced Query Formulation. *Knowledge and Information Systems*, *41*(3), 761–792. doi:10.100710115-013-0678-y

Google. (2015). Stop Words Project. Retrieved from http://code.google.com/p/stop-words/

Hahn, R., Bizer, C., Sahnwaldt, C., Herta, C., Robinson, S., Bürgle, M., . . . Scheel, U. (2010, May 3-5). Faceted Wikipedia Search. *Paper presented at the Proceedings of the 13th International Conference on Business Information Systems*, Berlin, Germany.

Hale, S. A. (2014, June 23-26). Multilinguals and Wikipedia Editing. *Paper presented at the Proceedings of the 2014 ACM Conference on Web Science*, Bloomington.

Hazel, P. (2017). PCRE - Perl Compatible Regular Expressions. Retrieved from http://www.pcre.org/

Hofmann, T. (1999). Probabilistic Latent Semantic Indexing. *Paper presented at the Proceedings of the 22th Annual International SIGIR Conference on Research and Development in Information Retrieval*.

Hofmann, T. (2001). Unsupervised Learning by Probabilistic Latent Semantic Analysis. *Machine Learning*, *42*(1), 177–196. doi:10.1023/A:1007617005950

Hofmann, T. (2003). Collaborative Filtering via Gaussian Probabilistic Latent Semantic Analysis. *Paper presented at the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. 10.1145/860435.860483

Hsieh, J.-W., Chen, L.-C., Chen, S.-Y., Chen, D.-Y., Alghyaline, S., & Chiang, H.-F. (2015). Vehicle Color Classification Under Different Lighting Conditions Through Color Correction. *IEEE Sensors Journal*, *15*(2), 971–983. doi:10.1109/JSEN.2014.2358079

Ishida, K., & Ohta, T. (2002). An Approach for Organizing Knowledge According to Terminology and Representing It Visually. *IEEE Transactions on Systems, Man and Cybernetics. Part C, Applications and Reviews*, *32*(4), 366–373. doi:10.1109/TSMCC.2002.806064

Ji, Z., Jing, P., Wang, J., & Su, Y. (2012). Scene Image Classification with Biased Spatial Block and PLSA. *International Journal of Digital Content Technology and its Applications, 6*(1), 398-404.

Jokinen, K., & Wilcock, G. (2012, May 21-27). Constructive Interaction for Talking about Interesting Topics. *Paper presented at the 8th international conference on Language Resources and Evaluation*, Istanbul, Turkey.

Klein, R., Kyrilov, A., & Tokman, M. (2011). Automated Assessment of Short Free-Text Responses in Computer Science using Latent Semantic Analysis. *Paper presented at the 16th Annual Joint Conference on Innovation and Technology in Computer Science Education*. 10.1145/1999747.1999793

Krestel, R., Fankhauser, P., & Nejdl, W. (2009, October 22-25). Latent Dirichlet Allocation for Tag Recommendation. *Paper presented at the third ACM conference on Recommender Systems*, New York. 10.1145/1639714.1639726

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An Introduction to Latent Semantic Analysis. *Discourse Processes*, *25*(1), 259–284. doi:10.1080/01638539809545028

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Landauer, T. K., McNamara, D. S., Dennis, S., & Kintsch, W. (2013). *Handbook of Latent Semantic Analysis*. London, UK: Psychology Press.

Lehmann, J., Isele, R., Jakob, M., Jentzsch, A., Kontokostas, D., Mendes, P. N., ... Bizer, C. (2015). DBpedia - A Large-scale, Multilingual Knowledge Base Extracted from Wikipedia. *Semantic Web Journal*, *6*(2), 167–195.

Li, M., Li, W. K., & Li, G. (2013). On Mixture Memory Garch Models. *Journal of Time Series Analysis*, *34*(6), 606–624. doi:10.1111/jtsa.12037

Liénou, M., Maître, H., & Datcu, M. (2010). Semantic Annotation of Satellite Images Using Latent Dirichlet Allocation. *IEEE Geoscience and Remote Sensing Letters*, *7*(1), 28–32. doi:10.1109/LGRS.2009.2023536

Lintean, M., Moldovan, C., Rus, V., & McNamara, D. (2010). The Role of Local and Global Weighting in Assessing the Semantic Similarity of Texts Using Latent Semantic Analysis. *Paper presented at the 23th International Florida Artificial Intelligence Research Society Conference*.

Liu, Z., Zhang, Y., Chang, E. Y., & Sun, M. (2011). PLDA+: Parallel Latent Dirichlet Allocation with Data Placement and Pipeline Processing. *ACM Transactions on Intelligent Systems and Technology, 2*(3), 26:21-26:18.

Lu, Y., Mei, Q., & Zhai, C. (2011). Investigating Task Performance of Probabilistic Topic Models: An Empirical Study of PLSA and LDA. *Information Retrieval*, *14*(2), 178–203. doi:10.100710791-010-9141-9

Lu, Z., Peng, Y., & Ip, H. H. S. (2010). Image Categorization via Robust PLSA. *Pattern Recognition Letters*, *31*(1), 36–43. doi:10.1016/j.patrec.2009.09.003

McInerney, J., Rogers, A., & Jennings, N. R. (2012). Improving Location Prediction Services for New Users with Probabilistic Latent Semantic Analysis. *Paper presented at the 2012 ACM Conference on Ubiquitous Computing*. 10.1145/2370216.2370420

Mesaros, A., Heittola, T., & Klapuri, A. (2011, August 29 - September 2). Latent Semantic Analysis in Sound Event Detection. *Paper presented at the 19th European Signal Processing Conference*, Barcelona, Spain.

Milne, D., & Witten, I. H. (2013). An Open-Source Toolkit for Mining Wikipedia. *Artificial Intelligence*, *194*(1), 222–239. doi:10.1016/j.artint.2012.06.007

Naughton, M., Stokes, N., & Carthy, J. (2010). Sentence-Level Event Classification in Unstructured Texts. *Information Retrieval*, *13*(2), 132–156. doi:10.100710791-009-9113-0

Nguyen, H. V., & Bai, L. (2011). Cosine Similarity Metric Learning for Face Verification. In Computer Vision, LNCS (Vol. *6493*, pp. 709–720). doi:10.1007/978-3-642-19309-5_55

Ozsoy, M. G., Alpaslan, F. N., & Cicekli, I. (2011). Text Summarization using Latent Semantic Analysis. *Journal of Information Science*, *37*(4), 405–417. doi:10.1177/0165551511408848

Porter, M. F. (2015). Snowball: A Language for Stemming Algorithms. Retrieved from https://goo.gl/VYLB7

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Siddiqui, A., Mishra, N., & Verma, J. S. (2015). A Survey on Automatic Image Annotation and Retrieval. *International Journal of Computers and Applications*, *118*(20), 27–32. doi:10.5120/20863-3575

Singh, K. P., Bebi, & Gulati, D. (2011). Technological March from Web 1.0 to Web 3.0: A Comparative Study. *Library Herald*, *49*(2), 146–157.

Somasundaram, K., & Murphy, G. C. (2012). Automatic Categorization of Bug Reports using Latent Dirichlet Allocation. *Paper presented at the fifth India Software Engineering Conference*, Kanpur, India. 10.1145/2134254.2134276

Speh, J., Muhic, A., & Rupnik, J. (2013). Parameter Estimation for the Latent Dirichlet Allocation. *Paper presented at the 2013 Conference on Data Mining and Data Warehouses*, Ljubljana, Slovenia.

Tan, P.-N., Steinbach, M., & Kumar, V. (2005). *Introduction to Data Mining*. Boston, Massachusetts: Addison-Wesley Press.

Wang, C., & Blei, D. M. (2013). Variational Inference in Nonconjugate Models. *Journal of Machine Learning Research*, *14*(1), 1005–1031.

Wu, F., & Weld, D. S. (2010, July 11-16). Open Information Extraction using Wikipedia. *Paper presented at the 48th Annual Meeting of the Association for Computational Linguistics*, Uppsala, Sweden.

Xu, C., Zhang, Y.-F., Zhu, G., Rui, Y., Lu, H., & Huang, Q. (2008). Using Webcast Text for Semantic Event Detection in Broadcast Sports Video. *IEEE Transactions on Multimedia*, *10*(7), 1342–1355. doi:10.1109/TMM.2008.2004912

Yeh, J.-Y., Keb, H.-R., Yang, W.-P., & Meng, I.-H. (2005). Text Summarization using a Trainable Summarizer and Latent Semantic Analysis. *Information Processing & Management*, *41*(1), 75–95. doi:10.1016/j.ipm.2004.04.003

This research was previously published in the International Journal of Data Warehousing and Mining (IJDWM),13(4); edited by David Taniar; pages 33-52; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 42
Co-Designing Participatory Tools for a New Age:
A Proposal for Combining Collective and Artificial Intelligences

José Luis Fernández-Martínez
Institute of Advanced Social Studies (IESA-CSIC) & University of Granada, Spain

Maite López-Sánchez
Universitat de Barcelona, Spain

Juan Antonio Rodríguez Aguilar
Artificial Intelligence Research Institute (IIIA-CSIC), Spain

Dionisio Sánchez Rubio
Universitat Politècnica de València (UPV) & Escola d'Art i Superior de Disseny de València EASD, Spain

Berenice Zambrano Nemegyei
MediaLab Prado, Spain

ABSTRACT

In the context of a citizen lab, this article describes how a vanguard of activists, designers, scholars and participation practitioners were involved in a participatory prototyping process. CoGovern was designed as an online participation tool whose focus is to incorporate citizen preferences in local policy making. It is aimed at supporting informed and transparent participatory processes while reducing the ability of sponsoring authorities to "cherry-pick" policy proposals and avoid providing explanations. This article proposes a decision-making process that incorporates artificial intelligence techniques into a collective decision process and whose result is mainly based on standard optimization techniques rather than vote-counting.

DOI: 10.4018/978-1-7998-0951-7.ch042

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. INTRODUCTION

The field of democratic innovations (Fung & Wright, 2001; Smith, 2009) is currently crowded with digital tools aimed at facilitating deliberative and participatory processes. Frequently, these tools emerge within activism environments characterised by a high level of civic engagement and political interest. In particular, most of these environments largely embrace the principles promoted by those social movements in defense of free and open source software such as opensource.org. More recent initiatives also promulgate the coproduction of knowledge between the general public and the scientific community such as Citizen Science (Socientize Consortium, 2013).

This article stems from one of those spaces: MediaLab Prado, a Citizen Lab located in Madrid (Spain). This experimental Lab organized a Workshop for Collective Intelligence on Democracy whose main purpose was to prototype digital tools addressing the challenges posed by citizen participation in the age of social media.

We address some of these challenges by focusing on two main questions: How can digital tools be designed in order to make participatory processes more transparent and accountable? Can artificial intelligence enhance participatory processes in an objective and publicly auditable manner, without deviating from participants' interests?

Firstly, this paper seeks to reflect how a vanguard of activists, designers, scholars and practitioners currently envision what – according to them – will be one of the most common forms of citizen participation in the coming years. Secondly, it proposes the basis for the future development of *CoGovern,* an online citizen participation tool whose focus is to incorporate citizen preferences in local policy-making. Thirdly, our tool proposal aims to ensure informed and transparent participatory processes to avoid that sponsoring authorities cherry-pick policy proposals and omit to provide explanations. And finally, for these purposes, this article proposes a decision-making process which incorporates artificial intelligence (AI) techniques in order to align citizen preferences, budgetary constraints and priorities, or strategic objectives about what policies and investments should be chosen for implementation.

The result has been the initial design of *CoGovern,* a prototype web application in which citizen participation is grounded in information fusion (Torra, V. & Narukawa, Y., 2007), argumentation theories (Awad, E. et al., 2015), and standard optimization techniques. In other words, the main substantive contribution is to propose a decision-making process in which the outputs – normally participatory policy proposals – are selected using optimization techniques. The central argument is that a proposal selection method based on the best possible combination, according to previously agreed and weighted criteria, instead of the traditional vote-counting system, can enhance participatory decision making.

The article proceeds as follows. The next section presents the sociological and technological research background. In particular, it introduces artificial intelligence techniques that can be applied to decision support in the context of political participatory processes. Next, we present the methodological approach which is based on a collective intelligence (CI) experiment; more specifically, a case study about a Citizen Lab is described. Subsequent section presents the resulting prototype design by focusing on its graphical design, functionalities, and automated decision making. Lastly, some implications of these types of participatory tools and future research paths are discussed.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. INTRODUCING AI IN PARTICIPATORY DECISION-MAKING PROCESSES

Political sociology often distinguishes three decision-making models in democratic regimes: representative, participatory, and technocratic (Bengtsson & Christensen, 2014). The major difference between them resides in the actors responsible for the decision making. Thus, in the representative model, a group of elected politicians are in charge of making decisions, whereas that is the responsibility of citizens and experts in the participatory and technocratic models, respectively. Recent public opinion research argues that most citizens prefer decision-making processes that involve both elected politicians and ordinary people; however, people believe that decisions are solely taken by elected politicians (Allen et al. 2015; Font et al. 2015, 2017).

On the contrary, in *Stealth Democracy*, Hibbing and Theiss-Morse (2002) question the existence of a real demand for more participation. They argue that people actually prefer not to participate but decisions to be taken "efficiently, objectively and without commotion and disagreement" (2002:143). Other participation criticisms revolve issues such as the lack of efficiency and the perception that participatory processes are highly time consuming. This article addresses this tension between a desire for more opportunities to participate and a demand of political processes to be more objective and efficient. Overall, our aim is to work towards convergence of diverting positions rather than their confrontation.

Despite this ambivalence results on citizen process preferences, "the participatory turn" has become rooted in political agendas (Bherer et al., 2016). As a result, during the last decades, participatory experiences have widespread around the world putting the stress on different aspects such as: prioritization of the aggregative component (e.g. voting) vs. deliberation; interested groups participation vs. open participation; top-down vs. bottom-up; or consultative vs. "more" binding processes.

However, over the past years, a new generation of participatory institutions is announced, mostly characterized by the use of the new information and communication technologies (ICTs). It is still early to assess whether the spread use of the ICTs entails a entirely new generation of participatory institutions. Nevertheless, what is certain is that, at least, they introduce new demarcation lines when designing participatory processes. For instance, according to Dahlberg's classification (2011), we can find four main categories of e-democracy tools: counter-publics, autonomous Marxists, liberal consumer and deliberative. Whilst the first would seek shaping groups of activists to perform contestatory actions, the second would be oriented towards the networking and self-organization outside the capitalist system (see also Chadwick, 2009:21). On the other hand, those under the category of liberal consumer would be oriented towards making requests and receiving support for them. Some scholars (London, 1995; Bohman, 1998) critically name these spaces as a "Marketplace of ideas" competing between them and introducing game elements (gamification) (Secchi & Spada, 2017). In contrast, those under the category of Deliberative would include argumentation systems to transform particular interests or concerns into common and consensual ones through the rational model of deliberation (Dahlberg, 2011). *CoGovern*, our participatory tool proposal, lies within this last category.

One of the most important changes that these new tools have introduced is the fact of increasing the number and diversity of participants. Although dilemmas on some key issues, such as anonymity and the digital divide, remain unsolved, large participation allows to build streamlined participatory processes as well as to scaling-up. However, this comes at a cost, since it also increases the complexity of the decision-making process. Taking into account this and other challenges facing participatory democracy, this article develops a theoretical decision-making process which incorporates AI techniques. In particular, it proposes a selection method of citizens proposals which determines the best possible combination (of

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

proposals) rather than conducting the traditional vote-counting ranking. But, can AI expected always to enhance participatory processes? We must be prudent about the most suitable application scope of AI when considering democratic processes, since there are many AI techniques, and we cannot foresee all the consequences of introducing general AI in our everyday political life. On the contrary, this article focuses on a specific aspect that can certainly benefit from the incorporation of a particular AI technique: the objective proposals' selection in participatory processes based on optimization techniques.

Nowadays, it is most often the case that authorities sponsor participatory processes to gather citizens preferences about what policies are needed and, once the consultation is finished, they might: (i) just cherry-pick a subset of proposals without giving any reasoned explanation; or (ii) prioritize the proposals in decreasing order of citizen support (normally gathered through voting), and choose the first ones in the ranking[1] so as to implement them as long as the budget lasts. Nevertheless, technology can enrich the whole participatory process. In particular, there are some areas of AI related to information aggregation, argumentation, and optimisation that can be applied. On the one hand, information aggregation (Torra & Narukawa, 2007) and argumentation (Awad et al., 2015) can be applied to opinion gathering. On the other hand, optimisation techniques (Hwang & Masud, 2012) are particularly suited for choosing the best subset of citizen proposals once their characteristics are quantified and selection criteria are clearly stated.

Regarding opinion gathering, voting is the hegemonic mechanism in democratic societies. It is simple and scales-up well. Thus, citizens are familiar with it and it allows the participation of a large number of citizens. Nevertheless, most times decisions are taken in a biased or even uninformed way. Ideally, arguments in favor and against a proposal should be thoughtfully pondered before voting for it. Forums in online participatory processes try to leverage these shortages, but they are unstructured and they just work well for small groups. Argumentation theory (Awad et al., 2015) is a research area concerned with reaching agreements through logical reasoning that consider supporting or/and attacking arguments. Most research work on argumentation is kept at a theoretical level; however, Deliberatorium, an on-line tool for collaborative deliberation (Klein, 2012) constitutes a remarkable exception. Moreover, aggregation operators, such as Weighted Ordered Weighted Aggregation operators (WOWA) (Torra & Narukawa, 2007), constitute mathematical methods that combine information in a richer and more general manner than standard methods such as the average or the mean computation.

As for choosing the best subset of proposals, optimisation methods have been long applied to economic contexts so to maximise the benefit in the selection. Combinatorial auctions (Cramton et al., 2006) for trading bundles of products in markets constitute a representative example. Most commonly, benefit is associated to monetary aspects, but in fact it can be computed with respect to several criteria, as long as they are quantified. The optimisation problems that consider different criteria simultaneously constitute the so-called multi-objective optimisation problems, and linear programming (Hwang & Masud, 2012) is a computational technique designed to solve them (i.e., to maximise the compound benefit). These methods compute the best solution (or, in other words, optimizes an objective function) that satisfies some additional restrictions.

In the subsequent sections of the article, we develop a method through which citizens can rate participatory proposals and also provide arguments in favour or against other citizens' arguments. Furthermore, proposals are selected by means of standard optimisation methods (Hwang & Masud, 2012). Specifically, proposals are chosen for implementation based on a set of criteria: citizen support, budget constraints, and, as we will detail later on, proposal alignment with pre-agreed priorities or strategic objectives.[2] In this manner, decisions are publicly known to be the best ones according to the previously established criteria. To the best of our knowledge, participatory processes do not use any of these methods to select

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

proposals. Our central argument is that these methods enhance transparency and effectiveness in citizen participation because it makes transparent the functioning rules before starting the participatory process. Thus, *CoGovern* prevents arbitrary actions by the authorities such as *cherry-picking* and, on the other hand, selects the best possible combination of proposals, taking into account citizen support, budget constraints, and a proposal's convergence with priorities or strategic objectives.

3. COLLECTIVE INTELLIGENCE AS A METHODOLOGICAL APPROACH: THE CASE OF A CITIZEN LAB

This section presents the methodological approach adopted in this article. For this purpose, it begins by explaining the starting point of this project. Second, it introduces the role played by a Collective Intelligence experiment (an international workshop) in its development.

3.1. Starting Point for the Project

The *CoGovern* project stems from a previous research named *Information fusion for norm consensus in virtual communities* (Serramia, 2016) developed at the University of Barcelona (UB) and the Artificial Intelligence Research Institute (IIIA-CSIC). This project was oriented to facilitate that members from a virtual community were able to decide their own community rules. Briefly, it aims to enable large-scale debates for collective decision making. The main idea is that users can express their support to a presented norm proposal by exposing arguments in favour and against it, as well as scoring (in a 1-5 points scale) each one of the arguments made by other users. By doing so, an argumentative structure (the so-called norm argument map) is naturally created. Aggregation methods are then used to, first, assess the collective support for each argument and, second, compute the collective support for the whole proposed norm. Two pilot tests were performed in order to evaluate the functionality of the norm argument map. Tests encompassed, respectively, 11 (Rodriguez-Aguilar et al., 2016) and 17 (Garcia-Heveling, 2017) people debating on norms within a prototyped football social network. Most users were male university students who reported having participated in virtual communities before. The average age was 25.5 years old. During the tests users debated normally about a few norms (they created and rated up to 76 arguments per norm); afterwards, a satisfaction survey asked them if resulting aggregated ratings were reasonable. In a scale from 1 to 5, the answers' mean was 3.36 respectively, which can be considered as a positive preliminary result if we take into account the usability deficiencies of our prototype.

Considering these modest tests as a starting point, the research team from UB and IIIA decided to look into the field of participatory processes to broaden the application of their research. Thus, in contrast with other projects in alpha or production stages[3], *CoGovern* is in a incipient stage whose goal is to clearly define both the project and the scope of its applications. This stage would correspond to a conceptualization phase encompassing the Envision and Speculate phases defined by Mario Špundak (2014).

3.2. Collective Intelligence at Medialab Prado

The project aforementioned (Serramia, 2016) was presented to the call for "Collective Intelligence for Democracy" by Medialab Prado. This call consisted on a two weeks international workshop aimed at exploring new ideas to design innovative participation tools. Medialab Prado is a Citizen Lab (Mazé, 2014;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hillgren et al., 2011; Seravalli, 2011) where individuals from different backgrounds – scientists, policy makers, lay citizens, programmers or designers – work together to solve complex social problems through participatory prototyping (Brodersen et al., 2008) and co-design. The interaction of these elements is what is known as collective intelligence. Following Chadwick, the core idea is that a distributed network of creators and contributors, the majority of them amateurs, can, using simple online tools, produce information goods that may outperform those produced by so-called authoritative, concentrated sources (2009, p. 21).

This international workshop applied work dynamics that come from the "hackathon" logic. According to Gerard Briscoe (2014), hackathon settings involve people working intensively during a limited period of time. The main goal is to generate ideas rather than develop fully operational tools. Documentation becomes crucial for these events, since it allows to have a registry of all the process which is fundamental to continue the work in the future. Another important factor of this methodology is the relationship established between the participants. *CoGovern* involved two main actors: project coordinators (who applied to the call by Medialab Prado) and external collaborators (who chose the project according to their preferences). The involvement degree between these two roles was balanced in order to guarantee an open participation (Manzini & Coad, 2015; Manzini & Rizzo; Lee, 2008). Finally, *CoGovern* team was composed by a wide range of profiles including architects, political scientists, designers and social workers of different ages and nationalities (Latin America and Spain).[4] This diversity caused the participation to respond to different interests and expectations.

Once the coordinator exposed the initial project, the collaborators concluded that there were already other tools addressing deliberative processes through aggregation methods, and only a few included some specific artificial intelligence techniques (see Table 1 in discussion section). Thus, an initial conceptual mapping was developed in order to find out some gaps in the existing participatory and deliberative tools. After comparing a number of digital tools, the team decided to focus on three dimensions: 1. how citizens can participate and express their preferences; 2. to what extent the decisions taken are binding; and 3. the existence of monitoring mechanisms to follow-up the decisions approved. Afterwards, the work team was divided into groups to work these dimensions focusing specially on the design of the different interface's sections: 1. participation: how to create proposals and how to structure and aggregate argumentation; 2. decisions (outputs): show how the final list of proposals that have been selected according to pre-established priorities, citizen support and budget; and 3. follow-up (outcomes): tracking the fate of the proposals (implementation, modification or cancellation). Finally, the resulting preliminary prototype conceptualization and design was presented (and discussed with other groups and participants) at the closing event of the workshop.

4. CO-GOVERN: A PARTICIPATORY TOOL PROPOSAL

Figure 1 details the overall participatory process in *CoGovern,* which can be described in terms of the following stages:

1. To release the functioning rules: Initially, the authority releases the necessary information to fully understand the process so to be aware of the "rules of game": informing about the prioritization criteria: i) citizen support; ii) budgetary aspects; and iii) alignment with a set of priorities or strategic goals. An information sharing event can help to disseminate and clarify the participatory process as well as to learn how the online tool works.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. To create proposals: Upon login, city residents and government can create proposals that are classified by administrative areas in case of a process in which different sectorial actions are aimed to be carried out. Users can also obtain information on the other proposals.
3. To deliberate and vote: A key feature of *CoGovern* is that it supports citizens to vote for and/or deliberate about proposals at their own convenience. The deliberative interface (Figure 2) shows the distribution of votes (it also offers the possibility of voting), as well as pro and cons arguments. Citizens can both provide new arguments and express their agreement, neutrality, or disagreement in relation to their peers' arguments. Argument information is aggregated[5] and added to the general voting by modifying the level of citizen support of each proposal.
4. To filter and fully specify proposals: After reviewing the contributions made by citizens, authorities' technical staff filter out illegal proposals and introduce both the cost and their strategic alignment (since most probably proposals will vary in their support to stated priorities or strategic goals) for the remaining proposals. Moreover, technicians are also required to specify those incompatibility relationships that hold between proposals. Such incompatibilities correspond to objective –and documentable– impediments for proposals to be simultaneously selected. Thus, for example, two proposals of alternative urban interventions designed for the same location are infeasible. This specification relates to the recommendations of the Participatory Budget Partners in UK,[6] that state that this indispensable technical input "might include reference to existing or planned spending proposals, to avoid duplication in allocating resources."
5. To select proposals: Once the data is gathered –citizen support (after computing voting and argumentative support), costs and strategic alignment– the proposals are selected according to an algorithm publicly known beforehand. Next subsection details how this selection process is automated.
6. To visualize the results: Users can inspect each proposal. Figure 3 depicts its state (selected or not) and its alignment with the considered criteria (citizen support; budgetary constraints, and alignment with pre-established priorities or strategic objectives) by means of colored bars.
7. To follow up selected proposals: Selected proposals are ordered according to its implementation degree. Users can also monitor their implementation schedule, deviations from the initial budgeting, and companies in charge. Once the project is completed, a report with all the information (Figure 4) will become available.

4.1. How Does Artificial Intelligence Operates in CoGovern?

Leaving aside the number of proposals to deal with, choosing the best combination of proposals constitutes a complex task, since selected proposals must satisfy incompatibility constraints and balance different criteria. Thus, for instance, it could be the case that the proposal with highest citizen support turns out to be so expensive that it alone depletes the overall available budget so that there is no budget left for implementing other interesting proposals. Alternatively, it could also be the case that the chosen proposals are not aligned with the administrative strategic objectives. For example, if promoting social cohesion is very important for the authorities, then a proposal dividing two neighborhoods should not be prioritised. This is so because co-governance is about seeking agreements between citizens and authorities, deciding upon those sets of proposals that are good enough for all the actors involved in the process. Overall, it is fundamental for transparency and accountability that the selection process is properly publicized, that it is objective, and that it encompasses rules for balancing different criteria that are clear from the very beginning, before the deliberation process starts.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. CoGovern proposal selection process

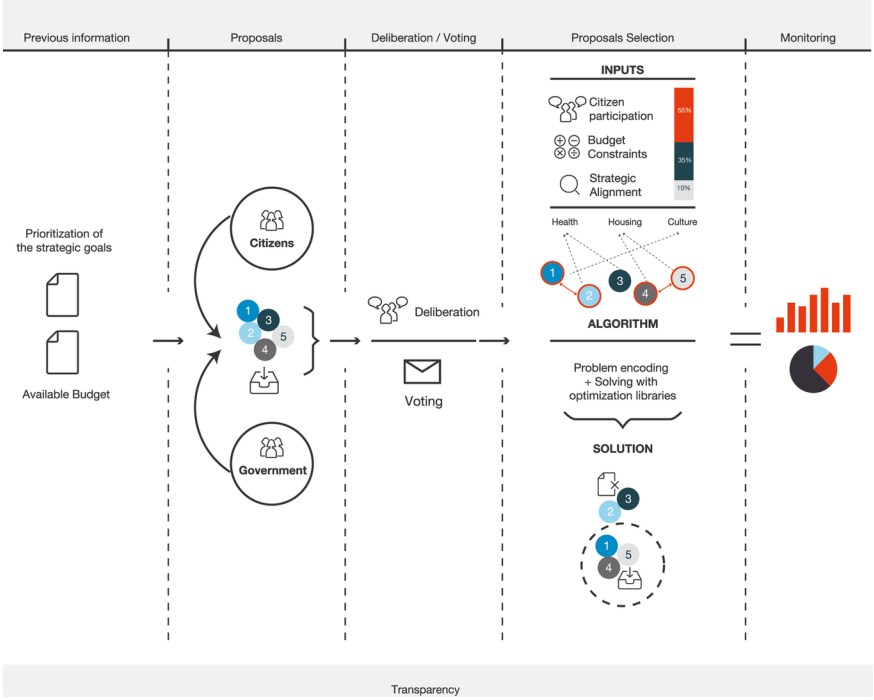
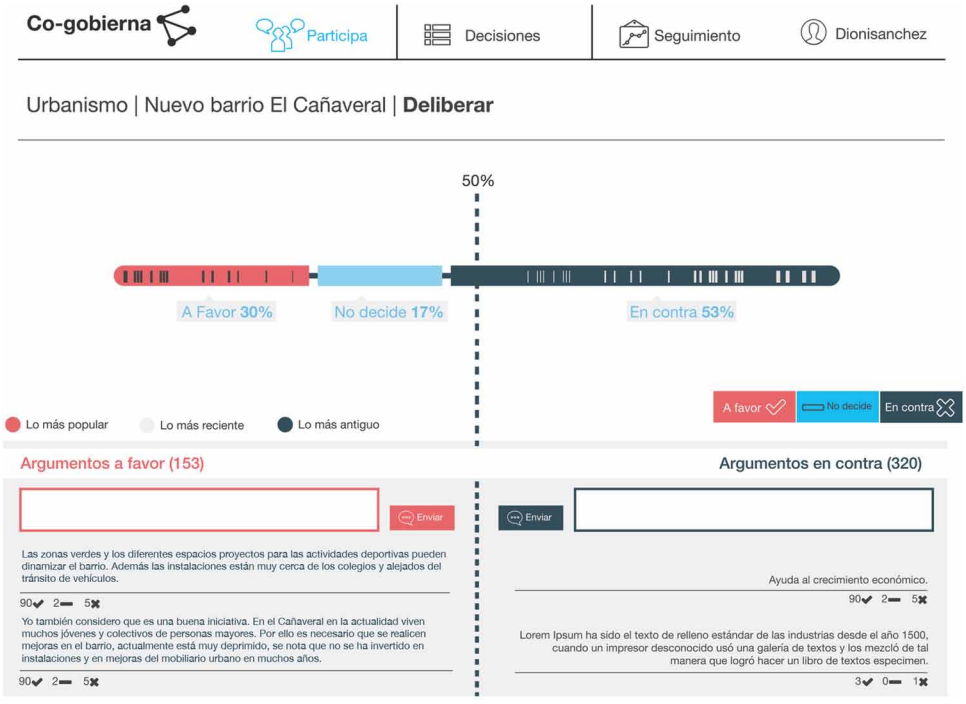


Figure 2. CoGovern participation interface for argumentation



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. CoGovern accountability interface

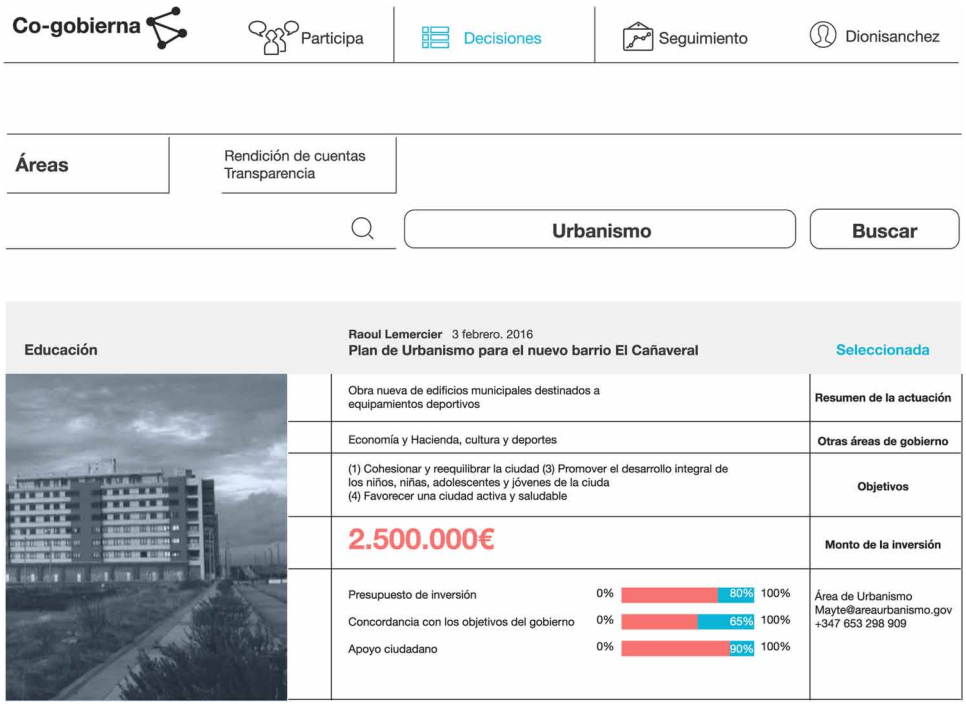


Figure 4. CoGovern monitoring interface



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Making explicit the selection process requires some details to be released. Firstly, the overall available budget is needed not only to know how much can be spent but also to characterize proposals accurately. Secondly, since several criteria are taken into account (citizen support, budgetary constraints, and alignment with priorities or strategic objectives), we need to know if all of them are considered to be equally important or if there are some preferences over some of them. If this is the case (e.g., citizen support is considered to be more important than cost) then the actual weights have to be clearly stated. Similarly, it is also necessary that priorities or strategic objectives are prioritized among them, so that they are presented in an orderly manner. Finally, in addition to the specific characteristics of each proposal, such as citizen support or implementation cost, it is also necessary that proposal incompatibilities are clearly displayed so that whenever a proposal is chosen, all participants understand that those other proposals that are incompatible with it cannot be possibly chosen.

Once this information becomes clear, it is possible to compute the best subset of proposals to implement. This selection process can be automatized by applying optimization techniques. More specifically, we propose to specify an optimization problem that seeks to maximize 1. the number of proposals; 2. the citizen support; 3. the support to administrative strategic objectives; whilst minimizing cost. Optimization problems can be encoded as linear programs that can be automatically solved by means of state-of-the-art solvers (i.e., optimization code libraries). In particular, considering a set of proposals P, we encode the problem by defining a set of binary decision variables $\{x_1,..x_{|P|}\}$ where each x_i e$\{0,1\}$ encodes the decision of whether proposal p_i will be selected (thus taking value 1) by the system or not (i.e., $x_i = 0$). Therefore, solving the problem amounts to solving the following linear program:

$$\max \sum_{i=1}^{|P|} x_i$$

$$\max \sum_{i=1}^{|P|} s\left(p_i\right) \cdot x_i,$$

being $s(p_i)$ a function that provides the support obtained by each proposal p_i

$$\min \sum_{i=1}^{|P|} c\left(p_i\right) \cdot x_i$$

being $c(p_i)$ a function reporting the cost of each proposal p_i

$$\max \sum_{i=1}^{|P|} v_p\left(p_i\right) \cdot x_i$$

being $v_p(p_i)$ a function that considers some values associated to the ordered strategic objectives so that, if we assume there are |O| strategic objectives, then we have $v_1 > \ldots > v_{|O|}$, where each v_i corresponds to the i-th strategic objective. Then, using an intermediate function f that specifies, for each proposal p_i, the strategic objectives it is aligned with, we compute $v_p(p_i) = \sum_{j \in f\left(p_i\right)} v_j$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

We can cast this problem specification as a single objective optimization problem that can be solved by the following linear program (refer to Lopez-Sanchez et al., 2017 for a more comprehensive and detailed explanation):

$$\max\left[\frac{w_r}{|P|}\cdot\sum_{i=1}^{|P|}x_i+\frac{w_s}{S_{max}}\cdot\sum_{i=1}^{|P|}s\left(p_i\right)\cdot x_i+\frac{w_v}{V_{max}}\cdot\sum_{i=1}^{|P|}v_p\left(p_i\right)\cdot x_i+\boldsymbol{w}_c\cdot\left(y-\frac{1}{b}\cdot\sum_{i=1}^{|P|}c\left(p_i\right)\cdot x_i\right)\right] \quad (1)$$

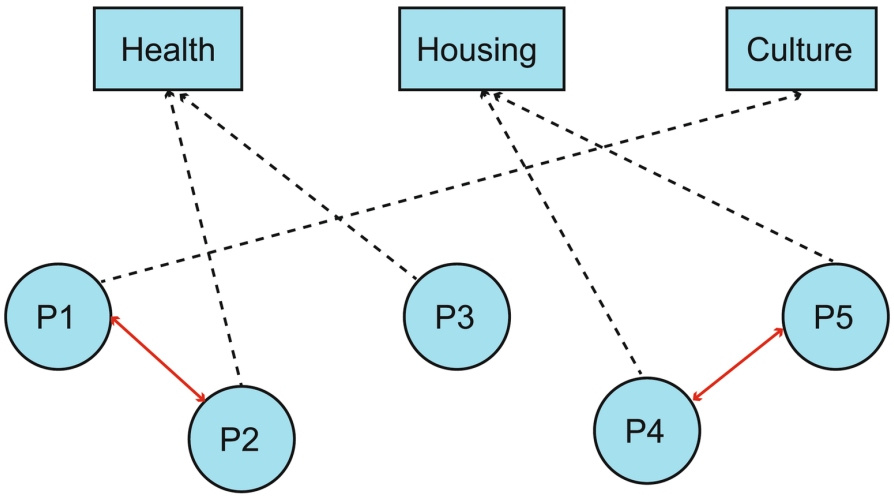
Subject to a number of constraints encoding maximum budget, incompatibility relationships or the importance awarded to criteria.

Figure 5 illustrates the resolution of this problem with an example: Imagine decision makers have the following priorities: 1. Promote health; 2. Provide access to affordable housing; and 3. Promote culture. We consider they are ordered in decreasing order of importance (that is, health is the most important priority, and it gets a value v_1=3 whereas culture is the least preferred one with value $v_3 = 1$). Additionally, we consider five different proposals promoting some of these values and described in Table 1. Notice that, as depicted in Figure 5, p_1 and p_2 are exclusive (i.e., they cannot be simultaneously chosen). This is so because they would be built in the same plot of land. Similarly, p_4 and p_5 are exclusive because urban architects constitute a scarce resource that can only work on a single urban plan at a time.

Last column in Table 1 shows how the standard way of choosing proposals would order them by just considering citizen support.

Considering a maximum budget of 5 million (b=5), the result would be to choose p_5 which has the highest support ($s(p_5)$ =0.9) and requires most budget ($c(p_5)$ =4). Unfortunately, next proposal in order (p_4, with 0.8 of support) cannot be included because it is incompatible with p_5. Subsequent proposal (p_2) cannot be included either because it costs 2 million and its addition to p_5 would exceed the maximum budget b=5. An alternative may be to add next proposal in the ordered list that would fit the available budget (i.e., p_1, which costs 1M). Thus, if we choose p_5 and p_1 we do not exceed the total budget.

Figure 5. Example of proposals (circles) and priorities (rectangles). Dashed (black) lines relate proposals to the priorities they promote. Solid (red) lines relate incompatible proposals



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Proposals example. Source: own elaboration

Proposal p_i	Description	Citizen Support $s(p_i)$	Cost $c(p_i)$	Strategic Objective $f(p_i)$	Standard Ranking
p_1	Build a library.	0.6	1	culture $f(p_1) = 3$	4
p_2	Build a new medical assistance center	0.7	2	health $f(p_2) = 1$	3
p_3	Improve ambulance services	0.4	1	health $f(p_3) = 1$	5
p_4	Rehabilitation of a deteriorated neighbourhood	0.8	2	housing $f(p_4) = 2$	2
p_5	Promotion of a new neighbourhood	0.9	4	housing $f(p_5) = 2$	1

Alternatively, we propose to consider the optimisation mechanism previously introduced that does not only consider citizen support, but it also combines both the cost and the priorities within a single evaluation function. In this case, any state-of-the-art solver such as CPLEX (CPLEX 2009), Gurobi (Gurobi) or GLPK (Makhorin, 2000) that is provided with the linear program encoded in Equation 1 would produce as a solution the combination composed by proposals p_2, p_3, and p_4, since it corresponds to the best possible proposal combination.

5. DISCUSSION

This article introduces a proposal for a digital tool for citizen participation. But, in which contexts and spaces can it be applied? Transferability and institutionalization have been central in the tool's design. The former is the capacity of a tool of democratic innovation to "operate in different political contexts, understood in relation to scale, political system or type of issue" (Smith, 2009:13), whereas institutionalizing means "moving from sporadic, episodic experiments toward embedded and routinized and easily replicable structures within government that last beyond the initial enthusiasm" (Russon-Gilman, 2016, p. 119). Taking into account these two principles three main decisions were made during the participatory prototyping process in order to shape the participatory device proposed in this article.

Firstly, the online tool proposed was conceived for two widespread types of participatory processes: participatory budgeting and strategic planning. Through participatory budgeting, authorities allocate a portion of the municipal budget for citizens to decide how to invest it. As for strategic planning, it projects the main lines of action in mid and long terms and seeks to fulfill governance objectives. Both participatory processes can be done on-line. Nevertheless, we should be prudent about the rapid spread of ICTs in participatory democracy. For example, some engaged citizens can perceive a loss of civic rewards as consequence of the digitalization of the participation. In this sense, Russon-Gilman alerts about "how ITCs could alter the character of deliberation and substantive participation within Participatory Budgeting (2016:129-130;138)". Obviously, a face-to-face deliberation cannot be equated to a process based on online argumentation. However, in our opinion, both have pro and cons. On one hand, argumentation allows a greater systematization, although the threat of the digital divide is still an issue. On the other hand, face-to-face deliberation has its own biases since some citizens are not able or willing to attend assemblies due to a lack of time, their different participatory styles, or simply, their fear to speak in public. The solution of the dilemma between online and offline participation seems to be an hybridization. For example, although proposals in participatory budgeting processes are uploaded

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

and voted through on-line applications, currently, some authorities (such as in Getafe, Spain) also offer technical support for those in need at designated physical locations (e.g., civic centres).

Secondly, the approach proposed in *CoGovern* highlights the importance of the convergence between citizens and authorities' perspectives. This recommendation is also shared by the Participatory Budgeting Partners in UK, among others. Nevertheless, it is worth noticing that alignment can lead to an excessive framing and, consequently, some groups may consider this type of participation too constrained. Therefore, before starting the participatory process it becomes key to explain that this alignment pretends to guarantee a high level of political compromise, which is essential for a successful implementation of the results of participatory processes. Moreover, it should be noted that priorities or strategic objectives, and their respective weights, may well be previously agreed between authorities and the community. In fact, this is in line with the principle of transferability which implies a certain degree of flexibility within and adaptation to specific contexts. For example, since 2016 the participatory budgeting of the city of Paris prioritizes projects made on deprived or sensitive areas, reserving 30% of the allocated project. *CoGovern* is much more open and simply makes explicit all the priorities or strategic objectives that will be considered in the selection process.

Thirdly, previous research -focused on democratic innovations- had already put the proposal selection problem on the table (Smith, 2009: 93; Font et al., 2017). These studies found out preliminary evidence about the existence of cherry-picking practices by authorities in the context of local participatory processes in Spain. This article has shown how this specific aspect within institutional participatory processes can benefit from certain AI techniques. However, it must be brought into discussion some considerations about the expansion of the use of algorithms for social purposes. Mattelart & Multigner (2007) – when referring to Gottfried Wilhelm Leibniz – alerted us to the use of algorithms for speculative purposes and their relation to modern capitalism. More recently, critical voices have related its use with the idea of cybercontrol. It is therefore imperative to seek new ways to make use of, visualize, and build the "algorithmic environment" (Álvaro, 2014). Can algorithms contribute to the social progress through its integration in social projects? Through codesign (Fuller, 2003), *CoGovern* has shown an attempt to bring the "algorithmic environment" closer to the society so as to be able to participate in its formulation. Enabling the dialogue between citizens, rulers and experts is an unavoidable task. It is a need, therefore, to create more spaces, such as citizens labs, in which citizens are confronted with the complexity of the institutions and technical issues, and institutions and experts adopt a sincere intent to (co)govern together with the people.

6. RELATED AND FUTURE WORK

This article proposes *CoGovern* as an online participation tool focused on incorporating citizen preferences to local policy making by means of artificial intelligence techniques. Thus, an unavoidable question has to do with what new elements has *CoGovern* brought into the landscape of participatory and deliberative tools. During the international workshop previously mentioned, invited experts introduced some of the most relevant participatory tools. Table 2 below summarizes the principal similarities and differences between these tools and *CoCovern*.

First, in relation to the main *goals* pursued by these different tools, *CoGovern* and *Decide Madrid* are both oriented to incorporate city residents preferences into policy decision-making processes at local level. The rest of tools, with the exception of the French tool *Parlement et Citoyens* (P&C), do not seek

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Comparison of different participatory tools. Source: own elaboration

Tools	Goals	Types of users	Online argumentation	How do citizens decide?	Monitoring	AI
CoGovern	Incorporate citizens preferences in policy decision-making processes.	City residents	Yes (pro and cons arguments)	Arguments / voting	Yes	Yes (optimal proposal selection)
Loomio	Help small groups make decisions through inclusion	Small groups	Yes (messages)	Arguments / voting	No	No
Baoqu	Solve complex discussions	Large groups	Yes (comments)	Groups of 5 to 25 users in subsequent levels that require qualified majority	No	No
Decide Madrid	Create proposals, deliberate and vote	City residents	Yes (comments)	Popular referendum	Feasibility reports	No
Appgree	Create representativeness of a whole through small groups	Large groups	No (answers to questions)	Representative voting for answers	No	(Demorank)
Quoners	Help groups decide on what they agree (or disagree) with	Large and small groups	Yes (pro and cons arguments)	Vote for multiple answers	No	No
Consider.it	Create a discussion forum: participants opinions grouped in favor or against any topic	Large groups	Yes (pro and cons arguments)	No voting (decision based on opinions)	No	(Doctoral research)
Parlement et Citoyens	Tool that allows citizens and deputies to collaboratively draft legislative proposals	Large groups	Yes (pro and cons arguments)	Arguments / voting	No	No

to influence political decisions. P&C allows citizens and deputies to collaboratively draft legislative proposals, therefore, it represents an interesting attempt to scaling-up in terms of political influence beyond local contexts. Second, regarding the *types of users,* three main groups were identified: small groups, large groups, and citizens residents. Small groups refer to, for example, co-workers, family, friends or activists, whilst large groups lack specific shared interests, locations, or strong ties. In contrast, citizens residents refer to the whole set of people officially registered in a particular city. *CoGovern*, therefore, falls clearly into the latter group in which is included *Decide Madrid* as well as other tools developed by local administrations (not included in the table), such as *Decidim Barcelona, Participa Getafe* and *Budget Participatif* in Paris. Third, as for argument functionalities, although all the tools considered here allow participants to include additional information supporting decisions, most of them are treated as comments. Just *CoGovern*, quoners, consider.it, and P&C do structure them as arguments in favour and against. Fourth, regarding *how do citizens decide,* some of them combine the use of online argumentation and voting systems. That is the case of *CoGovern*, *Loomio* and *P&C*. In *consider.it*, by contrast, decisions are not taken through traditional voting, since the aim is to aggregate opinions (not votes). On the other hand, in *Decide-Madrid,* those proposals that have received support from at least 1% of the whole population in Madrid can be voted online or in polling stations during "the referendum week". *Appgree* and *Quoners* offer different modalities of voting whereas *Baoqu* develops a scalable system for massive deliberation based on agreements in small groups which scales up to higher decisions levels. Fifth, the existence of *monitoring systems* to trace the fate of proposals is especially relevant for our goal of increasing transparency and accountability. As table 1 shows, current tools lack in this respect. Lastly, the most innovative aspect in *CoGovern* is the usage of *AI* in the decision-making process, which seems

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

to be in early development stages for current tools. *Appgre* and *Consider.it* are the tools that include more sophisticated computation or research. Nevertheless, rather than taking decisions, *Appgree* uses an algorithm (DemoRank) to randomly choose a small group of people as representative of the whole, whilst *Consider.it* graphically clusters people based on their opinions.

The future research of the *CoGovern* tool focuses mainly on developing the prototype until reaching an alpha phase, which will allow us to test it with groups of citizens and policymakers to detect which parts can be improved when analyzing the user experience. Specifically, as was done at the beginning of the project (see point 3.1), we should check the degree of satisfaction between citizens and policymakers with both the usability of the tool and the proposals selection method presented in this article. In relation to the latter, we plan to present the outcome of the decision (list of proposals) as a comparison of the traditional model of vote-counting and the optimization proposed by *CoGovern* in order to evaluate the user's reaction.

Beyond implementation of the current proposal, we also plan to include improvements such as the enrichment of the argumentation, so that pro and con arguments can also be counter-argumented (although simplicity forces us to prevent further counter-argurments). Finally, another future objective concerns having the rules of the game as transparent as possible, since the proposals selection process constitutes the most sensitive feature of *CoGovern*. In this sense, we consider it fundamental that we be able to communicate in a very intuitive manner about the internal functioning of the tool. To that end, we plan to explore visual metaphors to help explain not only the results of the selection process but also how decisions were taken.

REFERENCES

Allen, N., & Birch, S. (2015). Process preferences and British public opinion: Citizens' judgements about government in an era of anti-politics. *Political Studies*, *63*(2), 390–411. doi:10.1111/1467-9248.12110

Álvaro, S. (2014). El poder de los algoritmos: cómo el software formatea la cultura. *CCCBLab*.

Awad, E., Booth, R., Tohmé, F., & Rahwan, I. (2015). Judgement aggregation in multi-agent argumentation. *Journal of Logic and Computation*.

Bengtsson, Å., & Christensen, H. (2016). Ideals and actions: Do citizens' patterns of political participation correspond to their conceptions of democracy? *Government and Opposition*, *51*(2), 234–260. doi:10.1017/gov.2014.29

Bherer, L., Fernández-Martínez, J. L., García Espín, P., & Jiménez Sánchez, M. (2016). The promise for democratic deepening: The effects of participatory processes in the interaction between civil society and local governments. *Journal of Civil Society*, *12*(3), 344–363. doi:10.1080/17448689.2016.1215957

Bohman, J. (1998). Survey article: The coming of age of deliberative democracy. *Journal of Political Philosophy*, *6*(4), 400–425. doi:10.1111/1467-9760.00061

Briscoe, G. (2014). Digital innovation: The hackathon phenomenon.

Brodersen, C., Dindler, C., & Iversen, O. S. (2008). Staging imaginative places for participatory prototyping. *CoDesign*, *4*(1), 19–30. doi:10.1080/15710880701875043

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chadwick, A. (2009). Web 2.0: New challenges for the study of E-democracy in an era of informational exuberance. *I/S: A Journal of Law and Policy for the Information Society, 5*(1), 9-41.

Wikipedia. (2009). CPLEX. Retrieved 27 April 2017, from https://en.wikipedia.org/wiki/CPLEX

Cramton, P., Shoham, Y., & Steinberg, R. (2006). *Combinatorial Auctions*. Mass.: MIT Press Cambridge.

Dahlberg, L. (2011). Re-constructing digital democracy: An outline of four 'positions'. *New Media & Society*, *13*(6), 855–872. doi:10.1177/1461444810389569

Font, J., Smith, G., Galais, C., & Alarcón, P. (2017). Cherry-picking participation: Explaining the fate of proposals from participatory processes. *European Journal of Political Research*. doi:10.1111/1475-6765.12248

Font, J., Wojcieszak, M., & Navarro, C. J. (2015). Participation, representation and expertise: Citizen preferences for political decision-making processes. *Political Studies*, *63*(1 Suppl.), 153–172. doi:10.1111/1467-9248.12191

Fuller, M. (2003). *Behind the blip: Essays on the culture of software*. New York: Autonomedia.

Fung, A., & Wright, E. O. (2001). Deepening democracy: Innovations in empowered participatory governance. *Politics & Society*, *29*(1), 5–41. doi:10.1177/0032329201029001002

Garcia-Heveling, W. (2017) Citizen Participation: Norm consensus in virtual communities http://www.maia.ub.es/~maite/papers/memoriaWaltraud.pdf

Gurobi.com. (n.d.). Gurobi Optimization - The Best Mathematical Programming Solver. Retrieved 27 April 2017, from http://www.gurobi.com/

Hibbing, J. R., & Theiss-Morse, E. (2002). *Stealth democracy: Americans' beliefs about how government should work*. Cambridge University Press. doi:10.1017/CBO9780511613722

Hillgren, P., Seravalli, A., & Emilson, A. (2011). Prototyping and infrastructuring in design for social innovation. *CoDesign*, *7*(3-4), 169–183. doi:10.1080/15710882.2011.630474

Hwang, C., & Masud, A. S. M. (2012). *Multiple objective decision making—methods and applications: A state-of-the-art survey*. Springer Science & Business Media.

IBM. (n.d.). *IBM CPLEX Optimizer*. Retrieved 27 April 2017, from https://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/

Keshavarz, M., & Maze, R. (2013). Design and dissensus: Framing and staging participation in design research. *Design Philosophy Papers*, *11*(1), 7–29. doi:10.2752/089279313X13968799815994

Klein, M. (2012). Enabling Large-Scale Deliberation Using Attention-Mediation Metrics. *Computer Supported Cooperative Work*, *21*(4-5), 449–473. doi:10.100710606-012-9156-4

Lazer, D. (2015). Social sciences. the rise of the social algorithm. *Science*, *348*(6239), 1090–1091. doi:10.1126cience.aab1422 PMID:25953819

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lee, Y. (2008). Design participation tactics: The challenges and new roles for designers in the co-design process. *CoDesign*, *4*(1), 31–50. doi:10.1080/15710880701875613

London, S. (1995). Teledemocracy vs. deliberative democracy: A comparative look at two models of public talk. *Journal of Interpersonal Computing and Technology*, *3*(2), 33–55.

Lopez-Sanchez, M., Serramia, M., Rodriguez-Aguilar, J. A., Morales, J., & Wooldridge, M. (2017). Automating decision making to help establish norm-based regulations. Retrieved from https://arxiv.org/pdf/1703.09087.pdf

Makhorin, A. (2000). Department for Applied Informatics, Moscow Aviation Institute, Moscow, Russia. Retrieved from https://www.gnu.org/software/glpk/

Manzini, E., & Coad, R. (2015). *Design, when everybody designs: An introduction to design for social innovation*. MIT Press.

Manzini, E., & Rizzo, F. (2011). Small projects/large changes: Participatory design as an open participated process. *CoDesign*, *7*(3-4), 199–215. doi:10.1080/15710882.2011.630472

Mattelart, A., & Multigner, G. (2007). *Historia de la sociedad de la información*. Barcelona: Paidós.

Mazé, R. (2014). Our common future? political questions for designing social innovation. In *Proceedings of the Design Research Society Conference*.

Rodriguez-Aguilar, J. A., Serramia, M., & Lopez-Sanchez, M. (2016). Aggregation Operators to Support Collective Reasoning. In *Modeling Decisions for Artificial Intelligence*, *LNAI* (Vol. *9880).* Springer.

Russon-Gilman, H. R. (2016). *Democracy reinvented: Participatory budgeting and civic innovation in America*. Brookings Institution Press.

Secchi, M., & Spada, P. (2017). Democracy is a serious game. In *Proceedings of the General Conference at European Consortium for Political Research*, Oslo, Septembre 6-9.

Seravalli, A. (2011). Democratizing production: challenges in co-designing enabling platforms for social innovation. In *The Tao of Sustainability, International Conference on Sustainable Design Strategies in a Globalization Context*, October 27-29 (pp. 494-500).

Serramia, M. (2016). Information fusion for norm consensus in virtual communities. Retrieved from http://www.maia.ub.es/~maite/serramia_tfg.pdf

Serramia, M., Lopez-Sanchez, M., & Rodriguez-Aguilar, J. A. (2016). Aggregation Operators to Compute Norm Support in Virtual Communities. In *Proceedings of the 10th Workshop on Advances in Preference Handling (MPREF 2016) in IJCAI 2016 (25th International Joint Conference on Artificial Intelligence), New York, July 9.*

Smith, G. (2009). *Democratic innovations: Designing institutions for citizen participation*. Cambridge University Press. doi:10.1017/CBO9780511609848

Socientize Consortium. (2013). *Green paper on citizen science.*

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Špundak, M. (2014). Mixed agile/traditional project management methodology–reality or illusion? *Procedia: Social and Behavioral Sciences*, *119*, 939–948. doi:10.1016/j.sbspro.2014.03.105

Torra, V., & Narukawa, Y. (2007). *Modeling decisions: Information fusion and aggregation operators*. Springer Science & Business Media.

ENDNOTES

1. Rare exceptions, such as the one in Paris described in the discussion section, do also consider additional constraints that impose ad-hoc priorities.
2. As explained in the discussion, GoGovern has been designed to be applied mainly in two well-known types of participatory processes: participatory budgeting and strategic planning. Most often, we refer to priorities in the former and to strategic objectives in the latter, which may be more accurate.
3. Alpha projects: projects in demo phase that require a last revision to better define its functionalities. Projects in state of production: projects that are in the final phase of its development.
4. CoGovern project members, in alphabetical order: Nella Charms, Jennifer Clavijo, Ana Doria, José Luís Fernández-Martínez, Maite Lopez-Sanchez, Lizeth Ramírez, Juan A. Rodríguez-Aguilar, Jairo Salazar, Dionisio Sánchez, Marcelo Sánchez, Sayuri Susuki, and Berenice Zambrano.
5. By means of WOWA operators (see Serramia et al. 2016 and Torra & Narukawa, 2007)
6. https://pbnetwork.org.uk/wp-content/uploads/2016/10/PB-Mainstreaming.pdf (page 9)

This research was previously published in the International Journal of Public Administration in the Digital Age (IJPADA), 5(4); edited by Manuel Pedro Rodríguez Bolívar; pages 1-17; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 43
Exploring the Conceptual Nature of e-Business Projects

Benjamin Matthies
South Westphalia University of Applied Sciences, Germany

André Coners
South Westphalia University of Applied Sciences, Germany

ABSTRACT

E-business projects (eBPs) can be described as endeavors which transform traditional business functions into digital and mostly automated, inter-organizational e-business functions. Although studies exist which already increased the understanding of such endeavors, there is still no conceptualization that defines eBPs as their own kind of project, summarizing and categorizing all of their characteristic topics, problem areas, and tasks. This study attempts to close this gap by exploring the conceptual nature of eBPs. A large number of practical project reports was evaluated, using a latent semantic analysis for the exploratory summarization of the textual database. As a result, the study provides a comprehensive theoretical conceptualization including 12 core concepts and 74 specific sub-concepts, structured within the context of a system development life cycle. Such a project conceptualization provides a contribution to the practical management of eBPs, as well as a theoretical framework for further establishing eBPs as a unique project type.

1. INTRODUCTION

Transforming a company into a successful e-business has become one of the central strategic goals of modern corporate management (Ash & Burn, 2003; Chen, 2001; Gloor, 2012; Lientz & Rea, 2009). Such e-business transformations are basically about developing and implementing Internet-based technologies in order to pursue modern e-business models (Barua et al., 2001; Pateli & Giaglis, 2004; Wang et al., 2012). During this process, traditional business functions (such as sales, purchasing, and logistics) are expanded by integrating digital and inter-organizational e-business functions (such as e-commerce, e-procurement, or network-based supply chain management), typically resulting in fundamental changes

DOI: 10.4018/978-1-7998-0951-7.ch043

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

to the corporate value creation chain (Burn & Ash, 2001; Combe, 2012; Jackson & Harris, 2003). Therefore, these types of endeavors – so-called e-business projects (eBPs) – are particularly complex, and do not end after the purely technical implementation of new information technologies and systems. Instead, they rather initiate and accompany comprehensive corporate transformations, which extend beyond the technical aspects of typical IT projects in several ways (Chuang & Shaw, 2005). The consequences of eBPs require companies, for example, to align their corporate strategies profoundly before integrating e-business functions into the changing value creation; to coordinate the integration of new communication channels which change the way companies interact with customers; to collaborate with new stakeholders in inter-organizational alliances, to deal with new and often extensive data volumes; and, last but not least, to fundamentally redesign their business processes and the way employees execute their work (Balakian et al., 2002; Marca, 2006; Stoehr, 2002). This large number of diverse commercial, organizational, and technical considerations – which are often interdependent – ultimately results in an increased degree of complexity in such projects (Baccarini, 1996). As a consequence, even though eBPs basically encompass the implementation of e-business technologies, they can be described as comprehensive and externally-oriented business projects rather than as technical, internally-oriented IT projects. Accordingly, project management strategies, processes, and practices must be adjusted to match such complex undertakings (Phan & Vogel, 2012; Stoehr, 2002).

Because of their complexity, it is not uncommon for ambitious eBPs to fail (McLaughlin, 2009). As a result, research has shown an increased interest in exploring eBPs as a specific and unique type of project (see, e.g., Basu & Muylle, 2013; Clarke & Doherty, 2004; Lee et al., 2007; Phan, 2003; Phan & Vogel, 2012). Such studies deal with questions concerning, for example, which characteristics define eBPs, which specific challenges they encompass, and, of course, how to meet these. Nevertheless, although previous studies have significantly increased our understanding of eBPs, there is still no overall conceptualization – such as, for instance, the many conceptualizations that exist for general information systems development projects (see, e.g., Keil et al., 2000; Xia & Lee, 2005) – that structures eBPs as their own kind of project, summarizing and categorizing all of their characteristic topics, problem areas, and tasks. Such a project conceptualization could be of assistance in coming to understand the complexity of eBPs and, ultimately, to manage it (Gregory & Piccinini, 2013). In this regard, a conceptualization of eBPs would provide a basic content-based structure for their practical project management (e.g., during requirements analyses) and theoretical explorations (e.g., during success factors research).

This study addresses the research gap described and attempts to illustrate the complexity of eBPs through a comprehensive theoretical conceptualization. In this context, the study focuses on the following core question: Which specific topics, problem areas, and tasks characterize eBPs as a complex and unique type of project? Such a thematic project conceptualization is a recommended contribution for project management (Winter et al., 2006). In this respect, the contributions of this study differ from previous studies in two ways: (1) Unlike previous studies, this study is based on an evaluation of a large number of eBPs ($N = 355$). A collection of practical project reports will serve as a database for this purpose. An automated latent semantic analysis (LSA) will be used for the exploratory summarization of the primarily textual contents (> 2,000 pages). (2) For the first time, this study characterizes eBPs using a comprehensive thematic conceptualization including 12 core concepts and a total of 74 detailed sub-concepts. These thematic concepts are embedded in the context of a typical system development life cycle (SDLC), giving the conceptualization a founding framework. Finally, subsequent discussions work out and illustrate the conceptual nature of eBPs.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The research goal of this study may therefore be summarized as:

The development of a comprehensive thematic conceptualization of e-business projects based on a summary evaluation of practical project reports.

This article is structured as follows: Section 2 will discuss the basic characteristics of eBPs, presenting associated studies. Section 3 will introduce the methodology of the study, describing the technique of LSA and its application for the evaluation of project documentations. Section 4 will present the developed conceptualization of eBPs and discuss its thematic concepts. Section 5 will deal with its implications for practice and research, as well as with potential limitations of the study. Section 6 will provide a concluding summary and a short overview.

2. CHARACTERISTICS OF E-BUSINESS PROJECTS

Stoehr (2002) describes the nature of e-business projects (eBPs) as follows: "The similarity among all e-business projects is that their objective is to build up an e-business solution that centers around the internet" (p. 2). In this context, an e-business solution can be described as a development of digitally enabled information technologies and inter-organizational systems in order to pursue modern e-business models with mostly automated e-business functions (Pateli & Giaglis, 2004; Wang et al., 2012). The effects of eBPs therefore extend beyond the purely technical implementation of information technologies and systems, and typically imply a significant shift in corporate value creation and organizational structures and processes (Chuang & Shaw, 2005; Combe, 2012; Jackson & Harris, 2003). In order to gain a deeper understanding of the nature of eBPs, the authors explored the literature focusing on studies which examined and described eBPs from a more comprehensive perspective. The authors were able to find several scientific studies that have already tried to characterize eBPs as a complex and unique type of project (see Table 1)[1], providing discussions of their specific features. In general, the following central characteristics and associated challenges can be summarized on the basis of these studies:

- eBPs often radically change the business model of a company (Gordijn & Akkermans, 2001; Pateli & Giaglis, 2004). These changes are not restricted to exploiting new markets and sales channels, but relate also the entire value-creation chains (Amit & Zott, 2001; Burn & Ash, 2001). Therefore, eBPs require integrating the e-business solution into existing business functions in a specific manner, expanding them purposefully (Phan & Vogel, 2012). This is why developing an appropriately well-founded e-business strategy in advance, and aligning it with general company goals, is considered a central cornerstone for the success of eBPs (see, e.g., McLaughlin, 2009; van Grembergen & Amelinckx, 2002).
- eBPs generally have a strong external influence on their customers, especially on the way companies communicate and interact with them. This is why a customer-oriented focus is often emphasized as an important factor in conceptualizing and implementing e-business solutions (see, e.g., Phan, 2002).
- eBPs implement modern information and communications technologies (ICT) and inter-organizational systems (IOS), which are the foundation for facilitating collaborative, typically automated business processes between multiple networks and stakeholders (Chuang & Shaw, 2005; Shaw,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2003; Wang et al., 2012). eBPs therefore require intensive cooperation with diverse, mostly external, partners (see, e.g., Clarke & Doherty, 2004), as well as a high degree of standardization in order to harmonize their actions (see, e.g., Phan, 2009; Zhao et al., 2007).

- eBPs imply significant changes to business processes and to the way employees perform their work (Jackson & Harris, 2003). Therefore, an effective change management (Burn & Ash, 2001) and the targeted development of new responsibilities and skills are often emphasized tasks in eBPs (see, e.g., Phan, 2002; Tiwana & McLean, 2003).
- eBPs are often confronted by innovative but, at the same time, untested and continually changing technologies (Chang, 2010). As a consequence, eBPs also often have far shorter technical cycle times than traditional IT projects (Mehler-Bicher, 2002). These circumstances demand remarkable flexibility and creativity from project teams, due to potential technical or content-based variations (Daniel & Wilson, 2003; Lee et al., 2007).
- eBPs offer great potential for innovation, on both technical and business levels. This is why technical capabilities are relevant here, as well as a profound understanding of the business domain, the customer, the organization, and business processes. Consequently, the need for interdisciplinary project teams is a much-discussed requirement (see, e.g., Lee et al., 2015).

Earlier studies of eBPs have two things in common (see Table 1). First, previous studies were based almost exclusively on individual case studies and/or survey studies. No study has yet offered a broad analysis of multiple practical eBPs. Second, previous studies describe a fairly abstract picture of eBPs, focusing on the general meta-level. They mainly discuss characteristic challenges or recommend central success factors for managing eBPs. No study has yet provided a comprehensive picture of eBPs, in the sense of a full taxonomy including all thematic facets, specific problem areas, and tasks. This is an obvious gap in research, especially in view of the fact that such theoretical conceptualizations have existed for quite some time, in multiple variations, for other project types such as general information systems development projects (Keil et al., 2000; Xia & Lee, 2005) or enterprise resource planning (ERP)-system implementations (see, e.g., Somers & Nelson, 2004, Barki et al., 2005). Therefore, the current study attempts to close this gap by summarizing the conceptual nature of eBPs in a comprehensive theoretical conceptualization.

3. METHODOLOGY AND APPLICATION

3.1. Latent Semantic Analysis

Latent semantic analysis (LSA) will be used as the method for the exploratory evaluation of documented eBPs. LSA is an automated content analysis technique that allows comprehensive summaries of large volumes of texts (for an introduction, see, e.g., Deerwester et al., 1990; Dumais et al., 2004; Landauer et al., 1998). The basic idea behind LSA is that any document collection consists of certain contexts (e.g. individual documents, paragraphs, or sentences), as well as of words associated with these contexts. A structural analysis of these words, meaning their existence or non-existence in the individual contexts, allows the creation of statistical patterns of strongly-correlated words (i.e., words often used together in documents) as well as strongly-correlated contexts (such as documents that contain comparable sets of words). In this respect, the method of LSA is based on a factor analysis with the goal of reducing

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Related studies on e-business project characteristics

References	Research Focus and Contribution	Methods
Basu & Muylle (2007)	Development of an eBP planning process with a focus on established companies	Evaluation of the planning process in two practical eBPs
Burn & Ash (2001)	Development of a framework with 13 facilitators that lead to success of e-business change projects	Theoretical framework with evaluation in case studies with interviews
Chuang & Shaw (2005)	Development of an e-business implementation roadmap with key success criteria	Review of case descriptions, interviews, and survey study ($N = 43$)
Clarke & Doherty (2004)	Evaluation of eBP characteristics with a focus on IT-business relationships	Case study
Fong et al. (2002)	Characterization of the management of an eBP with phases and distinctive tasks	Review and summary of relevant literature
Kleist (2003)	Development of an approach for evaluating e-business information systems projects	Review and summary of relevant literature
Lee et al. (2007)	Exploration of the role of organizational capabilities in the context of successful e-business implementations	Survey study in Taiwanese firms ($N = 202$)
Lesjak & Vehovar (2005)	Exploration of factors that impact the evaluation of eBPs	Telephone survey among Slovenian companies ($N = 755$)
McLaughlin (2009)	Discussion of a failed eBP and recommendation of 10 success factors	Case study of a failed project with connected survey ($N = 346$)
Mehler-Bicher (2002)	Presentation of an evaluation approach for eBPs based on a profile scheme	Theoretical framework with evaluation in a practical case study
Phan (2002), (2003)	Discussion of eBP characteristics and recommendation of key success factors for developing and integrating e-business technologies	Case study at Intel Corp.
Phan (2009)	Discussion of eBP characteristics and recommendation of key success factors for agile developing and integrating e-business technologies	Examination of two case studies (Intel and Microsoft)
Phan & Vogel (2012)	Discussion of eBP characteristics and recommendations for a rapid e-business development project management with success factors	Case study at Intel Corp.
Shu-Hang et al. (2007)	Development of a requirement analysis pattern selection method for eBPs	Concept based on a literature study
Tiwana & McLean (2003)	Exploration of the uncertainty in eBPs with recommendations	Field study and survey study
van Grembergen & Amelinckx (2002)	Development of a generic e-business balanced scorecard for measuring and managing eBPs	Theoretical concept based on a literature study

the dimensionality of a large textual database through singular value decomposition (SVD). Basically, SVD works by separating the textual database into matrices with two groups of variables: on the one hand, eigenvalues of the individual contexts (i.e., documents, paragraphs, or sentences) and on the other, eigenvalues of individual words (i.e., terms). SVD then uses these values to calculate a set of factors which stand for groups of correlated words and correlated contexts with similar content. These extracted latent semantic factors, then, represent specific word usage patterns which can be interpreted as the thematic concepts inherent in the underlying document collection (Landauer, 2007; Valle-Lisboa & Mizraji, 2007). See Appendix B for a simplified illustration of the analytical procedure used in LSA.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In recent years, the application of automated content analysis techniques received growing attention in research for exploring large amounts of textual data (see Derbotoli et al., 2016). In particular, LSA offers this study three methodological advantages (see also Indulska et al., 2012; Evangelopoulos et al., 2012). First, it allows large collections of texts to be processed with a reasonable effort. Another advantage is the exploratory potential of LSA, since no manual, pre-defined analytic framework, for instance with pre-defined (a priori) variables or categories, is required. Third, the aforementioned advantages also help reduce negative human influence on the analysis, such as subjectivity or a lack of care in analytically coding texts. For these reasons, LSA is particularly well-suited for the purpose of this study: LSA facilitates an efficient, exploratory, and reliable identification of thematic concepts in a large collection of practical project reports.

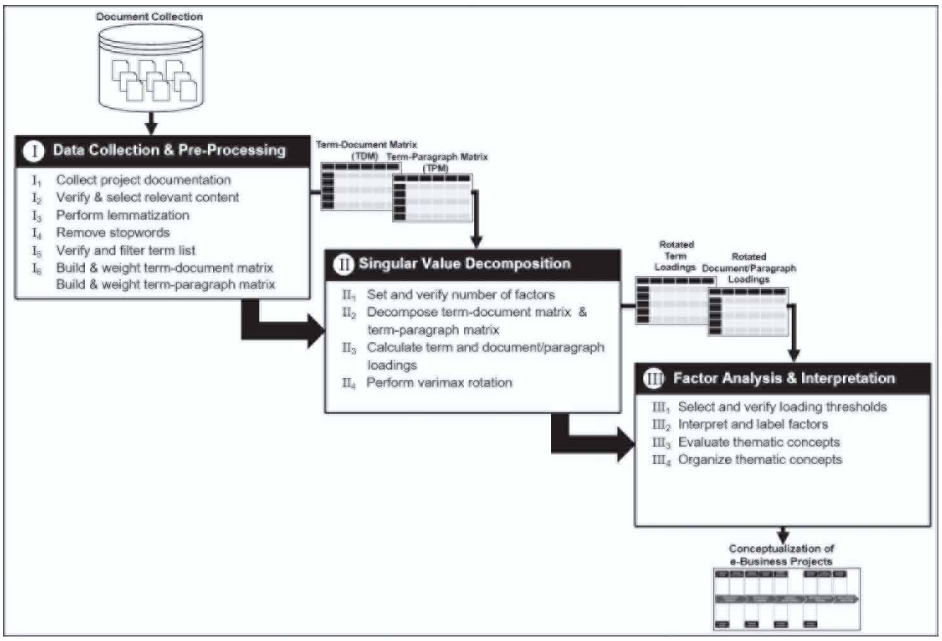
3.2. Data Source and Data Collection

This study uses practical project reports as a data source for exploring characteristic thematic concepts in eBPs. Project documentations are generally a rich source of textual codified project experiences (Disterer, 2002; Schindler & Eppler, 2003). In this context, such documentations usually describe five dimensions of project knowledge (see Zhao & Zuo, 2011): business domain knowledge (What characterizes the project from a business domain perspective?); project product knowledge (What characterizes the project with regard to the project outcome itself?); project engineering (technical) knowledge (What characterizes the project from a technical perspective?); organization management knowledge (What characterizes the project with regard to the management of organizations and relationships?); and project management knowledge (What characterizes the project with regard to its management?). Due to this comprehensive amount of codified knowledge, analysis of practical project reports is an approach which has already been used multiple times to explore thematic concepts (see, e.g., Choudhary et al., 2009; Herbst et al., 2014; Lormans et al., 2008). This approach is especially useful for examining a new phenomenon in practice for which little theoretical foundation currently exists (Benbasat et al., 1987).

This study uses a collection totaling 355 documentations of practical eBPs, compiled from collections of two specific initiatives. These are, first, the Sectoral e-Business Watch initiative[2] ($N = 271$), funded by the European Commission, which investigated the influence of e-business on companies, industries, and economies by assembling a comprehensive collection of practical project reports. Second, the study uses the database of the eXperience Online initiative[3] ($N = 84$), which also specializes in e-business and contains a comprehensive collection of project reports. These reports contain a retrospective summarization of the whole projects and should therefore represent a rich source of information for the purpose of this study. The reports cover a broad thematic range of eBPs, including, among others, e-commerce, e-procurement, and e-learning. The project reports generally follow a standardized format, describing the project background, the e-business solution implemented, the project procedure and success, and the lessons drawn from the project. The initiatives verified the quality, authenticity, and credibility of these documents. The reports were created between 2004 and 2012, mainly contain textual descriptions and vary between 4-13 pages written in English.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Analytical Process of the Latent Semantic Analysis



3.3. Analytical Process and Application

LSA is typically executed in a three-step process (see Müller et al., 2014; Sidorova et al., 2008): beginning with data collection and preparation, followed by statistical analysis of SVD, and ending in the final interpretation and definition of thematic concepts. This study completes two independent yet complementary analyses as a part of this process:

1. A more general analysis on a meta-level, which finally extracts a total of 12 thematic core concepts.
2. A more detailed analysis on a meso-level, which finally extracts a total of 74 specific sub-concepts.

The associated analytical process is summarized in the context diagram below (see Figure 1), and its application is explained more specifically in the following sections. All procedures described have been completed using Provalis Research´s QDA Miner and its text mining extension, WordStat 7.1.

3.3.1. Data Collection and Pre-Processing

During the first phase, an appropriate document collection was assembled and the textual database was prepared to derive an analyzable database.

(I_1) Collect project documentation: In the first step, a collection of documents potentially relevant for the purpose of analysis needed to be assembled. This study is based on a collection of 355 documented eBPs described in Section 3.2. All project reports were downloaded from the online databases and

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

transferred into a unified plain text format. No duplicates were found. Ultimately, a total of 2,019 pages (an average of 5.7 pages per report, min. 4 and max. 18 pages) were available.

(I_2) Verify & select relevant content: In the next step, relevant document contents were manually preselected and verified for relevance. During this review, irrelevant sections (e.g., company profiles, references, copyrights, annotation) were excluded from the following analysis, since irrelevant textual data could falsify the statistical calculations.

(I_3) Perform lemmatization: The specified document collection resulted in a database including a total of 18,308 unique words (or terms). To reduce the various forms of the same word, a lemmatization was conducted (as suggested by Evangelopoulos et al., 2012). This procedure unified grammatical variations of a word to their base or dictionary form (so-called lemmas) using comprehensive thesauruses (e.g., "analyzing," "analyzes," and "analyzed" to the base form, "analyze"). In this way the database was reduced to a term list of 13,341 remaining unique terms.

(I_4) Remove stopwords: In the next step, particular stopwords were removed from the database, meaning terms that do not carry any particular meaning and are therefore trivial (e.g., "and," "or," "in"). This reduced the term list to 11,584 remaining terms. For this purpose, the publicly available stopword list provided by Provalis Research was used. Additionally, terms with an occurrence frequency of fewer than 2 in the total document collection were excluded (as suggested by Sidorova et al., 2008), since such individual terms do not contribute to discovering patterns. In this way, the term list was further reduced to a total of 5,141 unique terms.

(I_5) Verify and filter term list: The term list was then manually reviewed in order to correct spelling errors, harmonize varying spellings, and filter out further irrelevant terms. This term filtering process is a critical aspect in the course of an LSA (Evangelopoulos et al., 2012). Here, two different term lists for two different analyses were specified (see Phase III). The first is a more comprehensive term list from which only additional trivial terms irrelevant for the topic to be analyzed were manually removed (e.g., "approximately," "obvious," "monthly," or "November"). Two independent analysts executed this filtering process in parallel after determining basic coding rules. In this case, special focus lay on nouns (e.g., "project"), verbs (e.g., "planning"), and adjectives (e.g., "expensive"). Pronouns, adverbs, and prepositions were removed due to their limited informative value (Baeza-Yates & Ribeiro-Neto, 1999). Subsequently, each filtered term list could be reviewed and compared for agreement in order to determine the inter-coder reliability. The agreement was 78%, indicating a generally acceptable level of reliability (Landis & Koch, 1977). The discovered differences were then discussed and corrected according to consensus. The resulting list of 2,423 unique terms formed the first database for the detailed analysis of thematic sub-concepts. The second term list was then reduced to an even more specific one, including a total of 806 central key terms. This term list was further specified through frequency filtering, meaning that only those terms were included which appeared in a minimum of 20% of the document collection (71 project reports). Iterative evaluations of the generated term lists and early analysis results were used for evaluating the suitability of this frequency threshold. The generated term list, including widespread, i.e., central key terms, formed the second database for the analysis of dominant core concepts within the document collection.

(I_6) Build & weight term-document/term-paragraph matrix: In the next step, a structured and quantitatively analyzable database could be created. This is typically a matrix in which the selected terms (in the rows), including their frequencies (in the cells), are organized in their respective contexts (in the columns). The segmentation of the context determines how the thematic concepts to be extracted

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

are distributed across the documents or document collection. If, for instance, the goal is to extract the dominant concepts within a whole document collection, it makes sense to segment the database according to complete documents. If, however, the goal is to identify more specific concepts – of which there may be many in an extensive document – it makes sense to use a finer segmentation based on individual paragraphs of the document. For this reason, both a term-document matrix (TDM) with 806 central terms as well as a term-paragraph matrix (TPM) with 2,423 terms were created for the two different analyses of this study. The TDM allowed the extraction of dominant core concepts (co-occurring terms) within the context of the overall document collection, while the TPM allowed the extraction of specific sub-concepts expressed within the context of individual document paragraphs (see also task III_4). The analysis of the TPM added specific insights to the study, since this analysis extracts co-occurrences of terms on paragraph level, thereby allowing the identification of more specific tasks and problems regularly discussed in specific passages of the reports.

TF-IDF indices were then calculated to create a more representative weighting of individual terms (see also Appendix B), in which the absolute frequency of a term in a document (TF = term frequency) was multiplied with the inverse frequency of the term in the overall document collection (IDF = inverse document frequency). This calculation therefore takes into account both the frequency of a term and its prevalence in the document collection, providing an index of representative weightings. Concretely, this means that the weight of very frequent, i.e., common, terms is reduced, and rarer terms are given more weight. The TDM and TPM weighted in this manner provided the analyzable database for the ensuing SVD in Phase II.

3.3.2. Singular Value Decomposition

During the second phase, the TDM and TPM were processed using the singular value decomposition (SVD) and specific factor loadings for individual terms and for their contexts (i.e., documents or paragraphs) were calculated.

(II_1) Set and verify number of factors: First, it was necessary to determine a specific number of latent semantic factors to be calculated. This means a determination of the dimensionality of the analysis was required. A small number of factors (a higher degree of aggregation) is more likely to bring general thematic concepts to light, while a larger number of factors (a lower degree of aggregation) will produce more detailed thematic concepts. However, it is generally difficult to determine an optimum number of factors. Here, Evangelopoulos et al. (2012) recommend the sensitive and iterative testing of alternative numbers of factors that result in a collection of meaningful concepts. Additionally, scree tests are suitable for reviewing the empirical foundation for the number of factors, in order to explain the variation included in a data set. As previously mentioned, this study is based on two different and complementary analyses. The first is a more general analysis, based on a TDM with 806 terms, intended to extract central core concepts. A 14-factor solution was evaluated for this purpose. The second is a more specific analysis, based on a TPM with 2,423 terms, intended to extract detailed sub-concepts. An 80-factor solution was evaluated for this purpose. In both analyses, iterative test runs determined that a larger number of factors would not reveal any

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

new meaningful concepts, and that a smaller number of factors would cause the loss of relevant concepts. Subsequent scree tests also confirmed the number of factors from an empirical perspective.

(II_2) Decompose term-document/term-paragraph matrix: Based on the factors determined, the TDM or TPM were divided into a unique matrix of term eigenvalues, a matrix of document or paragraph eigenvalues, and a diagonal matrix of singular values (see also Appendix B). These matrices formed the basis for determining individual term or document/paragraph loadings.

(II_3) Calculate term and document/paragraph loadings: The loadings for terms, documents, and/or paragraphs were then determined. Assigning specific high-loading terms and documents/paragraphs to individual factors allowed then the discovery of the specific thematic concepts in the document collection.

(II_4) Perform varimax rotation: To simplify the ensuing interpretation of latent semantic factors, in particular the composition of their associated high-loading terms, a varimax rotation was performed (as suggested by Sidorova et al., 2008). This is a statistical process for more strongly emphasizing variations of a term loading which simplifies the demarcation of high-loading terms or low-loading terms. The rotated term and document/paragraph factor loadings formed the basis for the ensuing analysis and interpretation of factors in Phase III.

3.3.3. Factor Analysis and Interpretation

In the third phase, specific high-loading terms for each factor were determined before two analysts interpreted the resulting bundle of words as thematic concepts and organized them into a comprehensive conceptualization.

(III_1) Select and verify loading thresholds: Before finally deriving the individual semantic factors, the threshold above which a term should be assigned to a factor had to be determined. The optimum level for this loading threshold is the subject of some discussion in the literature. With regard to this question, Evangelopoulos et al. (2012) recommend testing a varying threshold, then determining whether the resulting factors lead to meaningful results. Following this suggestion, loading thresholds of 0.58 (14-factor solution) and 0.21 (80-factor solution) were determined for the two analyses in this study.

(III_2) Interpret and label factors: The meanings of the individual latent semantic factors were then interpreted by reviewing their associated high-loading terms. Interpretations were completed in parallel by two independent analysts. The agreement of the respective interpretations showed an acceptable inter-coder reliability of 84%. Discrepancies were discussed and corrected by mutual consensus. In this context the analysts were able, for example, to interpret the correlated words "security," "transaction," "transmission," "data," and "protect," as the sub-concept, transaction security (F80_2). Another example is the core concept, e-business requirements (F14_7), composed, among others, from the high-loading terms "innovation," "requirement," "condition," "satisfaction," "quality," and "safety." In this process, the interpretation of unclear terms was supported by an accompanying keyword-in-context analysis (KWIC), which represents the words within their sentences, supporting an assessment of their meaning. Within the framework of the 14-factor solution, the analysts were able to identify a total of 12 central core concepts, since two factors seemed to duplicates and could be interpreted as the same topic (F14_1+3; F14_6+9). Within the frame-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

work of the 80-factor solution, the analysts were able to identify a total of 74 sub-concepts, since 4 concepts appeared twice (e.g., F80_9+31; F80_20+46), and 2 concepts could not be interpreted clearly (F80_28+40).

(III_3) Evaluate thematic concepts: The resulting concepts were then evaluated by an experienced domain expert for their meaningfulness and plausibility (as suggested by Evangelopoulos et al., 2012).

(III_4) Organize thematic concepts: The extracted concepts were then organized into a structured thematic conceptualization. The basic phases of a system development life cycle (SDLC; see Langer, 2004) provided the founding framework for this. These five phases delineate the typical context of an eBP, indicating the roles and interdependencies of the individual thematic concepts in a project process over time. First, the 12 core concepts were organized along these phases, providing main themes that characterize eBPs on a meta-level. In the following step, the specific 74 sub-concepts were then subordinated to these core concepts. These sub-concepts represent more specific tasks and problems on a meso-level. The analysts completed this organization using relevant literature related to the e-business discipline (see also Section 4). An illustrative example: the specific sub-concept communication channels (F80_53) was subordinated to the core concept e-business systems & technologies (F14_11) since it stands for a specific issue in the context of this main theme. Finally, the result was therefore a thematic conceptualization of eBPs, consisting of 5 project phases, 12 core concepts, and a total of 74 detailed sub-concepts (see also Figure 2 and Table 2 in Section 4).

4. RESULTS

4.1. Conceptualization of e-Business Projects

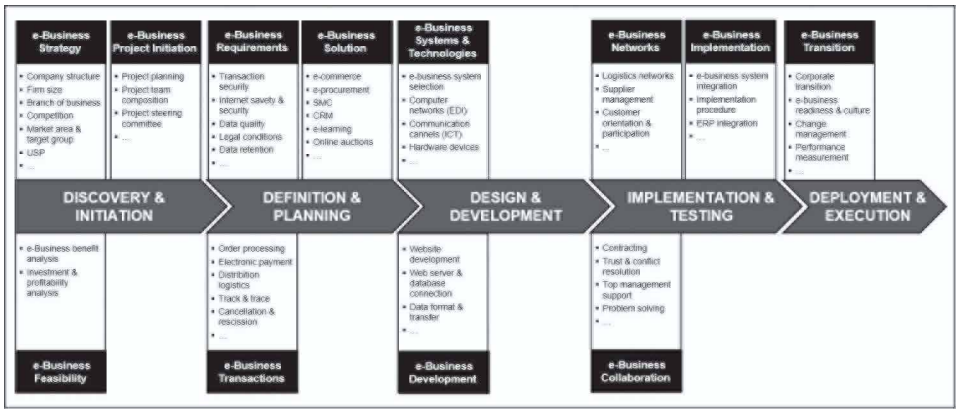
The following sections present and reflect on the developed conceptualization of eBPs (see Figure 2 and Table 2). This conceptualization describes the thematic complexity of eBPs on two different levels:

- 12 core concepts describe the central topical areas of eBPs.
- 74 detailed sub-concepts describe the detailed problem areas or tasks in eBPs.

Additionally, the basic phases of a system development life cycle (SDLC) provide structure and context. These phases of the SDLC are particularly appropriate for representing common eBPS, because they not only comprise technological areas of interest, but also organizational and business process re-engineering issues typically implied by such projects (Langer, 2004). The first project phase, Discover & Initiate, encompasses the underlying e-business strategy for eBPs, as well as the feasibility study on which project decisions are based. Furthermore, this phase involves the subsequent initiation of an eBP. Definition & Planning encompasses the conception of e-business solutions to be developed, the evaluation of associated requirements, as well as the analysis of involved business transactions, which may be either affected or created by the e-business solution. Design & Development encompasses the technical issues for realizing an e-business solution, such as an evaluation of existing e-business systems and technologies, as well as their ultimate technical conception and development. Implementation & Testing encompasses the implementation of e-business solutions in existing business operations, as well as their

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Conceptualization of e-Business Projects



coordination and harmonization with cooperating networks. Deployment & Execution encompasses the issues of an organizational transition into a functioning e-business, such as change management and the creation of a company-wide e-business culture.

The following sections will discuss more closely the five project phases and their respective thematic concepts. Moreover, related literature will be referred to (see also Table 3 in Appendix A) in order to reflect on the individual concepts and their meaning more theoretically, thereby deepening the understanding of eBPs.

4.2. Discovery and Initiation

During the first project phase, discovery & initiation, the purpose, goals, and scope of the project are defined. In particular, this phase is concerned foremost with delineating the company background relevant for decision making and for developing an e-business strategy (F14_12). An investigation of the advantageousness and practicability of the envisioned eBP is also part of this phase, in the sense of an e-business feasibility study (F14_8). Building on these considerations, the e-business project initiation (F14_4) can be conducted.

e-Business Strategy (F14_12): Transforming a company into a modern e-business can have a wide-reaching influence on its general business model as well as on its strategic corporate goals (Barua et al., 2001; Chen, 2001; Gloor, 2012). Therefore, an e-business solution should not be initiated blindly or without reflection; instead, it must be integrated into existing business practices in a controlled and directed manner. McLaughlin (2009, p. 48) summarizes this issue aptly: "[S]imply building an e-business system does not guarantee its success." This is why researchers in previous studies have regularly emphasized the importance of providing a profound e-business strategy to guide the e-business endeavors, a strategy which is aligned to the company's background and overall corporate goals (see, e.g., McLaughlin, 2009; Tiwana & McLean, 2003; van Grembergen & Amelinckx, 2002). Such a strategy further calls upon a company to clearly define what it wants to achieve through the project and – perhaps much more importantly – what it does not need to achieve (Phan & Vogel, 2012). In this respect, the e-business strategy is the basic framework which defines the project's purpose, goals, and scope.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The core concept, e-business strategy, encompasses such basic considerations of the company's situation, which may restrict the strategic framework of the e-business transformation in some instances. Several sub-concepts could be identified in this regard. For instance, the sub-concept, company size (F80_54), which may restrict the scope of an implemented e-business solution (Gengatharen & Standing, 2005; Taylor & Murphy, 2004). A company's structure (F80_12) may also influence the complexity of system and technology implementation (e.g., due to the number of locations and to decentralized functions). Further sub-concepts describe the general competitive situation, which should also be taken into consideration, including the branch of business (F80_59), the competitive environment (F80_7), as well as the market area and target group (F80_80). Finally, on the basis of such considerations, companies can then develop an appropriate e-business strategy (F80_37), as well as the unique selling proposition (F80_50) they aspire to.

e-Business Feasibility (F14_8): Companies should not integrate e-business solutions without critical reflection, but should weigh their advantages and disadvantages, opportunities, and risks carefully (Feng & Hong-Dan, 2010; Mehler-Bicher, 2002; Shu-Hang et al., 2007). On the one hand, the decision about e-business investments requires a rational evaluation of costs and benefits, and thus an objective consideration of the added value (Kleist, 2003; van Grembergen & Amelinckx, 2002). On the other hand, this not only concerns evaluating common economic indicators (revenue growth, for instance), but also evaluating IT-related aspects such as the continued development of operational efficiency (KLeist, 2003).

The core concept of e-business feasibility, therefore, encompasses an evaluation of the profitability of an investment into an e-business solution (F80_25). This means a review of whether the planned cost reductions or revenue growth justify the sometimes substantial investments in technology and organizational change (Kleist, 2003). Additionally, the sub-concept of e-business benefits (F80_24) was identified, describing the operational advantages, such as business process automation and efficiency increases (see, e.g., Ash & Burn, 2003), implied by the e-business solution.

e-Business Project Initiation (F14_4): On the basis of the antecedent considerations, the eBP can then be specified and initiated. The planning of eBPs involves, in particular, the concrete definition of a project assignment, meaning the goals the project is considered to achieve (Basu & Muylle, 2007). Furthermore, the responsibilities and allocated resources are defined (Chuang & Shaw, 2005; Fong et al., 2002). This also encompasses the formation of a project team, which generally should include professionals with interdisciplinary expertise (Lee et al., 2015).

The core concept, e-business project initiation, describes the aforementioned issues and encompasses related sub-concepts revealed, such as the initial planning of project goals, timeframes, costs (F80_47), the setting up of a project steering committee (F80_1), as well as the compilation of an appropriate operational project team (F80_30).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.3. Definition and Planning

The project phase, definition & planning, is focused foremost on making the e-business solution (F14_1+3) to be developed as concrete as possible. Additionally, this phase involves considerations concerning the related e-business requirements (F14_7) and e-business transactions (F14_5).

e-Business Solution (F14_1+3): The e-business solution to be developed and implemented forms the core of any eBP (Balakian et al., 2002; Stoehr, 2002). E-business solutions, in this context, can be described as means for the expansion of traditional business functions through various digital and network-based e-business functions (Shaw, 2003; Shaw, 2000; Wang et al., 2012). The core concept of e-business solutions (F14_1+3) encompasses different forms of such e-business functions, including e-commerce (F80_61) and online marketing (F80_31), e-procurement (F80_56), network-based supply chain management (F80_57), customer relationship management (F80_36), e-learning (F80_69), e-auctions (F80_48), production management (F80_46), mobile business (F80_72), and online media & entertainment (F80_16). Furthermore, several enabling or supporting tasks and functions of an e-business solution were identified as sub-concepts, such as a help desk and incident management (F80_64), a product & service portfolio designed for e-business (F80_8), and document management (F80_7) and license management (F80_51). Securing the intellectual rights to e-business solutions that have been developed (e.g., technical innovations or brands) is another concept revealed in the course of this study (F80_36).

e-Business Transactions (F14_5): e-business solutions facilitate and/or require the development of automated and inter-organizational transactions (Wang et al., 2012). These transactions, meaning especially their operational effects, need to be profoundly harmonized with existing transactions (Phan & Vogel, 2012). In this context, the core concept, e-business transactions, describes such automated business processes facilitated through e-business solutions, in particular through the following sub-concepts: order processing (F80_4+33), distribution logistics (F80_58), electronic payment (F80_14), cancellation & recession (F80_68), authentication processes of users and customers (F80_3+75), outbound and inbound transactions (F80_67), and tracking & tracing of consignments (F80_70).

e-Business Requirements (F14_7): The developers of e-business solutions may be confronted with several requirements, in technical, business-related, or also legal contexts (see, e.g., Gilioli-Rotondaro, 2002; Gordijn & Akkermans, 2001; Jones et al., 2000). The core concept, e-business requirements, summarizes such demands and expectations, relevant, for example, within the framework of a requirements analysis (see Moertini et al., 2014; Shu-Hang et al., 2007). In the course of this study, several sub-concepts could be identified with respect to such requirements, such as internet security and safety (F80_26+72), transaction security (F80_2), the encryption of information (F80_11), a user-friendly interface (F80_32), and data quality requirements (F80_63). Legal conditions (F80_43) may imply additional, sometimes restrictive demands in eBPs, for instance in relation to the requirements of internet fraud prevention (F80_34) or data retention regulations (F80_42).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.4. Design and Development

The project phase, design & development, describes technical aspects for facilitating an e-business solution. This includes considerations of both e-business systems & technologies (F14_11) and the technical realization of the envisioned concepts within the framework of an e-business development (F14_14).

e-Business Systems & Technologies (F14_11): From a technical standpoint, e-business is based on integrating information and communication technologies (ICT) and inter-organizational e-business systems (Wang et al., 2012). For the technical realization of the envisioned e-business functions, various solutions from different suppliers are available to companies. In this context, the core concept, e-business systems & technologies, represents a summarization of such technical solutions and involves a fundamental consideration of how to understand and evaluate this technical diversity. Several corresponding sub-concepts could be subsumed under this topic, such as the concepts describing e-business systems (F80_21), hardware & software technologies (F80_66), hardware devices (F80_22), and communication channels (F80_53). Considerations of specialized systems such as recommender systems (F80_45) may also play a role in this context.

e-Business Development (F14_14): Sooner or later, every eBP will include the technical development of e-business systems and technologies (see Phan, 2009; Phan &Vogel, 2012). In this respect, the core concept, e-business development, summarizes the issues associated with such a technical development of e-business solutions. This includes, for instance, developing websites (F80_49), connecting these to web servers and databanks (F80_5), integrating one's own systems into larger computer networks (F80_38), or developing company-specific business rules (F80_74). Handling and exchanging data – often large amounts of new data – are general challenges in eBPs (Wang et al., 2012; Zhao et al., 2007). This includes aspects of a standardized data format and data transfer (F80_29), data collection (F80_13), data replication and storage (F80_52), as well as real-time data exchange and data synchronization (F80_6).

4.5. Implementation and Testing

The project phase, implementation & testing, includes the operational introduction and start-up of e-business technologies and systems. Besides the actual e-business implementation (F14_10) into operating practices, this also requires interaction and coordination with a large number of stakeholders in e-business networks (F14_6+9), as well as the accompanying need for cooperative e-business collaboration (F14_13).

e-Business Networks (F14_6+9): Implementing e-business technologies opens up new forms of network-based cooperation. As Chang (2010) says: "Company value chains are transformed into horizontally integrated components" (p. 104). This means that e-business integrates mutually dependent partnerships along the whole value-creation chain. Thus, the very nature of e-business already implies that companies implementing an e-business solution have to interact with various stakeholders in such network-based environments (Clarke & Doherty, 2004; Zhao et al., 2007).

The core concept of an e-business network encompasses alliances and stakeholders with whom cooperation and coordination may be required in an eBP. For example, an eBP may require close collaboration with a company's suppliers (F80_15), in particular with e-procurement implementations (see Barua et

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

al., 2001; Chuang & Shaw, 2005; Krasner, 2000). Additionally, close coordination with larger logistics networks (F80_79) and their players may be necessary (see Boone & Ganeshan, 2007). Furthermore, customers are also important stakeholders (F80_18), and involving them early on is essential for ensuring that an eBP is consistently customer-oriented (Gilioli-Rotondaro, 2002).

e-Business Collaboration (F14_13): As already mentioned, implementing e-business solutions is always accompanied by intensive interaction with partners in inter-organizational alliances (Chang, 2010). Acting in these (external) partnerships, some of which may include heterogeneous or new stakeholders, can represent a complex organizational challenge in eBPs (Burn & Ash, 2001; Jackson & Harris, 2003). In this regard, Chang (2010) emphasizes the importance of an effective collaboration in such a dependent relationship, which requires, for example, the existence of mechanisms to facilitate a reliable collaboration, as well as the existence of mutual trust and personal commitment (see, e.g., OECD, 2005, for a description of a so-called "trust infrastructure" in e-business).

Against this backdrop, the core concept, e-business collaboration, encompasses the various tasks and capabilities required for acting in such a complex project environment. This includes sub-concepts describing an effective, prompt problem-solving (F80_79), the ability to overcome larger conflicts and to ensure binding cooperation based on trust (F80_10), and contracting (F80_76), which is an issue for determining binding obligations in a supplier cooperation (Boone & Ganeshan, 2007). Furthermore, an absolute support from the top management (F80_39) is another important prerequisite in such strategically-relevant cooperation, for instance in executing decisions (Phan, 2009; Phan & Vogel, 2012).

e-Business Implementation (F14_10): Implementing e-business systems and technologies implies operational consequences for business processes and the way employees work (see Burn & Ash, 2001; Jackson & Harris, 2003). The core concept, e-business implementation, encompasses such issues related to the well-planned introduction of new e-business solutions and the seamless transfer from traditional business functions to new and automated e-business functions. This core concept encompasses sub-concepts such as the operational integration of e-business systems (F80_20+46), the selected implementation procedure (F80_19), for example, piloting, as well as the harmonization with existing ERP-systems (F80_60). A good employee coordination (F80_44) is also an issue to mention in such complex organizational interventions, since this can often have a major influence on work content and processes.

4.6. Deployment and Execution

The project phase, deployment & execution, encompasses the consequences of the organizational transformation that companies are often confronted with after the implementation of e-business solutions, summarized in the core concept, e-business transition (F14_2).

e-Business Transition (F14_2): eBPs imply business-related and organizational transformations (see Barua et al., 2001; Burn & Ash, 2001; Gloor, 2012; Jackson & Harris, 2003). Kalakota et al. (1999) describes this aptly: "the creation and implementation of an e-business project is inextricably linked to the management of change" (p. 60). Therefore, an eBP does not just stop at a pure technical implementation of e-business systems and technologies, but also includes the sensitive transition

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

of traditional business practices into networked, automated business practices, and thus into a functioning e-business. In general, this results in the necessity to fundamentally rework prior business processes as part of a business process reengineering (Earl, 2000). This step may also include both questioning the breadth and depth of services offered by the company as well as externally pooling and potentially outsourcing company functions (Phan & Vogel, 2012). Of course, such fundamental structural and procedural upheavals require a high level of readiness for change from employees. As a result, e-business transitions require employees to be introduced to new processes and responsibilities so that they can acquire the knowledge they need (Daniel & Wilson, 2003; Lee et al., 2007; Tiwana & McLean, 2003). This is why the literature regularly emphasizes the importance of sound change management and trainings in the context of eBPs (see Burn & Ash, 2001). Finally, both measuring general e-business values and assessing the specific performance of e-business solutions are also frequently discussed challenges (see Kleist, 2003; Lesjak & Vehovar, 2005; van Grembergen & Amelinckx, 2002).

The core concept, e-business transition, encompasses a variety of issues involved in handling this type of organizational transformation. In this context, sub-concepts could be identified which describe the business process reengineering (F80_55) and the organizational transition into networked, automated business practices (F80_33), as well as considerations relating to the outsourcing of business functions (F80_65). Further sub-concepts describe the issues of an e-business culture and readiness (F80_27) and a change management, including employee training (F80_41), as well issues of measuring the attained performance of e-business (F80_78).

Table 2. Conceptualization of e-Business Projects

(I) DISCOVERY & INITIATION	(II) DEFINITION & PLANNING	(III) DESIGN & DEVELOPMENT	(IV) IMPLEMENTATION & TESTING	(V) DEPLOYMENT & EXECUTION
F.14_12 e-BUSINESS STRATEGY	**F.14_1+3 e-BUSINESS SOLUTION**	**F.14_11 e-BUSINESS SYSTEMS & TECHNOLOGIES**	**F.14_10 e-BUSINESS IMPLEMENTATION**	**F.14_2 e-BUSINESS TRANSITION**
F80_12 COMPANY STRUCTURE	F80_7 DOCUMENT MANAGEMENT	F80_21 e-BUSINESS SYSTEM SELECTION	F80_19 IMPLEMENTATION PROCEDURE	F80_27 e-BUSINESS CULTURE & READINESS
F80_37+45+2 e-BUSINESS STRATEGY	F80_8 PRODUCT & SERVICE PORTFOLIO	F80_22 HARDWARE DEVICES	F80_20 e-BUSINESS SYSTEM INTEGRATION	F80_33 CORPORATE TRANSITION
F80_54 FIRM SIZE	F80_9+31 ONLINE MARKETING	F80_45 RECOMMENDER SYSTEMS	F80_60 ERP-SYSTEM INTEGRATION	F80_41 CHANGE MANAGEMENT
F80_59 BRANCH OF BUSINESS	F80_16 ONLINE MEDIA & ENTERTAINMENT	F80_53 COMMUNICATION CHANNELS (ICT)	F80_44 STAFF COORDINATION	F80_55 BUSINESS PROCESS REENGINEERING
F80_77 COMPETITION	F80_35 INTELLECTUAL PROPERTY PROTECTION	F80_66 HARDWARE AND SOFTWARE TECHNOLOGIES	**F.14_6+9 e-BUSINESS NETWORKS**	F80_65 OUTSOURCING
F80_80 MARKET AREA & TARGET GROUP	F80_36 CUSTOMER RELATIONSHIP MANAGEMENT	**F.14_14 e-BUSINESS DEVELOPMENT**	F80_15 SUPPLIER MANAGEMENT	F80_78 PERFORMANCE MEASUREMENT
F80_50 UNIQUE SELLING PROPOSITION (USP)	F80_46 PRODUCTION MANAGEMENT	F80_5 WEB SERVER & DATABASE CONNECTION	F80_71 LOGISTICS NETWORKS	
F.14_8 e-BUSINESS FEASIBILITY	F80_56 e-PROCUREMENT	F80_6 REAL-TIME DATA EXCHANGE & SYNCHRONIZATION	F80_18 CUSTOMER ORIENTATION & PARTICIPATION	
F80_24 e-BUSINESS BENEFIT ANALYSIS	F80_57 SUPPLY CHAIN MANAGEMENT	F80_29 DATA FORMAT & DATA TRANSER (STANDARDIZATION)	**F.14_13 e-BUSINESS COLLABORATION**	
F80_25 INVESTMENT & PROFITABILITY ANALYSIS	F80_61 e-COMMERCE	F80_38 COMPUTER NETWORKS (EDI)	F80_10 TRUST & CONFLICT RESOLUTION	
F.14_4 e-BUSINESS PROJECT INITIATION	F80_51 LICENCE MANAGEMENT	F80_49 WEBSITE DEVELOPMENT	F80_23 EMERGENCY RESPONSE PLAN	
F80_1 PROJECT STEERING COMMITTEE	F80_64 HELP DESK & INCIDENT MANAGEMENT	F80_52 DATA REPLICATION & STORAGE	F80_76 CONTRACTING	
F80_30 PROJECT TEAM COMPOSITION	F80_48 ONLINE AUCTIONS	F80_13 DATA COLLECTION	F80_39 TOP MANAGEMENT SUPPORT	
F80_47 PROJECT PLANNING	F80_69 e-LEARNING	F80_74 BUSINESS RULE DEVELOPMENT	F80_79 PROBLEM SOLVING	
	F80_72 MOBILE BUSINESS			
	F.14_5 e-BUSINESS TRANSACTIONS			
	F80_3+75 AUTHENTIFICATION			
	F80_4+73 ORDER PROCESSING			
	F80_14 ELECTRONIC PAYMENT			
	F80_17 FRONT-END AND BACK-END INTERACTION			
	F80_58 DISTRIBUTION LOGISTICS			
	F80_67 OUTBOUND & INBOUND TRANSACTIONS			
	F80_68 CANCELLATION & RESCISSION			
	F80_70 TRACK & TRACE			
	F.14_7 e-BUSINESS REQUIREMENTS			
	F80_2 TRANSACTION SECRITY			
	F80_26 INTERNET SECURITY			
	F80_11 ENCRYPTION & INTEGRITY CHECK			
	F80_32 USER INTERFACE			
	F80_34 INTERNET FRAUD PREVENTION			
	F80_42 DATA RETENTION			
	F80_43 LEGAL CONDITIONS			
	F80_63 DATA QUALITY			

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

5. DISCUSSION

5.1. Contributions

This study provides two contributions: one methodological, and the other content-based. From a methodological perspective, this study is the first to use an examination of a large number of practical eBPs. The information potential of practical project documentations has been exploited for this purpose. For their explorative summary, latent semantic analysis (LSA) has proven itself an effective and efficient solution for extracting concrete thematic concepts from this extensive textual database.

From the content-related perspective, this study is the first to deliver a comprehensive thematic conceptualization that contributes to characterizing eBPs as a unique type of project. The dimensionality of this conceptualization, including 12 core concepts and 74 detailed sub-concepts, as well as its fundamental contextualization within the framework of a system development life cycle (SDLC), contributes to managing the thematic complexity in eBPs on various levels.

5.2. Implications for Practice

The conceptualization of eBPs can be of concrete practical use in three different ways. First, the conceptualization represents a comprehensive summary of the topics potentially relevant for planning eBPs. In particular, as part of a requirement analysis (see Mehler-Bicher, 2002; Shu-Hang et al., 2007), this type of thematic reference can provide structured leads on the problem fields and tasks that need to be taken into account. Secondly, the conceptualization provides a thematic structure for knowledge management within the framework of eBPs. This is true both for structured reflection on and documentation of finished projects in the sense of a post-project review (see Disterer, 2002), as well as for storing collected project knowledge in structured knowledge databases. In particular, the domain-specific knowledge of eBPs could be explicitly modeled in respective ontology-based databases, which for their part allow a storage and retrieval of specific knowledge fragments in computational form (Staab et al., 2001). This allows future projects to retrieve context-oriented experiences according to specific problem areas or tasks. Third, in its suggested form, the conceptualization provides categories and required keyword sets to implement an e-business specific dictionary, which can provide the basis for computer-supported text analyses (i.e., for automated coding and categorization of documents).

5.3. Implications for Research

Implications for research include, first, evaluating the results of this study in order to discover any incompleteness or misinterpretation in the conceptualization (see limitations in Section 5.4). This means, in particular, that individual concepts can still surely be expanded, corrected, or specified. In this regard, the suggested conceptualization could be reviewed, for instance, using the criteria clarity, coherence, extendibility, encoding bias, and commitment suggested by Gruber (1995). The composition of e-business networks (F14_6+9), for example, could certainly be expanded (e.g., through public service networks). Experimental applications in practical case studies, for instance, or analytical reviews using other collections of data (i.e., project documents) would be suitable means for this.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

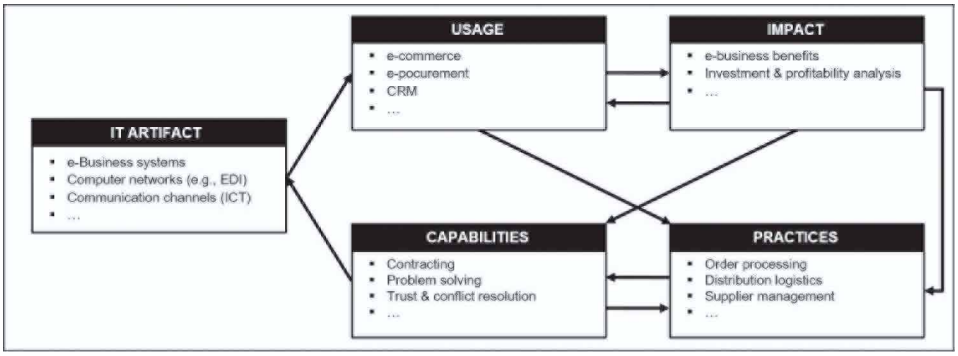
In addition, the developed conceptualization provides thematic leads for continuing theoretical investigations of eBPs that, for instance, could be useful for researching success factors. In particular, research perspectives could focus on investigating and describing correlations between the individual concepts. Potential interdependencies can already be assumed between multiple topics, such as the influence of legal conditions on data security and retention, or the connection between an organizational e-business culture and change management within the framework of the e-business transition. A deeper understanding of these interdependencies would contribute to the understanding of eBPs. The organization of thematic concepts into Benbasat and Zmud's (2003) nomological net depicted in Figure 3 provides one starting point for investigating these relationships. It represents relationships between the IT artifacts of an eBP (e.g., EDI), their application (e.g., e-procurement) and influence (e-business benefits), and their organizational practices (e.g., order processing) and capabilities (e.g., contracting).

5.4. Limitations

This study is of course not without its limitations. First, potential limitations with respect to the database used must be addressed. Knowledge codified in practical project reports may not always be objective and complete. One aspect frequently discussed in this context is the deliberate withholding of experiences (especially mistakes) out of fear of negative sanctions (see Cheng et al., 2009). This limitation may, however, be countered by the fact that the content to be described was standardized through the respective initiatives, and that the quality, authenticity, and credibility of the project reports was reviewed. Additionally, the document collection assembled for the study only represents a small sample of all eBPs actually completed in practice. Other collections of documents could potentially deliver different results. Nevertheless, the study does use a comparatively large number of documented projects from two different initiatives.

Second, despite the advantages discussed in Section 3.1, the methodological limitations of the LSA must be addressed. First, the LSA handles the textual database in an unreflected manner as a "bag of words," omitting a consideration of the content-based context of the words processed or, in other words, their actual meaning (Indulska et al., 2012). The LSA does partially address this problem by allowing the determination of the words' assumed meanings through the context of the bundles of associated words (e.g., the individual terms, "electronic," "business," and "strategy" suggest the topic of an "e-business strategy" in combination). Nevertheless, the interpretive depth of the LSA is limited, and it

Figure 3. Clustering of e-Business Project Themes



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

does leave questions of "why?," "how?," and "with what effect?" unanswered to a large extent: Why was an e-business strategy relevant, and what exactly does it look like? This study tries to make at least an initial attempt at answering these questions on the basis of a literature-based discussion. Nevertheless, future studies could build on it by offering a more exact definition and an explanation of these thematic concepts (e.g., by using case studies or literature reviews). One additional disadvantage of LSA is the diversity of analytical parameters to be determined (in particular a selection of the number of factors and determination of the loading threshold), which could have a significant influence on analysis results. For this reason, the suitability of the parameters used was subjected to thorough reviews.

Third, potential limitations result from the fact that LSA includes multiple manual procedures (i.e., manually filtering terms, interpreting and labeling thematic concepts), which require interpretation and could therefore involve a certain subjective bias. To ensure a reliable analysis nonetheless, such interpretive procedures were completed in parallel by two independent analysts, then reviewed for consistency.

6. CONCLUSION

The goal of this study was to illustrate the complexity of eBPs as a whole, providing a comprehensive summarization of all specific topics, problem areas, and tasks. In this context, the study aims at contributing to the characterization of e-business projects as a unique type of project. For this purpose, the study developed a thematic conceptualization including 12 core concepts and 74 specific sub-concepts, structured within the context of the typical phases of a system development life cycle (SDLC). The study provided general evidence that eBPs are complex endeavors whose thematic context does not end with a purely technical implementation of new technologies and systems, Instead, they rather introduce and accompany more extensive business-related and organizational changes.

First of all, this study provides a contribution to the e-business discipline regarding the practical management of eBPs, since project management strategies, processes, and practices generally need to adjust to the complexity of such endeavors. In this regard, the developed conceptualization represents a thematic reference for planning eBPs (e.g., in the context of requirement analyses) and for associated knowledge acquisition in the course of such projects. Additionally, this research provides a theoretical framework for further establishing eBPs as a unique project type. This includes in particular deeper explorations of associated characteristics, challenges, and success factors in such projects.

REFERENCES

Alter, S. (2001). *Information systems: Foundation of e-business* (4th ed.). NJ, USA: Prentice Hall PTR.

Amit, R., & Zott, C. (2001). Value creation in e-business. *Strategic Management Journal*, *22*(6-7), 493–520. doi:10.1002mj.187

Ash, C. G., & Burn, J. M. (2003). Assessing the benefits from e-business transformation through effective enterprise management. *European Journal of Information Systems*, *12*(4), 297–308. doi:10.1057/palgrave.ejis.3000476

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Baccarini, D. (1996). The concept of project complexity - a review. *International Journal of Project Management*, *14*(4), 201–204. doi:10.1016/0263-7863(95)00093-3

Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. England: Addison Wesley Longman Limited.

Balakian, W., Young, K., & Veerapaneni, R. (2002). *Managing e-Business Projects*. Bloomington, IN: PMP.

Barki, H., Oktamis, S., & Pinsonneault, A. (2005). Dimensions of ERP Implementations and their impact on ERP project outcomes. *Journal of Information Technology Management*, *16*(1), 1–9.

Barua, A., Konana, P., Whinston, A. B., & Yin, F. (2001). Managing e-business transformation: Opportunities and value assessment. *Sloan Management Review*, *43*(1), 36–44.

Basu, A., & Muylle, S. (2007). How to plan e-business initiatives in established companies. *MIT Sloan Management Review*, *49*(1), 28–36.

Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The case research strategy in studies of information systems. *Management Information Systems Quarterly*, *11*(3), 369–386. doi:10.2307/248684

Benbasat, I., & Zmud, R. W. (2003). The identity crisis within the IS discipline: Defining and communicating the discipline's core properties. *Management Information Systems Quarterly*, *27*(2), 183–194.

Boone, T., & Ganeshan, R. (2007). The frontiers of e-Business technology and supply chains. *Journal of Operations Management*, *25*(6), 1195–1198. doi:10.1016/j.jom.2007.02.002

Burn, J., & Ash, C. (2001). Managing Change for e-business Success. In *Proceedings of the 14th BLED Electronic Commerce Conference*, Slovenia.

Chang, H. L. (2010). A roadmap to adopting emerging technology in e-business: An empirical study. *Information Systems and E-Business Management*, *8*(2), 103–130. doi:10.100710257-009-0111-y

Chen, S. (2001). *Strategic management of e-business*. New York, NY, USA: John Wiley & Sons, Inc.

Cheng, M. M., Schulz, A. K. D., & Booth, P. (2009). Knowledge transfer in project reviews: The effect of self-justification bias and moral hazard. *Accounting and Finance*, *49*(1), 75–93. doi:10.1111/j.1467-629X.2008.00271.x

Choudhary, A. K., Oluikpe, P. I., Harding, J. A., & Carrillo, P. M. (2009). The needs and benefits of Text Mining applications on Post-Project Reviews. *Computers in Industry*, *60*(9), 728–740. doi:10.1016/j.compind.2009.05.006

Chuang, M. L., & Shaw, W. H. (2005). A roadmap for e-business implementation. *Engineering Management Journal*, *17*(2), 3–13. doi:10.1080/10429247.2005.11415281

Clarke, S., & Doherty, N. (2004). The importance of a strong business-IT relationship for the realisation of benefits in e-business projects: The experiences of Egg. *Qualitative Market Research: An International Journal*, *7*(1), 58–66. doi:10.1108/13522750410512886

Combe, C. (2012). *Introduction to e-business*. Oxford, UK: Elsevier Ltd.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Crawford, C., & Ferguson, G. (1970). A General Rotation Criterion and its Use in Orthogonal Rotation. *Psychometrika*, *35*(3), 321–332. doi:10.1007/BF02310792

Daniel, E. M., & Wilson, H. N. (2003). The role of dynamic capabilities in e-business transformation. *European Journal of Information Systems*, *12*(4), 282–296. doi:10.1057/palgrave.ejis.3000478

Debortoli, S., Junglas, I., Müller, O., & vom Brocke, J. (2016). Text mining for information systems researchers: an annotated topic modeling tutorial. *Communications of the Association for Information Systems*, *39*(1), 7.

Deerwester, S., Dumais, S., Furnas, G., Landauer, T., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American Society for Information Science*, *41*(6), 391–407. doi:10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASI1>3.0.CO;2-9

Deise, M. V., Wright, A., & Nowikow, C. (2000). *Executive's guide to e-business: from tactics to strategy*. New York, NY: John Wiley & Sons, Inc.

Disterer, G. (2002). Management of project knowledge and experiences. *Journal of Knowledge Management*, *6*(5), 512–520. doi:10.1108/13673270210450450

Dumais, S. (2004). Latent Semantic Analysis. *Annual Review of Information Science & Technology*, *38*, 189–230.

Earl, M. J. (2000). Evolving the E-Business. *Business Strategy Review*, *11*(2), 33–38. doi:10.1111/1467-8616.00135

Evangelopoulos, N., Zhang, X., & Prybutok, V. R. (2012). Latent semantic analysis: Five methodological recommendations. *European Journal of Information Systems*, *21*(1), 70–86. doi:10.1057/ejis.2010.61

Feng, O., & Hong-Dan, Z. (2010). Study on the Feasibility of the e-Business Model. In *Proceedings of the International Conference on E-Business and E-Government* (ICEE) (pp. 369-372). 10.1109/ICEE.2010.101

Fong, B. C., Fong, A. C. M., & Chan, C. H. (2002). E-Business Project Management: Managing the Planning Process. *International Journal of the Computer, the Internet and Management, 10*(1), 31-41.

Gengatharen, D. E., & Standing, C. (2005). A framework to assess the factors affecting success or failure of the implementation of government-supported regional e-marketplaces for SMEs. *European Journal of Information Systems*, *14*(4), 417–433. doi:10.1057/palgrave.ejis.3000551

Gilioli-Rotondaro, R. (2002). Defining the customers expectations in e-business. *Industrial Management & Data Systems*, *102*(9), 476–482. doi:10.1108/02635570210450154

Gloor, P. (2012). *Making the e-business Transformation*. London, UK: Springer Science & Business Media.

Gordijn, J., & Akkermans, H. (2001). Designing and evaluating e-business models. *IEEE Intelligent Systems*, *4*(4), 11–17. doi:10.1109/5254.941353

Gregory, R. W., & Piccinini, E. (2013). The Nature Of Complexity In Is Projects And Programmes. In *Proceedings of the 21st European Conference on Information Systems (ECIS)*, Utrecht, Netherlands.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Gruber, T. R. (1995). Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies*, *43*(5), 907–928. doi:10.1006/ijhc.1995.1081

Herbst, A., Simons, A., vom Brocke, J., Müller, O., Debortoli, S., & Vakulenko, S. (2014). Identifying and Characterizing Topics in Enterprise Content Management: A Latent Semantic Analysis of Vendor Case Studies. In *Proceedings of the 22nd European Conference on Information Systems (ECIS)*, Tel Aviv, Israel.

Indulska, M., Hovorka, D. S., & Recker, J. (2012). Quantitative approaches to content analysis: Identifying conceptual drift across publication outlets. *European Journal of Information Systems*, *21*(1), 49–69. doi:10.1057/ejis.2011.37

Jackson, P., & Harris, L. (2003). E-business and organisational change: Reconciling traditional values with business transformation. *Journal of Organizational Change Management*, *16*(5), 497–511. doi:10.1108/09534810310494900

Jones, S., Wilikens, M., Morris, P., & Masera, M. (2000). Trust requirements in e-business. *Communications of the ACM*, *43*(12), 81–87. doi:10.1145/355112.355128

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, *23*(3), 187–200. doi:10.1007/BF02289233

Kalakota, R., & Robinson, M. (1999). e-Business: Roadmap for Success, MA, USA: Addison- Wesley Longman.

Keil, M., Mann, J., & Rai, A. (2000). Why software projects escalate: An empirical analysis and test of four theoretical models. *Management Information Systems Quarterly*, *24*(4), 631–664. doi:10.2307/3250950

Kleist, V. F. (2003). An approach to evaluating e-business information systems projects. *Information Systems Frontiers*, *5*(3), 249–263. doi:10.1023/A:1025693109442

Krasner, H. (2000). Ensuring e-business success by learning from ERP failures. *IT Professional*, *2*(1), 22–27. doi:10.1109/6294.819935

Landauer, T. K. (2007). LSA as a theory of meaning. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 3–32). Mahwah, NJ, USA: Lawrence Erlbaum.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2-3), 259–284. doi:10.1080/01638539809545028

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. doi:10.2307/2529310 PMID:843571

Langer, A. M. (2008). *Analysis and Design of Information Systems* (3rd ed.). London, UK: Springer.

Lee, C. K., Cha, K. J., & Park, B. (2015). Identifying Required Job Skills of E-Business Project Managers Based on Corporate Homepages and Job Search Websites. *International Journal of Software Engineering and Its Applications*, *9*(5), 75–84. doi:10.14257/ijseia.2015.9.5.08

Lee, C. P., Lee, G. G., & Lin, H. F. (2007). The role of organizational capabilities in successful e-business implementation. *Business Process Management Journal*, *13*(5), 677–693. doi:10.1108/14637150710823156

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lesjak, D., & Vehovar, V. (2005). Factors affecting evaluation of e-business projects. *Industrial Management & Data Systems*, *105*(4), 409–428. doi:10.1108/02635570510592334

Lientz, B., & Rea, K. (2009). *Start Right in E-Business*. San Diego, CA: Academic Press.

Lormans, M., van Deursen, A., & Gross, H. G. (2008). An industrial case study in reconstructing requirements views. *Empirical Software Engineering*, *13*(6), 727–760. doi:10.100710664-008-9078-4

Marca, D. A. (2006). Project Management of e-Business Initiatives. In *Proceedings of the International Conference on e-Business*, Setugal, Portugal.

Martin, D., & Berry, M. (2007). Mathematical foundations behind latent semantic analysis. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 33–57). Mahwah, NJ, USA: Lawrence Erlbaum.

McLaughlin, S. (2009). The imperatives of e-business: Case study of a failed project. *The Journal of Business Strategy*, *30*(1), 40–49. doi:10.1108/02756660910926966

Mehler-Bicher, A. (2002). Evaluating e-business investments by defining profiles. In *Proceedings of the 9th European Conference on IT Evaluation* (pp. 317-324).

Moertini, V. S., Heriyanto, S., & Nugroho, C. D. (2014). Requirement analysis method of e-commerce websites development for small-medium enterprises. *International Journal of Software Engineering & Applications*, *5*(2), 11. doi:10.5121/ijsea.2014.5202

Moslehi, A., Linger, H., & Tanner, K. (2011). Towards a KM framework for e-business projects within SMEs. In *Proceedings of the 17th Americas Conference on Information Systems (AMCIS)*, Detroit, Michigan, USA.

Müller, O., Schmiedel, T., Gorbacheva, E., & vom Brocke, J. (2016). Towards a typology of business process management professionals: Identifying patterns of competences through latent semantic analysis. *Enterprise Information Systems*, *10*(1), 50–80. doi:10.1080/17517575.2014.923514

OECD. (2005). *ICT, E-Business and SMEs. OECD Digital Economy Papers*. Paris: OECD Publishing.

Pateli, A. G., & Giaglis, G. M. (2004). A research framework for analysing eBusiness models. *European Journal of Information Systems*, *13*(4), 302–314. doi:10.1057/palgrave.ejis.3000513

Phan, D. D. (2002). E-business success at Intel: An organization ecology and resource dependence perspective. *Industrial Management & Data Systems*, *102*(4), 211–217. doi:10.1108/02635570210423253

Phan, D. D. (2003). E-business development for competitive advantages: A case study. *Information & Management*, *40*(6), 581–590. doi:10.1016/S0378-7206(02)00089-7

Phan, D. D. (2009). Managing Agile E-Business Development: How Intel and Microsoft did it? In *Proceedings of the XIX ACME International Conference of Pacific Rim Management,* New York, USA (pp. 390-403).

Phan, D. D., & Vogel, D. (2012). Rapid E-Business Development Project Management at Intel. *International Journal of E-Business Development*, *2*(3), 97–102.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Schindler, M., & Eppler, M. J. (2003). Harvesting project knowledge: A review of project learning methods and success factors. *International Journal of Project Management*, *21*(3), 219–228. doi:10.1016/S0263-7863(02)00096-0

Shaw, M. J. (2000). Building an e-business from enterprise systems. *Information Systems Frontiers*, *2*(1), 7–17. doi:10.1023/A:1010037718452

Shaw, M. J. (2003). *E-business management*. Dordrecht: Kluwer.

Shu-Hang, G., Yu-Qing, L., Mao-Zhong, J., Jing, G., & Hong-Juan, L. (2007). A requirement analysis pattern selection method for e-business project situation. In *Proceedings of the IEEE International Conference on e-Business Engineering (ICEBE)* (pp. 347-350). 10.1109/ICEBE.2007.46

Sidorova, A., Evangelopoulos, N., Valacich, J. S., & Ramakrishnan, T. (2008). Uncovering the intellectual core of the information systems discipline. *Management Information Systems Quarterly*, *32*(3), 467–482.

Somers, T. M., & Nelson, K. G. (2004). A taxonomy of players and activities across the ERP project life cycle. *Information & Management*, *41*(3), 257–278. doi:10.1016/S0378-7206(03)00023-5

Srinivasan, S. (2004). Role of trust in e-business success. *Information Management & Computer Security*, *12*(1), 66–72. doi:10.1108/09685220410518838

Staab, S., Studer, R., Schnurr, H. P., & Sure, Y. (2001). Knowledge processes and ontologies. *IEEE Intelligent Systems*, *16*(1), 26–34. doi:10.1109/5254.912382

Stoehr, T. (2002). *Managing e-business Projects: 99 Key Success Factors*. Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-56090-3

Taylor, M., & Murphy, A. (2004). SMEs and e-business. *Journal of Small Business and Enterprise Development*, *11*(3), 280–289. doi:10.1108/14626000410551546

Tiwana, A., & McLean, E. R. (2003). The tightrope to e-business project success. *Communications of the ACM*, *46*(12), 345–350. doi:10.1145/953460.953520

Valle-Lisboa, J. C., & Mizraji, E. (2007). The uncovering of hidden structures by latent semantic analysis. *Information Sciences*, *177*(19), 4122–4147. doi:10.1016/j.ins.2007.04.007

van Grembergen, W., & Amelinckx, I. (2002). Measuring and managing e-business projects through the balanced scorecard. In *Proceedings of the 35th Annual Hawaii International Conference on System Sciences (HICSS)*, Hawaii, USA. 10.1109/HICSS.2002.994395

Wang, S., Li, L., Wang, K., & Jones, J. D. (2012). e-Business systems integration: A systems perspective. *Information Technology Management*, *13*(4), 233–249. doi:10.100710799-012-0119-8

Winter, M., Smith, C., Morris, P., & Cicmil, S. (2006). Directions for future research in project management: The main findings of a UK government-funded research network. *International Journal of Project Management*, *24*(8), 638–649. doi:10.1016/j.ijproman.2006.08.009

Xia, W., & Lee, G. (2005). Complexity of Information Systems Development Projects: Conceptualization and Measurement Development. *Journal of Management Information Systems*, *22*(1), 45–83. doi:10.1080/07421222.2003.11045831

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Zhao, D., & Zuo, M. (2011). Knowledge Transferred across Projects and Its Characteristics in an IT Service Enterprise. In *Proceedings of International Conference on Information Technology, Computer Engineering and Management Sciences (ICM)*, Nanjing, Jiangsu, China (pp. 267-270). 10.1109/ICM.2011.139

Zhao, K., Xia, M., & Shaw, M. J. (2007). An integrated model of consortium-based e-business standardization: Collaborative development and adoption with network externalities. *Journal of Management Information Systems*, *23*(4), 247–271. doi:10.2753/MIS0742-1222230411

ENDNOTES

1. In order to identify relevant studies specifically focusing on eBPs, a keywords search in three literature databases (Ebsco BSC, Wiley Online Library, Emerald Insight) was conducted, using search terms such as "e-business" (and other terms related to sub-topics) and terms related to project management combined by a logical AND (search fields: title, abstract, and keywords). Potentially relevant articles have been selected by reviewing titles and abstracts. However, it should be noted that this literature review, with its specific databases and search criteria, cannot find every potentially relevant paper.
2. See http://ec.europa.eu/enterprise/archives/e-business-watch/ for a presentation of the initiative.
3. See http://www.experience-online.ch/cases/experience20.nsf/en/index/ for a presentation of the initiative.

This research was previously published in the Journal of Electronic Commerce in Organizations (JECO), 15(3); edited by Pedro Isaías; pages 33-63; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

APPENDIX A

Results of The Latent Semantic Analysis

Table 3. Core Concepts Derived from the 14-Factor Solution

No.	Label	High-Loading Terms (excerpt)	Related Literature
F14_1	e-business solution	Web; concept; basis; online; sell; crm; service; launch; shop; account; site; server; retail; supply; link; transfer; turnover; mail; share; set	Balakian et al. (2002); Shaw (2003); Stoehr (2002); Wang et al. (2012)
F14_2	e-business transition	Structure; resource; innovation; corporation; site; firm; team; task; future; report; optimize; skill; motivation; responsibility; outsource; sap; maintenance; component; background; external; contract	Barua et al. (2001); Burn & Ash (2001); Jackson & Harris (2003)
F14_3	e-business solution	Business; electronic; concept; basis; design; corporation; internet; service; firm; delivery; commerce; distribution; demand; decision; development; service	Balakian et al. (2002); Shaw (2003); Stoehr (2002); Wang et al. (2012)
F14_4	e-business project initiation	Team; project; task; responsibility; specialist; resource; size; plan; name; specification; report; detail; experience; office; establish; structure; staff; potential	Basu & Muylle (2007); Chuang & Shaw (2005); Fong et al. (2002)
F14_5	e-business transactions	Sell; share; transaction; transfer; link; internet; market; customer; mail; account; external; user; relationship; staff; launch; server; unit; standard; confidential; basis	Wang et al. (2012)
F14_6	e-business networks	Internet; concept; connection; account; service; coordination; set; basis; infrastructure; process; server; international; global; logistics; launch; level; learn; location; country; effective	Clarke & Doherty (2004); Zhao et al. (2007)
F14_7	e-business requirements	Innovation; size; risk; customer; requirement; condition; schedule; satisfaction; access; cost; creation; security; role; quality; safety; availability; storage; optimize; payment; responsibility	Moertini et al. (2014); Shu-Hang et al. (2007)
F14_8	e-business feasibility	Benefit; company; chance; investment; potential, euro; business; innovation; advance; aim; administration; client; area; approach; analysis; reduction; application	Feng & Hong-Dan (2010); Kleist (2003); Lesjak & Vehovar (2005); Mehler-Bicher (2002)
F14_9	e-business networks	Background; connect; computer; supplier; network; access; chain; member; link; EDI; record; control; automation; catalogue; confirmation; geographical; area; environment	Clarke & Doherty (2004); Zhao et al. (2007)
F14_10	e-business implementation	Innovation; test; pilot; installation; component; learn; manage; implementation; department; challenge; personnel; procedure; structure; connection; method; optimize; form; maintenance; ERP	Chuang & Shaw (2005); Lee et al. (2007)
F14_11	e-business systems & technologies	Tool; standard; technology; definition; software; system; application; source; basis; database; ERP; solution; resource; enterprise; technical; shop; small	Chang (2010)
F14_12	e-business strategy	Corporation; business; background; firm; challenge; decision; management; target; turnover; strategy; goal; financial; budget; aim; market; competition	Amit & Zott (2001); Chen (2001); McLaughlin (2009); Phan & Vogel (2012)
F14_13	e-business collaboration	Partnership; market; participant; contract; definition; condition; coordination; logistics; regulation; responsibility; mail; negotiation; location; link; help; priority; convince; trust; requirement; reduction	Clarke & Doherty (2004); Zhao et al. (2007)
F14_14	e-business development	Component; development; equipment; entry; investment; configuration; exchange; program; connection; database; technical; specialist; expertise; expense; installation; format; interface; document	Phan (2009); Phan & Vogel (2012)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Sub-Concepts Derived from the 80-Factor Solution

No.	Label	High-Loading Terms (excerpt)
F80_1	project steering committee	committee; steering; project; team; responsibility; department; delegate; hire; task; role; people; employee; technician; interdisciplinary; play
F80_2	transaction security	security; protect; secure; safety; transmission; data; transaction; insecurity; feeling; confidential; authorization; encode; exchange; virtual; ssl; network; guarantee; ensure; access; client; private
F80_3	authentification	assertion; identity; certainty; verification; legislation; claim; cautious; authentication; valid; freely; infringe; login
F80_4	order processing	order; incoming; subsystem; receipt; desk; goods; browser; delivery; markup; confirmation; imminent; advise; outgoing
F80_5	web server & database connection	server; SQL; Microsoft; firewall; database; Linux; https; script; web; access; Unix; PHP; interface; VPN; host; connection
F80_6	real-time data exchange & synchronization	Realtime; synchronization; data; clever; GPRS; emulate; technician; device; offline; mobile; byte; terminate; update; head; combination; exchange
F80_7	document management	acrobat; adobe; document; header; replicate; reader; download; store; photocopy; postcard; official; letter; desktop; index; publication
F80_8	product & service portfolio	range; offer; service; product; customer; price; broad; shop; article; special; internet; catalogue
F80_9	online marketing	advertising; campaign; internet; generator; marketing; promotion; dealer; specialist; logo; magazine; poster; invitation; sales; newspaper; agency
F80_10	trust & conflict resolution	conflict; resolve; skepticism; disappear; surface; earn; trust; opening; lose; attractive; affiliate; balance; realization; commitment; attract; coordination; overcome
F80_11	encryption & integrity check	encrypt; integrity; encryption; signature; certificate; https; authenticity; certification; protocol; SSL; configurable; firewall; secure; interval; root
F80_12	company structure	HQ; branch; location; main; headquarter
F80_13	data collection	scanner; data; assemble; convey; frame; barcode; reasonable; dock; zone; compromise; arrive; deposit; label
F80_14	electronic payment	payment; credit; card; paybox; cancellation; pay; authorization; cash; mobile; phone; debit; repayment; accumulate; extranet; transaction
F80_15	supplier management	purchase; supplier; procurement; negotiation; buyer; association; catalogue; conform; vendor; seller; volume; article; agreement; price; cooperation; department; decentralize
F80_16	online media & entertainment	game; unlimited; DVD; CD; freely; library; book; keyword; author; bookshop; video; discover; entertainment; poster; protective
F80_17	front-end and back-end interaction	frontend; backend; communication; compensation; registration; notification; person; insure; Intershop; reproduce
F80_18	customer orientation & participation	customer; convince; indicate; development; tailor; involve; response; skeptical; oppose; evidence; technician; pilot
F80_19	implementation procedure	phase; pilot; project; function; involve; test; smooth; ensure; transition; change; initial; positive; early; automate; implementation
F80_20	e-business system integration	integration; shortcoming; technology; online; chance; determination; compatibility; incompatible; insufficient; fail; council; attempt; merge; coordination; intervention
F80_21	e-business system selection	system; solution; development; RFQ; vendor; negotiation; cost; proceeding; subject; source; logical; collective
F80_22	hardware devices	toner; printer; accessory; peripheral; portable; multifunctional; camera; fax; pc; scanner; range; breadth
F80_23	emergency response plan	emergency; dispatcher; intervention; failure; accelerate; response; resolve

continues on following page

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Continued

No.	Label	High-Loading Terms (excerpt)
F80_24	e-business benefit analysis	cost; reduce; save; reduction; time; error; low; efficient; postage; increase; improve; productivity; quality; high; significant; process
F80_25	investment & profitability analysis	quantify; benefit; difficult; profitability; raise; investment; turnover; satisfaction; increase; financial; capex; target; productivity; calculation; question; awareness
F80_26	internet security	virus; trojan; ISDN; inspection; protective; modem; domain; legal; hacker; script
F80_27	e-business culture & readiness	staff; questionable; enthusiastic; employee; knowledge; skepticism; apparent; discover; serious; notification; offset; culture; ict; workload
F80_28	-	dependency; clever; century; fail; apparent; commodity; accommodate; rely; Realtime; economy; emulate
F80_29	data format & data transer standardization	format; xml; file; ascii; convert; data; ftp; csv; edi; electronic; document; excel; upload; standard; transfer; interchange
F80_30	project team composition	member; team; staff; training; employee; consultant; skill; project; technician; competency; recommendation
F80_31	online marketing	advertisement; search; engine; animation; banner; recipe; google; unlike; disadvantage; offline; affiliate; play; zone; role
F80_32	user interface	interface, user; friendly; accept
F80_33	transition process	smooth; ensure; transition; change; structure; control; coordination; date; process; guarantee; automate; fully; cause; lose; minimize; method
F80_34	internet fraud prevention	fraud; abuse; manipulation; scene; doubt; aggressive; author; reasonable; advertisement; chip; attempt; live; count
F80_35	intellectual property protection	property; intellectual; protect; copyright; trademark; denote; interpretation; patent; sensitive; object
F80_36	customer relationship management	customer; relationship; management; satisfaction; service; primary; private; loyal; consumer
F80_37	e-business strategy	e-business; solution; strategy; commerce; procurement; success; factor; activity; supply; shop
F80_38	computer networks	network; interconnect; EDI; data; programmable; logic; gate; array; semiconductor; simulate; controller
F80_39	top management support	chief; officer; executive; declare; director; contact; CIO; reinforce; manage; revise; merge
F80_40	-	definable; redefine; disruption; root; sustain; distinctive; dominate; freely; resemble
F80_41	change management	employ; people; change; process; compatibility; chance; training; management; revelation; merge; competency; skill; decentralize; merger; fair; cultivate; sustain
F80_42	data retention	deletion; delete; data; bookkeeping; cancel; story; correction; absence; careful; stage
F80_43	legal conditions	federation; authority; government; legislation; local; council; legal; regulation; administration; incompatible; resident; department
F80_44	staff coordination	staff; coordination; urgent; compromise; perception; accordance; skill; dispatch; technician; assign; vacation; job; decision
F80_45	recommender systems	recommendation; rating; calculation; compute; similarity; filter; formula; numeric; personalization; suitable
F80_46	production management	production; control; accounting; capacity; machine; management; calculation
F80_47	project planning	endeavour; project; manager; coordination; responsibility; management; organization; department; task; division; realization; restriction; assume; employee; budget; technical; early; initiation; stage
F80_48	online auctions	auction; submission; bid; customer; expire; stipulate; offer; opening; online; favourable; letter; tender; reject; plausibility; public; valid

continues on following page

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Continued

No.	Label	High-Loading Terms (excerpt)
F80_49	website development	www; http; html; website; index; source; URL; PHP
F80_50	unique selling proposition	unique; sell; proposition; homogenous; custom; practical; role; side; conversation
F80_51	licence management	fee; pay; licence; charge; attention; commission; member
F80_52	data replication & storage	data; replication; replicate; middleware; interval; object; cd; store; backup; accordance; minimize; copy; tape; synthetic; disk; compose; redundant; aggregation
F80_53	communication channels	fax; telephone; mail; phone; call; order; email
F80_54	firm size	small; size; large; company; scale; number; employee; enterprise; medium; volume; economy
F80_55	business process reengineering	process; reengineering; adapt; sequence; change; evaluate; redesign; consecutive; interpret; objective; setting; task; variable
F80_56	e-procurement	inventory; stock; procurement; online; manage; warehouse; tender; level; control; quantity; management; deliver
F80_57	supply chain management	chain; supply; value; management
F80_58	distribution logistics	logistics; transport; service; provider; delivery; warehouse; centre; shipment; deliver; job; post; sales; distribution; goods; status; outsource
F80_59	branch of business	industry; automotive; health; target; sector; group; branch; association; construction
F80_60	erp-system integration	sap; ERP; system; resource; planning; enterprise; Mysap; integrate; solution; software; Netweaver; data; process; supply; chain; management
F80_61	e-commerce	shop; online; shopping; catalogue; commerce; book; website; click; retail; customer; article
F80_62	e-business strategy	strategy; corporation; vision; future; plan; pursue; goal; decision; follow; strategic; time
F80_63	data quality	data; incorrect; incomplete; simulate; recognition; danger; mistake; handwriting; identical; stipulate; correct; factor; interpretation; success
F80_64	help desk & incident management	hotline; help; resolve; problem; arise; incident; stable; practical; technician; induction; practise; contact; uncertainty; person
F80_65	outsourcing	core; competency; competence; concentration; outsource; focus; house
F80_66	hardware and software technologies	open; source; software; memory; processor; hardware; intel; disk; Linux; augment; HD; camera; tape; printer; Microsoft; Unix; PHP
F80_67	outbound & inbound transactions	outbound; inbound; reconciliation; dispatch; carrier; advice; receipt; metre; invoice; interchange; message; confirmation
F80_68	cancellation & rescission	occasion; recipient; inconsistency; cancel; complaint; incident; intention; paybox; moderate
F80_69	e-learning	knowledge; learning; group; illustrate; communication; show; example; teach
F80_70	track & trace	tag; RFID; track; trace; status; record; forwarder; interplay; consignment; GPRS, reader; barcode; label; read; delivery; postal; parcel; antenna; parcel; receive; frequency; WLAN; radio; scanner
F80_71	logistics networks	partner; organization; fulfilment; distributor; solution; party; provider; cooperation; involve; service; association; inflexible; mandate; credibility; member; board; realization; supervisor; extranet; cross; border
F80_72	mobile business	mobile; phone; device; dispatch; app
F80_73	order processing	order; release; receive; check; confirmation; mail; automation; plausibility; reject; goods; dispatch; invoice; receipt; send; payment; delivery
F80_74	business rule development	rule; business; define; set; assign; model; subject; base; pricing

continues on following page

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Continued

No.	Label	High-Loading Terms (excerpt)
F80_75	authentication	password; username; log; authentication; id; payroll; unauthorized; user; login; identity; assertion; access; session
F80_76	contracting	draft; announcement; document; contract; duplicate; tender; courier; predefine; agreement; formalize; vary; informal; framework; negotiation; permission; office
F80_77	competition	competition; advantage; competitor; price; market; position
F80_78	performance measurement	KPI; key; performance; role; endeavour; orientation; success; key; quality
F80_79	problem solving	problem; solve; arise; cause; overcome
F80_80	market area & target group	significant; market; group; target; area; share; leader; geographical; domestic

APPENDIX B

Illustrative Analysis Example

This appendix is intended to illustrate the analytical procedure of latent semantic analysis (LSA) using a simplified analysis example. For other exemplary analyses see also Müller et al. (2014) or Sidorova et al. (2008). Additionally, Martin and Berry (2007) provide a more detailed introduction to the mathematical process of LSA.

1. Data Collection and Pre-Processing:

The analysis begins by collecting and preparing a database suitable for the purpose of the analysis. This simplified demonstration will use the text fragments represented in Table 5 from a total of six exemplary project reports.

Table 5. Document Collection

ID	Project Reports	Project Descriptions (fragments)
project_#5	Development of the use of electronic media at Schweiz Tourismus (ST)	The current website is therefore based on a **content management system** (CMS)…
project_#22	Efficient and targeted information supply for members of Santésuisse	In addition to the functionalities of a traditional **content management system**, this application platform also makes it possible...
project_#25	Extending the functionality of an already existing ERP system at Agro AG	**Electronic-Business** will support competitive **strategy** and enable innovative **business** relationships…
project_#27	Electronic procurement of MRO materials and services at UBS AG	The suppliers provide **content management** for the article **data** and interfaces for **data** interchange…
project_#42	Mass customisation through B2B integration at Kaved AG	The company pursues a range of strategic **corporate** objectives as part of its **electronic-business strategy**.
project_#67	Business process support thanks to electronic instruments at OPO Oeschger AG	**Electronic-Business** is a vitally important factor in **corporate strategy**.
project_#82	Reengineering course management processes at sunrise	A **content management system** with standardized formats…

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

These text fragments will then be transferred into a structured data format during pre-processing. To do so, irrelevant terms (such as "on," "in," "and") are filtered out in the course of stopword removal. Additionally, terms that appear only once in the database and which do not contribute to the identification of patterns are also filtered out. For the purpose of this simplified database, we will forego a lemmatization, *i.e.,* unification of various forms of terms. Finally, seven terms (highlighted) are transferred into a *term document matrix* (TDM), in which the terms and their respective frequencies are associated to their respective contexts - in this case, the documents (see Table 6).

Instead of a simple term frequency, the more representative TF-IDF weighting will be used for further analyses (see also Sidorova et al., 2008). An associated TF-IDF transformation is based on a calculation of the weight (w) of each term (i) and document (j):

$$w_{ij} = tf_{ij} * idf_i \quad (1)$$

tf_{ij} is the raw term frequency of a term i in the document j and is multiplied by the inverse document frequency, providing the prevalence of a term and $idf_i = \log_2 (N/n_i) + 1$. Here, N is the number of documents in the collection; n_i is the frequency of the respective term i in the entire document collection. Such an inverse document frequency transformation gives rarer terms more weight and reduces the influence of very frequent, *i.e.,* common, terms (see Table 7).

2. Singular Value Decomposition

The *singular value decomposition* (SVD) uses the weighted TDM to identify specific groups of strongly correlated terms (*i.e.,* words that frequently appear together in documents) and strongly correlated documents (*i.e.,* documents that contain comparable words). To do so, the SVD determines individual factors that are associated with respective high-loading terms or high-loading documents. The word usage patterns expressed in these latent semantic factors stand for thematic concepts, which can then be interpreted as such in the following phase.

Table 6. Term-Document Matrix (with Term Frequencies)

		Document						
		project_#5	**project_#25**	**project_#27**	**project_#42**	**project_#22**	**project_#67**	**project_#82**
Term	Business		2		1		1	
	Content	1		1		1		1
	Corporate				1		1	
	Data			2				
	Electronic		1		1		1	
	Management	1		1		1		1
	Strategy		1		1		1	
	System	1				1		1

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 7. Term-Document Matrix (with TF-IDF Weightings)

		Document						
		project_#5	**project_#25**	**project_#27**	**project_#42**	**project_#22**	**project_#67**	**project_#82**
Term	Business		0.7360		0.3680		0.3680	
	Content	0.2430		0.2430		0.2430		0.2430
	Corporate				0.5441		0.5441	
	Data			1.6902				
	Electronic		0.3680		0.3680		0.3680	
	Management	0.2430		0.2430		0.2430		0.2430
	Strategy		0.3680		0.3680		0.3680	
	System	0.3679				0.3680		0.3680

The following will briefly describe the analytic background of SVD (see also Evangelopoulos et al., 2012; Landauer et al., 1998). Essentially, the underlying TDM (A) is divided into a product of three individual matrices by the SVD:

$$A = U\sum V^T \tag{2}$$

U here stands for the term eigenvectors and V stands for the document eigenvectors, whereas the superscript ^{T} identifies a transposition. $\sum$ stands for a diagonal matrix of singular values, which can be determined from the square root of the term eigenvalues and document eigenvalues. Building upon this $U\sum$ provides the factor loadings for the terms and $V\sum$ provides the associated factor loadings for the documents.

A factor rotation can make the interpretation of the factors easier by emphasizing their respective associated terms or document loadings without falsifying their basic assertions. This simplifies the differentiation between significant (high-loading) and insignificant (low-loading) terms of the individual semantic factors. This example will use, as suggested by Sidorova et al. (2008), a varimax rotation. To put it simply, this involves multiplying the matrix of unrotated term loadings (L_T) or document loadings (L_D) with the transformation matrices (M_T or M_D) in order to obtain matrices with rotated, more meaningful factor loadings (L^*_T and L^*_D). For a more detailed introduction of varimax rotations, see, for example, Crawford and Ferguson (1970) or Kaiser (1970).

After using the procedure illustrated here, the factor loadings represented in Table 8 can be obtained. The already rotated term loadings for two semantic factors are represented. The ensuing scree test (Figure 4) displays the eigenvalues of multiple factors in descending order and describes the extent to which an additional factor will contribute to explaining the variation included in the data set. Based on the scree plot, the selection of the first two factors seems justified.

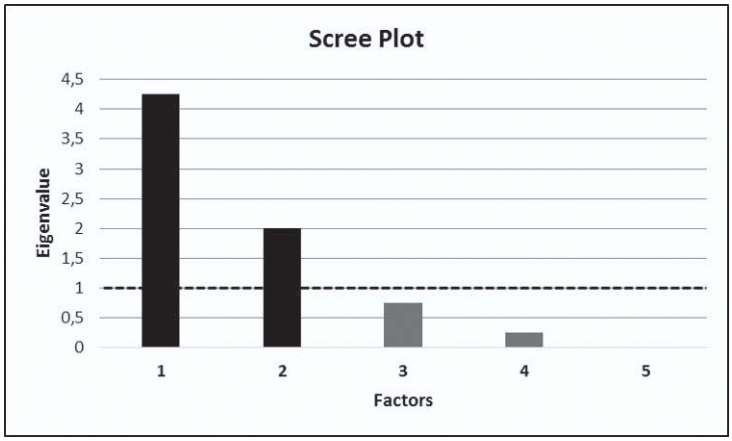
From Table 8, it can be seen that a specific group of high-loading terms (> 0.80) is associated with factor 1 ("business", "corporate," "electronic," and "strategy"), while factor 2 pools a different group of high-loading terms ("content," "management," and "system"). In the next phase, the assertions of these associated sets of terms are interpreted and labeled as thematic concepts.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 8. Term Loadings (after Varimax Rotation)

Terms	Factor 1	Factor 2
BUSINESS	**0.97577**	0.15068
CONTENT	0.17986	**0.93280**
CORPORATE	**0.85981**	0.09677
DATA	0.06022	0.20465
ELECTRONIC	**0.97577**	0.15068
MANAGEMENT	0.17986	**0.93280**
STRATEGY	**0.97577**	0.15068
SYSTEM	0.14086	**0.93839**

Figure 4. Scree Plot



3. Factor Analysis & Interpretation:

The first step of factor analysis is determining a threshold for factor loadings. This is a specific limit at which term loadings may be associated with a factor. In this illustration, we should assume a loading threshold of 0.80, so that the two factors can be associated with the terms shown in Table 9. Interpreting the compositions of associated bundles of terms then allows the interpretation of thematic concepts. The terms in factor 1, accordingly, clearly refer to the topic of an "e-business strategy." Factor 2, on the other hand, can be described as the thematic concept, "content management system." Therefore, the results of this simplified analysis illustrate how LSA can extract primarily automated thematic concepts from textual collections of documents.

Table 9. Labeled Factors with Related High-Loading Terms

Factor	Label	High-Loading Terms	Related Documents
1	e-Business Strategy	Strategy; Electronic; Business; Corporate	project_#25; project_#42; project_#67
2	Content Management System	Management; Content; System	project_#5; project_#22; project_#27; project_#82

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 44
Three-Dimensional Virtual Environment and NPC: A Perspective about Intelligent Agents Ubiquitous

Fabricio Herpich
Federal University of Rio Grande do Sul, Brazil

Felipe Becker Nunes
Federal University of Rio Grande do Sul, Brazil

Gleizer Bierhalz Voss
Federal Institute of Education Science and Technology Farroupilha, Brazil

Roseclea Duarte Medina
Federal University of Santa Maria, Brazil

ABSTRACT

The use of intelligent agents aware of the individual characteristics and context of students, allow to offer a suitable support to the real requirements. Allied to this, the implementation of these agents in the three-dimensional virtual environments, tend to transcend the existing potential in the interactions with the learning objects it contains and also to expand the alternatives of building the students' knowledge. Throughout this chapter it will be presented the development of intelligent agent called ELAI, by using the NPC strategy on the virtual world platform OpenSimulator. The ELAI provides support for teaching Computer Networking (CN), being sensitive to the context of learners to their level of expertise. In order to maximize the flexibility of interactions between the student, through the student's avatar and the NPC, an interconnection with a chatterbot was established, whose knowledge base was increased by files in AIML inherent to the topic of CN.

DOI: 10.4018/978-1-7998-0951-7.ch044

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

With the increasing use of Information and Communication Technologies (ICT) in the educational field, many requirements were emerging and changing in the current scenarios, making it necessary to reflect on the new computing paradigms in education. Given this conjuncture, research on intelligent agents in immersive environments applied to education become crucial, as it is highlighted in researches done by Iglesias & Luengo (2004), Callaghan et al. (2009), Garrido et al. (2010), Esteves et al. (2010), (Vilela, 2011) and Griol et al. (2014), since they are able to give students the ability to be immersed in the environment and interact with the learning objects, besides allowing the intelligent support to the students, assisting on their process of teaching and learning.

In this same line of this research, De Lucia et al. (2009), Bainbridge (2010), Dalgarno and Lee (2010), Ai-Lim Lee (2011), Nelson and Erlandson (2012), Childs and Peachey (2013) and Richards and Taylor (2015), discuss and points out the creation of immersive virtual environments aimed at education. The implementation of these environments demand that several factors are considered for development, such as: educational objectives, teaching strategies based on learning theories, friendly and instructional design, objects, experiments and exercises able to encourage the interaction and collaboration among users, besides passing the feeling of being immersed in the environment. Because, according to Medina (2004), learning obtained through the personal experiences of the participants and their interactions with other participants, becomes more productive, consolidated and dynamic.

Intelligent agents integrate fully to the development of these immersive virtual environments, as they offer support to its users. The relevance of intelligent agents focused on education is discussed by Soliman and Guetl (2010), in which the authors claim that in virtual learning environments, students have great flexibility in the face of numerous learning opportunities and, therefore, it is necessary to support and an intelligent orientation. Therefore, as Russel & Norvig (1995), intelligent agents must have the ability to perceive the environment in which they belong and, through sensors, operate adopting the best possible actions to solve given situations.

Furthermore, immersive experiences tend to further the engaged students with the objectives proposed in the environment, and thus it is possible to say that the intelligent agent can significantly contribute to the students' learning process. According to Soliman and Guetl (2010), the agent can act as a teacher, learning facilitator or as a colleague in collaborative environments. The agent can interact with students at different moments and places, guiding the student in the virtual environment, explaining topics, asking questions, giving feedback, helping students to collaborate with other students, offering support to the learning personalized.

This chapter of book presents the development of an intelligent agent, whose name is Intelligent Agent adaptive to the Level of Expertise of Students (ELAI). This tutor was developed through the conception of Non Player Characters (NPC), that is a perspective existing in the development platform of virtual worlds OpenSimulator and that was implemented in the Teaching Computer Networks in a Free Immersive Virtual Environment (TCN5) virtual world (Voss, 2014). This agent has as an objective to contribute in the students' teaching in disciplines of Computer Networks, presenting the ability and aspects of intelligent agents with characteristics of context-aware in the level of students' expertise to collaborate more active in this process. In order to maximize the flexibility of the interactions between the student, through his avatar and the NPC, an interconnection of the tutor with a chatterbot was established, whose knowledge base was expanded with topics about computer networks.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Throughout this chapter it will be demystified the processes related to the implementation of a Non-Player Character in the platform of virtual worlds OpenSimulator and its interconnection with a knowledge based of an external chatterbot, implemented in the Pandorabots platform. As well as detail, the necessary steps to demonstrate that the chatterbot knowledge base can be created and adapted, by adding new terms related to area of interest of the developer.

BACKGROUND

Many researches discuss and present results addressing the use of ICT in the education sector. The use of these alternatives which assist on the process of teaching and learning, has taken forms and means that go beyond the real and reaching the virtual elements, influencing the students with methods that help in getting a reliable knowledge that many times, for hazard issues and intrinsic of their cost, end up becoming unviable.

In this conjecture, this section presents the main concepts related to the use of NPC with characteristics of intelligent agents in 3D Virtual Environments for supporting the processes of student's teaching and learning in Computer Networks. Allied to this, also will be presented concepts related to ubiquitous computing and context-aware, which were indexed to the intelligent agent.

Ubiquitous Computing and Context-Aware

Ubiquitous computing is the area of research related to integration of computing the actions, attitudes and/or behaviors inherent to humans, aiming to make the Human-Computer Interaction (HCI) invisible. In this sense, Portella (2008) states that one of the goals of ubiquitous computing is make omnipresent computing resources and integrated into the daily rotine of contemporary life.

The term was coined by Weiser (1991), which stated: "The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it" (Weiser, 1991). The essence of this vision, according to Portella (2008), is to create environments with numerous computing resources and communication, integrated transparently to people.

However, the context-aware computing, according to Dey (2001), is any information that can be used to characterize the situation of entities that are considered relevant to the interaction between a user and an application. Herpich et al. (2014b) affirm that context-aware computing is characterized by performing the collection of many information involving the user, i.e., computational context, physical and time. Thus, information is collected about the environment in which it finds its location and the computing device used, for example.

These applications use context information to provide personalized services to users, such as the adaptation of content and tools in accordance with the preferences of the same, and therefore considered a context-aware environment (Herpich et al., 2014b). For Baldauf et al. (2007), these environments are able to adapt its operations to the current context without explicitly requiring user intervention, thus seeking it to maximize their usability and effectiveness, taking into account the environmental context. As applications examples is possible, to identify the tour guides, restaurants, smart homes, among others.

In this regard, Knappmeyer et al. (2013) define the context-awareness as the ability of services or applications adapt to a specific context. In addition, for both, this information, which are adapted, should be a lifecycle defined by the author as: acquisition, reasoning, presentation, organization, and finally the application or service acts on it.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ubiquitous Learning

Also known as u-learning, the ubiquitous learning is defined by Sakamura & Koshizuka (2005) as that in which is possible to learn about anything, anytime and anywhere, using ubiquitous computing technology and infrastructure. According to Hwang et al. (2009), the ubiquitous learning is an innovative method of teaching, but Sung (2009) adds that the u-learning is more than that, because essentially it translates into a vision of learning that is directly linked to the day-to-day of users, not limited only in classrooms, but e.g., at home, at work, in the library, in museum and also in a daily user interactions.

Yahya et al. (2010) consider the u-learning a new learning paradigm, which incorporates the features of e-learning and m-learning. In addition, the authors define it as a learning that takes place in a ubiquitous computing environment, which allows learning the right thing, in place at the time and on the right way. In addition, the authors conducted a survey on the different characteristics identified by researchers, presenting them in a brief summary, as it follows:

- **Permanence:** The information remains available, unless the students purposely remove it.
- **Accessibility:** Information is available whenever students need to use it by accessing them through different devices.
- **Immediacy:** The information may be obtained by students immediately.
- **Interactivity:** Students can interact with peers, teachers and specialists efficiently and effectively through different media.
- **Awareness of context:** the environment can adapt to the real situation of learners, to provide adequate information for students.

Related to u-learning, Yang et al. (2006) established a model to formally define the contexts regarding the student's needs in u-learning environments. The authors elect five factors of context-aware that need to be considered in these environments, which are: who, where, when, how and what. Relating who is the student, their preferences, knowledge capacity, among other particularities of apprentices, and segment into three entities involved, being them the factors mentioned above, the learning environment and the context model. In addition, it presents the specificities of each entity involved and the roles assigned to each:

- **Who:** the learning environment considers as the learner, setting the context model with the personal characteristics of the student, their preferences, social relations, accessibility and experiences.
- **Where:** the learning environment considers as being the student and plays the context model with location information, devices and the network that students are using.
- **When:** the learning environment considers as being the student, setting the context model as the calendar of this student.
- **How and What:** the learning environment considers as being the learning content of the students and sets the context model as the profile of this content and the description of the service that will be offered.

Basically, Hwang et al. (2009) claim that an ubiquitous learning system should be situated in real-life context, providing a systematic, economical and authentic teaching and guidance for the learners, and opportunities of practicing learning tasks as well as reviewing learning processes, enabling a potential of applying it to complex science experiments.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Three-Dimensional Virtual Environment: Virtual Worlds (VW)

Also known as immersive virtual environments or metaverse, virtual worlds are tools that simulate the real world environment in a three-dimensional (3D), providing the user with a controlled experimental environment of numerous possibilities. According to Bainbridge (2010), VW are defined as persistent online computer-generated environments where people can interact in a way comparable to the real world, either for work or for leisure.

Nelson and Erlandson (2012) and Griol et al. (2014) claim that VW are three-dimensional spaces simulated by computer, which can be exploited in first or third person, through a graphical representation called avatar. Besides offering a combination of simulation tools, these environments also provides a sense of immersion, of opportunities for communication and collaboration that instigate its application in education.

The increased use of VW in different fields, allowed their inclusion in the educational field, from which emerge new opportunities for use of these computational resources as part of support and even of motivation in the process. In this sense, according Valente & Mattar (2007), 3D virtual worlds enable the inclusion and practice of activities for experiential learning, simulation, modeling of complex scenarios, among others, with opportunities for collaboration and co-creation that cannot easily be experienced on other platforms. Accordingly, Medina (2004) reinforces this idea by stating that the learning gained thbrough the personal experiences of the participants and their interactions with other participants, becomes more productive, dynamic and consolidated.

Through the VW, is possible too, to give students the experimental practice of a particular domain of knowledge, which can be accomplished through the use of virtual labs with 3D objects, intelligent agents, among other factors, which according to Nunes et al. (2014), consist of the extension of real laboratories, thus offering new opportunities without costs and risks; the operation of virtual machines in controlled environments; has easier access because it is an environment available at all times; immediate feedback; the expertise distributed among students, where the exchange of experiences with each other is possible; the student-centered learning; and also encouraging the practice of authorship.

All these factors, in addition to technological advances led to the construction of a wide range of virtual laboratories applied to different fields of knowledge (e.g., physics, chemistry, mathematics, medicine, among others). Therefore, some platforms options to the development of virtual worlds are the pioneer of Second Life, the OpenSimulator and OpenWorderland.

In this work, the tool used was the OpenSimulator development platform, that allows the creation and configuration of the virtual worlds, combined with SLOODLE tool, which according to Peachey et al. (2010), integrates virtual world and web-based virtual learning environments technologies. In this case, it was arranged two different technologies: the virtual environment of teaching and learning Moodle and the virtual world. The interconnection provided to the SLOODLE, allows that the content present in Moodle, is inserted into the virtual world. To this study, it was selected the virtual world TCN5 (Voss, 2014), which addresses the teaching of computer networks through simulations in 3D.

Artificial Intelligence and Intelligent Agents

Artificial Intelligence (AI) is an area of computing that for years have been devoted to propose methods, techniques and tools that may be able to represent human knowledge in artificial systems. Besides, according to Liu et al. (2010), the artificial intelligence is the science of research, design and application

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

of intelligent machines or intelligent system to simulate intelligent human capabilities and extension of human intelligence.

Intelligent agents are considered a subfield of artificial intelligence and can be defined as those responsible for adopting the best action, among numerous possible, to do something about a situation. According to Silva (2004) emerged in the 90s as an evolution of object-oriented programming.

However, many authors differ on what the actual characteristics that define an intelligent agent, because of this, there are several concepts, definitions and categorizations on the subject. To Wooldridge and Jennings (1995), an agent it is a computer system that is situated in an environment that is capable of autonomous action in this environment, and aims to achieve your project goals. Already Russel & Norvig (1995) address that an intelligent agent can be described as anything that is able to perceive their environment through sensors and acts on it with their effectors.

The definition given by Hayes-Roth (1995), the author claims that intelligent agents continuously perform three functions, which are: perception of dynamic conditions of the environment; action to change the environmental conditions; and reasoning to interpret perceptions, solve problems, make inferences, and determine actions.

As there are many concepts and definitions on the subject, there are also numerous properties and characteristics that an agent has. In this sense, Franklin and Graesser (1996) present these properties to agents and affirm that to be effectively considered agents, it must meet the first four aspects among the eight listed, which are:

- **Autonomy:** Operate without direct human intervention and have some kind of control over their actions and internal states;
- **Interaction:** Can interact with other agents or humans;
- **Reactivity:** Can perceive their environment and respond to changes that occur in it;
- **Proactivity:** No need to act only in response to a stimulus, but taking initiative;
- **Continuous Activity:** The agent is in continuous process of activity;
- **Adaptation:** Can change their behavior based on some information;
- **Mobility:** Power capacity to move or be transported;
- **Personality:** Ability to convey emotional states;

Therefore, intelligent agents in the educational context, are widely used as tools to support students in order to support the interaction with the environment that are situated, providing personalized learning, as can be seen in Rissoli and Santos (2013), Dos Santos (2003), Aguiar et al. (2011). According to Tyugu (2011), the intelligent agents are software components that have some intelligent behavior characteristics that makes them special: pro activity, understanding of an agent communication language (ACL), reactivity through the ability to make some decisions and take action. To Garrido et al. (2010), the agents that have educational purposes, are able to communicate, collaborate, discuss and guide other students or agents.

Non-Player Character

Non Player Character, as clarified in its own definition, these are characters which their actions are defined through programming sets, being composed of specific scripts and functions to perform certain activities, i.e., they are not characters played by real users (people), but by predefined commands.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Basically, the NPC are computer-controlled characters and, according to Cunha & Giraffa (2001), share the same fundamental characteristics of agents. Besides, they are widely used in games like virtual characters, making these more interactive, real and challenging games.

In the context of this work, the OpenSimulator platform to development of virtual worlds, allows that the users create these characters, which are programmable entities and represented by avatars, which can be implemented by inserting scripts, in order to conduct pre-defined activities. For the programming of NPC, according to Vecino et al. (2012), is possible to use the language of script OpenSimulator Scripting Language (OSSL) and Linden Scripting Language (LSL), own of the platform, which are basically specific functions for the management and control of the actions of NPC.

These autonomous entities can be used in many situations, to provide guidance and information to other real user avatars, providing greater interaction, or incorporating new features, e.g. movements, sensors, communication, implementation tasks, simulation of human behavior, integration with other resources and to have some intelligence, through the implementation of rules on artificial intelligence.

Chatterbots

The chatterbots are artificial entities capable of conversational interaction, also known as conversational agents, robots of conversation and chatbot. According to Pilastri and Brega (2009), are computer programs that attempt to simulate conversations with users, in order to at least temporarily, lead a human being to think that are talking to someone else.

Sganderla et al. (2003), define that the chatterbots are intelligent agents developed to simulate a conversation by exchanging text messages and / or audio, similar to the virtual chats. And according to Comarella and Café (2008), they are used for the most diverse purposes, since to a relationship as a "virtual friend", even for commercial use.

To play this role of dialogue with the user, the chatterbot has the knowledge base in which performs multiple queries, filtering these searches through keywords defined in the notation established by language Artificial Intelligence Markup Language (AIML). Thus, when the conversational agent has the knowledge being debated, it processes the request and expresses his answer according to what was returned in your search. However, in circumstances where the chatterbot does not have certain information on its knowledge database, it will strive to lead or keep the conversation with the user, therefore, instead of answering the user will question him in order that him rephrases your question, until the expressions used by him are found in the knowledge base of conversational agent.

It is possible cite as an example of chatterbots agents: ALICE (2015), Eliza (2015), Elektra (2015) and Apple's Siri (2015). In the context of this work, was used the Pandorabots platform (PandoraBots, 2015), that offers free services for creating, managing and developing the chatterbot and their knowledge base, in addition to hosting the source code and make available the chat service on an allocated space on the server.

RELATED WORK

In this chapter, the reader was presented to the views with initial conceptions about Three-dimensional Virtual Environments and Intelligent Agents using concepts of U-Learning Systems and sensitive to context. With these base, the problems and difficulties as well as the research that has been done in the area of Three-dimensional Virtual Environments linked to Intelligent Agents could be presented and detailed.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

There are many scientific studies that employ the use of intelligent agents in their guidelines, with the intention of supplementing more effectively the process of teaching and learning. Allied to this, the growing implementation of three-dimensional virtual environments, new proposals and research objects are emerging, incorporating concepts of artificial intelligence along these metaverses, however the inclusion of these concepts in virtual worlds is relatively new. Against this background, the authors analyzed some researchs that made the use of these techniques and at the same time give input the proposal discussed in this book chapter.

Vilela (2011) presents a proposal that have the intention to validate approaches to solve issues related to individual training members for tasks performed as a team. Therefore, it makes use of intelligent agents able to play the team member roles for carrying out such tasks, which were developed in a 3D virtual space with the use of OpenSimulator tool.

In this perspective, the research was applied in a case study for training mechanics Portuguese

Air Force in operations for the maintanance of aircraft engines in the F-16 model. To aim can be achieved, it was developed in addition to the intelligent agents, 3D objects related to the specific parts and components of aircraft fuselages e.g., screws, motors, among others. Vilela (2011) points out also that the application of intelligent agents helps in resolving issues related to the limitations of members in teams of tasks. Moreover, it constitutes a very valuable tool for training and vocational education.

Sgobbi et al. (2014) presents a work whose proposal envisions organize immersive environments to facilitate learning activities using virtual devices with the support of artificial characters. Therefore, it discusses the use of OpenSimulator platform for the development of virtual artifacts as well as implementation of agents. To realize the interaction of students with the artifacts available in the environment and also with the agent, the authors present a series of commands to be used throughout the implementation. Moreover, present methods used to connect the artificial characters with an external chatterbot, whose base of knowledge was defined in AIML and hosted on Pandorabots service. Through this connection, attaches itself to the artificial characters the ability even to maintain a dialogue with the student and can even address specific topics and related to virtual artifacts arranged on the environment.

In addition to the three-dimensional virtual environments listed above, it is possible cite as examples some recent approaches to the implementation of immersive environments in development platform of virtual worlds OpenSimulator:

- **The Virtual Laboratory for Software Engineering Education** (Herpich et al., 2014a), which enables the practice of Unified Modeling Language (UML) modeling software called for through a tool processes GenMyModel (2015) and exposure problems through the Problem Based Learning (PBL) methodology using the OpenSimulator platform.
- **The Algorithms Virtual Laboratory** (Nunes et al., 2013), which presents and exposes a lot of exercises, materials, contents and activities so that students can respond and test their knowledge through the development of logic programming in an online tool, known as (IDEOne, 2015). The laboratory was implemented using the OpenSimulator tool and validated in a class of Graduation, in the Algorithms discipline.
- **The Chemistry Laboratory** (Nunes et al., 2014), which enables a space with countless contents and issues related to chemistry, where students are immersed and can interact with videos, slides, 3D objects, chats, quizzes and practical experiments with the aid of a NPC, which has the role of advise and guide the navigation and perform activities by students.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In relation to the paper of Vilela (2011), it presents a solution for individual training team members, in order to empower them in operations for the maintenance of aircraft engines of the Portuguese Air Force. Therefore, it combines immersive 3D virtual environment to intelligent agents, which have the role to simulate the other team members. However, the work proposed in this chapter covers the use of intelligent agents as aware of the context of guides students in order to guide students on the environment and also answer their questions about issues related to computer networks, and to adapt to the level of knowledge of these students.

The study by Sgobbi et al. (2014) presents an immersive environment in order to facilitate active learning activities through virtual artifacts with the support of artificial characters. However, the research conducted by the authors, had an emphasis on intrinsic aspects of the implementation and development of these artificial characters. While the research presented throughout this chapter, sought to implement, support and assess the NPC to use with intelligent agents characteristics in the teaching and learning topics related to Computer Networks. In summary, the study presented in this chapter presents an intelligent agent with ubiquitous features, which allows interaction of students with an NPC aware of the learner's characteristics, thus enabling the adaptation of the ELAI in terms of experience and the environment cognitive style of the learner.

CASE STUDY: AN APPROACH TO INTELLIGENT AGENTS UBIQUITOUS

This section details the issues involved in the implementation of ELAI intelligent agent, presenting the related particularities for the development of NPC in the platform development of virtual worlds OpenSimulator and even exposing how to improve the implementation of OSSL scripts to assign to the NPC intelligent agent characteristics.

Therefore, the reader is presented to a case study involving the study areas and relating them to the present work. Thus, the themes of Virtual Environments dimensional through the use of Virtual Worlds are discussed; It is implemented in this environment concepts of Artificial Intelligence and Intelligent Agents through the NPC coding; and finally they are related topics on Ubiquitous Computing and Context Aware, which are related to both the three-dimensional environment as the ELAI.

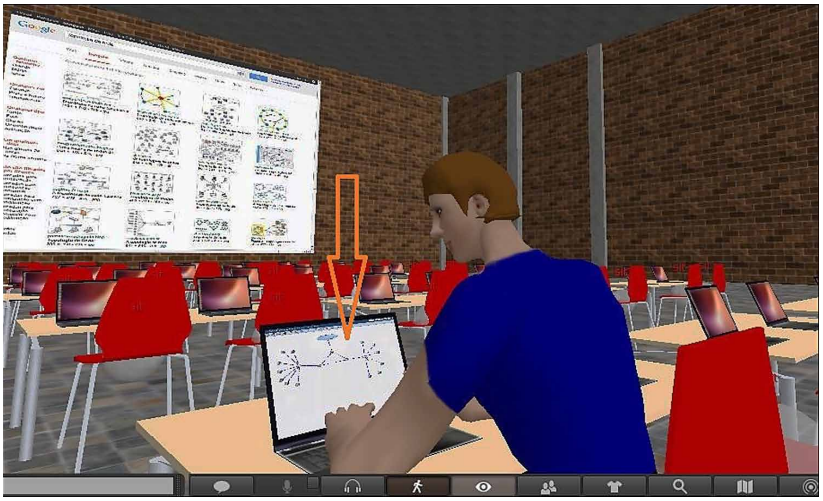
This chapter aims to present readers a discussion of the intrinsic potential in intelligent agents embedded in three-dimensional virtual environments. To perform the implantation of both an intelligent agent called ELAI, which is integrated in conversational agents (chatterbot) and inherent characteristics of intelligent agents, the same being developed using NPC functions and inserted through an avatar in the virtual environment TCN5, which is characterized by being an immersive environment, where students can explore the knowledge related to Computer Networks and obtain learning through 3D objects, activities, contents and external tools for practice (Figure 1). This agent has its features aware of the level of expertise of the context of the student aimed at teaching Computer Networks.

In this sense, to achieve the main goal, several necessary specific objectives were developed, among them are:

- Modeling and the development of intelligent agent itself, giving it the context-aware computing features to the level of knowledge students;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Student interaction with a learning tool in the TCN5 environment
Source: Voss (2014)



- In order to optimize their ubiquitous concepts have been adapted the rules for identification of expertise level to the students, presented by Possobom (2014), so that it was possible to formulate this student contextual information in order to make the ELAI and the TCN5 context-sensitive levels of experience of its members;
- In line with the last presented objective, rules were developed for the ELAI capture the student's experience level and can thus provide support with sensitivity to the current context highlighted to the student;
- Artificial intelligence rules were implemented to increase the knowledge-base of the intelligent agent, specifically focusing on Computer Networks and considering the three levels of expertise covered in this study: basic, intermediate and advanced;
- Finally, there was the evaluation of users via the System Usability Scale (SUS) proposed by (Brooke, 1996).

The development and use of immersive and three-dimensional virtual environments geared to teaching, seeks to improve and facilitate user interaction with the content seen in class. Through these environments, it is possible extend the experimentation practices beyond the normal learning environment, encouraging communication among students to solve problems. In this context, in order to enrich the immersive experiences, the authors intend to provide an intelligent agent aware of the context and the idiosyncrasies of each student, especially considering the cognitive style and the level of expertise, making it more attractive environment for users, because its characteristics and close to the individual preferences of the students.

Thus, the main impact provided by these adaptation concerns positively on student learning that make use of it, since it allows better use of educational content available in the environment. Where the student is building his knowledge, with the presence and action of the intelligent agent, which is able to detect its the level of knowledge, helping it dynamically as the accessibility of content that best fit their level of experience.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

This adaptation, according Pozzebon (2003), is justified since the teaching mediated by computer using AI techniques increases learning efficiency, allowing the student selects the variables as content, time, place and volume of matter be learned. Moreover, Soliman & Guetl (2010), claim that in virtual learning environments students have great flexibility in the face of more learning opportunities and, therefore, it is necessary to support and intelligent guidance.

Methodology

The intelligent agent ELAI is inserted in an immersive virtual environment called TCN5 (Voss, 2014), which was developed in the OpenSimulator platform and is geared to teaching of Computer Networks and has context-aware features, being sensitive to the context of cognitive style of students, respecting the characteristics and adapting their content according to user preferences.

During this section, the steps involved in this research are described, which were developed in order to elucidate the intrinsic steps to each process adopted throughout his achievement as well as to clarify the forms of established ratings:

Step 1: The first stage was characterized by the definition of the objectives that this work contemplates, and the literature review on the topics discussed in order to ground and deepen related expertise. Subsequently it was defined the tools used and the technological infrastructure resources used for the development of research.

Step 2: During the second stage, there was NPC characters coding in immersive virtual environment, therefore, implemented the development of intelligent agent ELAI. In addition, it performed the development of new classes of knowledge in AIML in order to let him expert on the topics of Computer Networks.

Step 3: During this step, it was inserted content and activities related to Computer Networks in each cognitive area. These educational materials were added respecting the expertise levels, these being classified as: basic, intermediate and advanced. Considering the disparities between levels of experience, the authors sought to provide a greater number of educational resources to students, as well as the feasibility of evaluating them while performing these activities.

Step 4: It occurred to evaluation, which was conducted through a questionnaire directed by students of undergraduate programs in Computer Science and Information Systems as well as students of the Post-Graduate Master in Computer Science at Federal University of Santa Maria (UFSM), making a total of fourteen students of Computer area involved in the evaluation.

To accomplish this was followed by a short script in order to maintain a uniform investigation with stakeholders, it was decided therefore an Educational Route, which, according to Bassani and Behar (2005), allows the appropriation and monitoring of the construction process knowledge. The pedagogical route followed by students was structured as it follows: [a] Access the Moodle VLE; [b] Access to TCN5; [c] Register in Moodle by SLoodle; [d] Meet the TCN5 environment; and [e] Simulate talks with ELAI.

At the end of this educational journey, the authors applied a usability test in order to evaluate the ELAI intelligent agent, referred to as System Usability Scale method (Brooke, 1996), which provides a questionnaire consisting of ten multiple-choice questions and it uses its own logic to estimate and measure the level of usability based on the responses obtained in questionnaires. The results of these analyzes are presented throughout the following sections.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Development of Intelligent Agent ELAI and Interconnection to Pandorabots

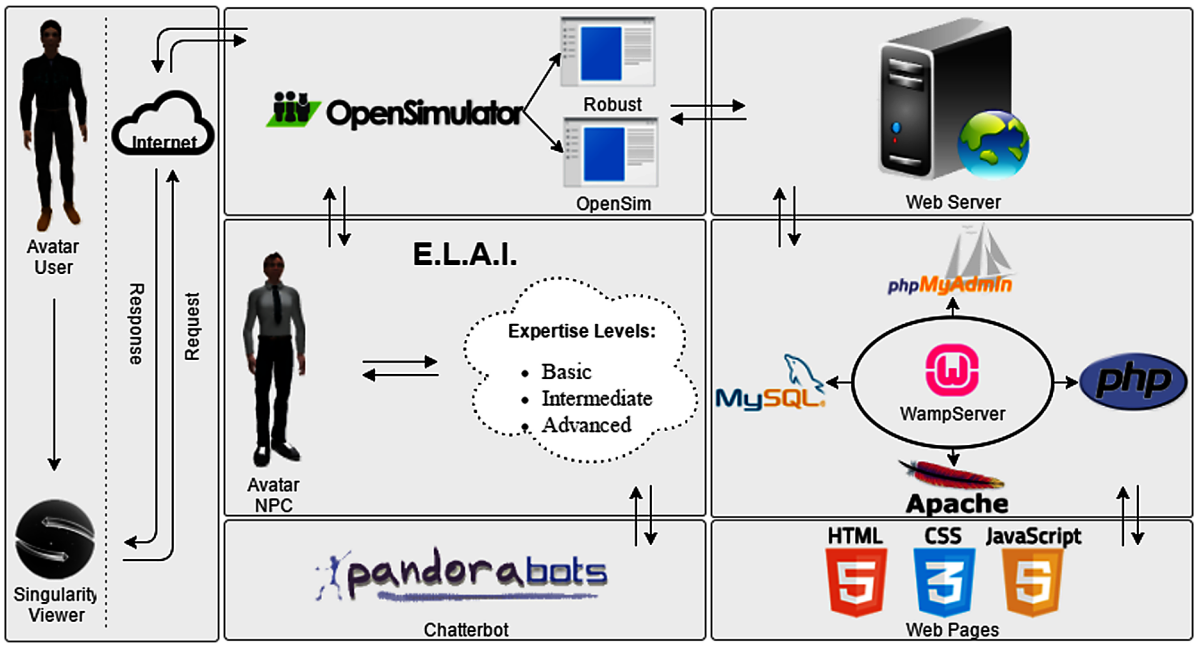
The present study aims to demystify the process related to the implementation of a Non-Player Character in the virtual world OpenSimulator and its interconnection with a knowledge base of an external chatterbot, hosted on Pandorabots web site. As well as the detail demonstrates the necessary steps, so that the knowledge base can be adapted, by inserting new terms relating to the area of interest addressed by the user.

In this sense, to offer a complete view of the work and its extension, it was drawn up an architecture which clarify all the technologies used, their interconnections and inferences that aim to meet the main goal of the study.

Figure 2 shows a scenario with all technologies used in this study, showing an operating cycle which the user follows during your navigation. Firstly, the user performs access to the immersive virtual environment TCN5, whose was developed in the OpenSimulator platform and hosted on a web server. Once entering in the environment, the user is immersed and has your contextualized representation by an avatar, where is possible carried out the proposed activities, interacting with 3D objects and simulations, also can talk about your doubts with intelligent agent ELAI. The ELAI is composed by the conversational agents (chatterbots), that explains some subject and answers the student, based on the findings made by the agent related on the level of knowledge of the student.

Other information that composed the architecture are about the web server, your services (PHP language of programing, MySQL database and the Apache web server) and at last, the web pages that was developed/installed to advise the activities of students, for example, Moodle and the module SLoodle.

Figure 2. Architecture of the project and its elements
Source: Authors, 2015



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

For success of its development of this research, it was installed the OpenSimulator metaverse on a server, and then later was performed the necessary settings that enable the implementation of NPC and enable the commands and functions inherent to these characters. Therefore, it is first necessary to make some modifications in certain lines of code in the file "OpenSim.ini", as mentioned on the official web site of OpenSimulator, which are shown in Figure 3.

It is important to highlight that, with including these commands to the source code of "OpenSim.ini", the user will be performing configuration changes that impact and, consequently, in the functioning of the virtual world, so it is necessary to know exactly what changes and what can be made. By using the expression "Enable = true", the user activates the XEngine script engine and with the command "AllowOSFunctions = true", enables the use of the functions of OSSL language, enabling the creation of NPC. Regarding "OSFunctionThreatLevel = VeryLow", this is used to set the level of threat and their allocation necessary for the performance of the implemented scripts. In this case, the level was set too low, so there was no obstruction.

At the end of installation and configuration of OpenSimulator metaverse, it was possible to start the NPC creation process, which is developing OSSL scripts, which allow both control the character as well as adding movements, animations, appearances and several actions in response to the environment and its users. To carry out the implementation of the NPC is previously necessary to create a primitive object (prim), setting to a 3D object (e.g. box), so that is possible to encode and attach scripts to this object, as it is possible to see in Figure 4.

The expression "initPos" is related to the start position of the NPC, that is, the local that the agent is created in the virtual world, in this case it was defined a position with one point away from the 3D object that contains the script. The variable "NPC" contains the "osNpcCreate" command that makes the creation of the NPC in the virtual world, in this case being assigned the name "Intelligent Agent or ELAI" to him, with the initial position set previously.

Figure 3. Settings for the implementation of NPC
Source: Authors, 2015

```
752  ;; Required changes to source code OpenSim.ini
753  [XEngine]
754      Enabled = true
755      AllowOSFunctions = true
756      OSFunctionThreatLevel = VeryLow
757  [NPC]
758      Enabled = true
```

Figure 4. Coding an NPC in OpenSimulator
Source: Authors, 2015

```
76      state_entry()
77      {
78          initPos = llGetPos() + <1,1,1>;
79          npc = osNpcCreate("Intelligent Agent", "", initPos, masterKey);
80          osNpcSay(npc, "Hi, " + name);
81      }
```


EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

At the time of its inception, the NPC will speak in the internal chat virtual world a message previously set through the command "osNpcSay", which in this case was a "Hi" with the avatar name closest to it. With this process done, it will be created in the metaverse and perform the actions that were defined in the related script. Such actions can be classified as transactions, speeches and reactions of the NPC according to information from sensors that detect the presence of some avatar to your playing field.

As it is described in Figure 4, the command "OsNpcSay" allows to insert predefined lines to be performed by the NPC through internal chat. As for the movements, these can also be set through this OSSL language that defines what position the NPC should move or other movement of actions may be included (e.g., run, jump, fly, dance).

On the issue involving the use of sensors, the position information can be changed based on some own environmental information or of its members, for example. The scripts also enable the implementation of sensory functions to the NPC, in which is possible to capture user information when they step into the sensor range of action and perform certain actions, such as walking to the user meeting, send a message by chat or voice, or make some gesture or animation.

Another possibility offered by these implementations involving the use of scripts is the NPC's ability to perform external requests to the environment, thus allowing the information query on different sites eg database, online content repositories or even seeking some information about the user that is interacting. In this sense, there is the possibility of linking the NPC with a conversational agent (chatterbot) and enhancing the capacity for interaction between user-NPC and giving it the ability to respond students with more complementary and full answers.

The interconnection with conversational agents also occurs through scripts in OSSL language, therefore, first is necessary the creation of a chatterbot. This work was used to Pandorabots tool for enabling the hosting of conversational agent for free, allow online changes, provide information and access logs of the conversations, and provide a default knowledge base with AIML files since the precursor ALICE bot.

The default class that is provided as standard, it contains general data about unexceptional topics of the day-to-day, being made available in order to assist the general user. If any question is made to the NPC and he does not know this answer, the base is scheduled to induce the user to ask the similar questions until it can be answered your question, even if in general.

In cases that want to use a custom database with specific topics on a content, customization of the data contained must be performed by inserting the terms and their meanings as it deems necessary. In order to demonstrate such an adaptation of the knowledge base in this study it was created some files in AIML notation topics related to Computer Networks area, as it is possible to see in Figure 5, to provide a sample users that performed the tests.

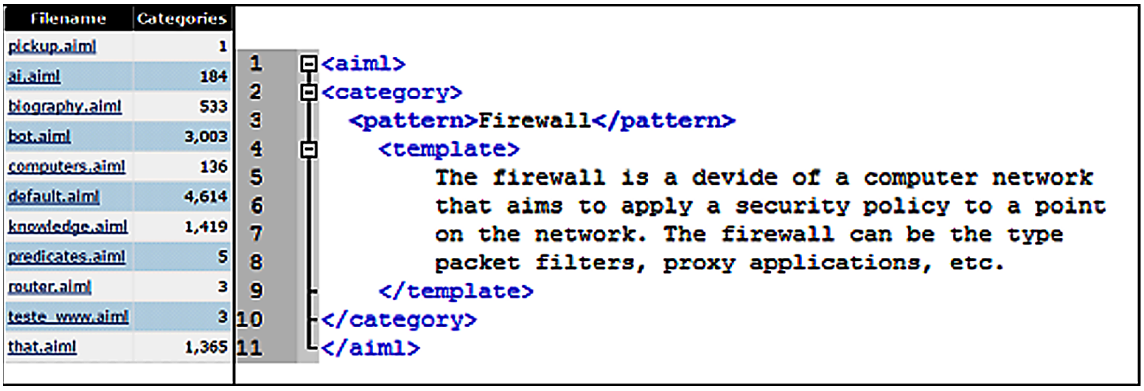
In this class, there is an example in which it is reported a keyword in "<pattern>" with the reference "Firewall" and the knowledge of its reference is reported in the tag "<template>". Thus, when the chatterbot is interrogated, filtering it is performed using a keyword, which in this case is the term "firewall" and its counterpart is the knowledge contained in the tag "<template>".

With the creation of the NPC and the customized AIML knowledge base, the interconnection of NPC with chatterbots should be done. Therefore, the scripts have been adapted in OSSL language. Figure 6 shows the interconnection script, where it is possible to check a line, which targets at twice the code for better interpretation.

At first section it is informed the chatterbot identifier code previously implemented in Pandorabots and held the requisition to the conversational agent, so the required parameters are sent: the chatterbot identifier code, the message itself and the request method were used. During the second step is performed

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. AIML files in PandoraBots
Source: Authors, 2015



Filename	Categories
pickup.aiml	1
ai.aiml	184
biography.aiml	533
bot.aiml	3,003
computers.aiml	136
default.aiml	4,614
knowledge.aiml	1,419
predicates.aiml	5
router.aiml	3
teste_www.aiml	3
that.aiml	1,365

```
<aiml>
<category>
  <pattern>Firewall</pattern>
    <template>
        The firewall is a devide of a computer network
        that aims to apply a security policy to a point
        on the network. The firewall can be the type
        packet filters, proxy applications, etc.
    </template>
</category>
</aiml>
```

Figure 6. Script to creation of the intelligent agent
Source: Authors, 2015

```
string botid = "bc3c78cc7e3428e8";
requestid = llHTTPRequest("http://www.pandorabots.com/pandora/talk-xml?botid="
   +botid+"&input="+llEscapeURL(_msg)+"&custid="+cust,[HTTP_METHOD,"POST"],"");

cust_begin=llSubStringIndex(body, "custid=");
cust=llGetSubString(body, cust_begin+8, cust_begin+23);
that_begin = llSubStringIndex(body, "<that>");
that_end = llSubStringIndex(body, "</that>");
reply = llGetSubString(body, that_begin + 6, that_end - 1);
newreply = SearchAndReplace(reply, "%20", " ");
reply = newreply;
newreply = SearchAndReplace(reply,"&quot;","\"");
reply = newreply;
newreply = SearchAndReplace(reply,"&lt;br&gt;","\n");
reply = newreply;
newreply = SearchAndReplace(reply, "&gt;", ">");
reply = newreply;
newreply = SearchAndReplace(reply, "&lt;", "<");

osNpcSay (npc, newreply);
```

using information returned by chatterbot in order to pass the message to the user only returned after the query in the knowledge base of conversational agent. Thus, at the end of this analysis, the message to the user is presented, through the command “osNpcSay (npc, newreply)”, as it is possible see in Figure 7.

Subsequent to all these procedures mentioned when the user interacts with the NPC in the virtual world, automatically it will be interacting with the conversational agent implemented in Pandorabots and can thus make inquiries and ask questions about Computer Networks topics. To demonstrate the interconnection described in this section, Figures 7 and 8 present an interaction between the NPC and the user avatar in the virtual world.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Student interaction with ELAI
Source: Authors, 2015



It is possible to see in Figure 7 that the user held by the channel "/16" chat the following question: "What is a Firewall?"; and the NPC in turn, made an appointment request in chatterbot knowledge base, returning to information that had been previously set, as shown in Figure 5.

From the development of these features, it becomes possible to perform various implementations in order to improve issues related to usability and interactivity of the NPC towards users. In this perspective, it was also attributed characteristics of intelligent agents to the NPC, making them proficient and skilled at multitasking. Therefore, when considering the criteria listed by Franklin & Graesser (1996) to intelligent agents, e.g. interaction, autonomy, reactivity, pro-activity, ongoing activity, adaptability, mobility and personality, it is possible expand the effectiveness of the NPC.

In this sense, to meet the requirements listed above, it was implemented by the NPC, sensory functions in order to enable these features. To perform activation, a function was inserted to establish a coverage radius for the intelligent agent could capture the movements of users who enter this perimeter, including enabling the capture of information from these users. Based on this, when the user through your avatar approach the intelligent agent, the sensor enables the perception of the environment; the activities on the environment itself or its users; the autonomy to perform certain activities; reactivity to perceive the environment and respond to its changes; proactivity to take initiative; mobility related to the ability to move; and ongoing activity being always active process.

In addition to meet these features, the intelligent agent can meet mobility requirements and reactivity, as when the student approaches and enters in the NPC's sphere of action, it performs the reaction move up to meet the student's position. In addition, if the student moves within that perimeter, the smart agent will follow him/her in order to provide a sense of support and monitoring, and it can offer for help. At the same time that the agent identifies the student's proximity perimeter, a welcome message is issued, establishing a communication channel through the chat to provide a user / agent interaction.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. ELAI answering the questions aware of the level of expertise
Source: Authors, 2015



Finished the implementation procedure and interconnection of these technologies, it is clear both the complexity as the existing potential for their use in the education sector, so it is important to measure and estimate their relevance in real situations, considering the interaction of the users towards the intelligent agent.

Evaluation of Intelligent Agent ELAI

In order to measure the opinions of students about their work, a questionnaire was applied aimed at evaluating the characteristics related to intelligent agent requirements, usability and interaction of students with the intelligent agent ELAI and verify the adjustments made to the context learner.

To assess the responsiveness and classify the level related of the ease of use of ELAI, it was decided to use the System Usability Scale, which according to Brooke (1996), it is a scale consisting of ten questions with a global perspective of subjective usability evaluations. To evaluate the agreement or disagreement of the respondent user, the authors used the scale of Likert (1932). The choice of SUS scale occurred not only for providing a questionnaire with questions and ready alternatives but, for having a method and a logic able to circumvent the very positive assessments of the respondents, as well as very negatives, in order to determine and achieve a result even more expressive and considerable assessed work.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In this sense, to achieve this target, the SUS has its own way of calculating your score. Brooke (1996) exposes in his research, which is primarily necessary to perform subtraction of the following:

- **Questions of Odd Number (1, 3, 5, 7 e 9):** The score is established by the value entered in the user's response, and subtracted the amount of 1 (User response - 1).
- **Questions of Pair Number (2, 4, 6, 8 e 10):** The score is composed of the amount set on 5 and subtracted the value of the response informed by user (5 - User response).

Subsequent to the calculation performed, are added all the scores of the questions answered by the user and this result is multiplied by 2.5, so that it will be possible to estimate the overall value of SUS. These SUS global values are set from 0 to 100 and divided on a scale of seven adjectives classification, in order to evaluate the result. Regarding the means and classification of adjectives and SUS score, Bangor et al. (2009) has seven possibilities with intervals between values, which can be attributed to the assessed work, namely: Best Imaginable, Excellent, Good, OK, Poor, Awful and Worst Imaginable.

In the evaluation of this work, as there were fourteen members involved in the assessment, there was the sum of all each user issues (Column "Sum of answers by User" in Figure 9) and then multiplying the value of 2.5, determining thus the total amount of SUS for each of them (Column "SUS" in Figure 9). At the end of this multiplication, averaging was performed to define the final value of SUS, the thus performed is the sum of all individual global values of each student and dividing them by the total number of users participating in the evaluation.

From this, it was obtained the final result of SUS equivalent to 73.04. Which corresponds to the concept Good on average ratings and scores of SUS presented by Bangor et al. (2009). The overview of the results by both student as a matter can be seen in Figure 9.

Figure 9. Results of the SUS
Source: Authors, 2015

Questões / Usuários	Q1	S	Q2	S	Q3	S	Q4	S	Q5	S	Q6	S	Q7	S	Q8	S	Q9	S	Q10	S	Soma das respostas por usuário	SUS
1	5	4	4	1	2	1	4	1	3	2	3	2	4	3	2	3	4	3	2	3	23	57,5
2	4	3	2	3	4	3	4	1	5	4	2	3	2	1	4	1	3	2	2	3	24	60,0
3	5	4	1	4	5	4	1	4	3	2	3	2	4	3	2	3	3	2	1	4	32	80,0
4	4	3	3	2	4	3	4	1	3	2	4	1	5	4	3	2	3	2	4	1	21	52,5
5	5	4	2	3	5	4	2	3	4	3	2	3	5	4	2	3	4	3	2	3	33	82,5
6	4	3	1	4	5	4	1	4	4	3	1	4	5	4	1	4	5	4	1	4	38	95,0
7	3	2	2	3	4	3	2	3	3	2	3	2	4	3	3	2	4	3	1	4	27	67,5
8	4	3	1	4	5	4	2	3	4	3	1	4	4	3	2	3	3	2	2	3	32	80,0
9	5	4	1	4	5	4	1	4	4	3	2	3	5	4	5	0	4	3	1	4	33	82,5
10	4	3	2	3	5	4	1	4	5	4	2	3	5	4	2	3	5	4	2	3	35	87,5
11	5	4	1	4	5	4	2	3	5	4	1	4	5	4	1	4	5	4	2	3	38	95,0
12	4	3	2	3	4	3	4	1	5	4	3	2	4	3	3	2	4	3	2	3	27	67,5
13	4	3	1	4	2	1	4	1	4	3	1	4	3	2	4	1	5	4	4	1	24	60,0
14	4	3	5	0	2	1	4	1	5	4	3	2	5	4	4	1	5	4	3	2	22	55,0
Soma total das respostas		46		42		43		34		43		39		46		32		43		41		73,04

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Regarding the first question (Q1), users were asked whether they would like to use the ELAI often in immersive virtual environment TCN5. The total of fourteen students, five were categorical in stating positively that would like to use it. Eight pointed that they agree partially with its often use.

These responses present positive results about the proposed work because expose the willingness of users frequently use the ELAI. Only one of the evaluated pointed out something other than "Agree" or "Partly agree", which used the "indifferent", that it is a neutral alternative, ie being neither positive nor negative related to the proposed work. With regard to the eight students who agreed partially, possibly the fact that their assessments will be partial, is assigned the novelty imposed by working in their daily lives, and they go on to work with an avatar accompanying and guiding the user through this process.

Regarding the number of question two (Q2) was asked users they were unnecessarily complex

ELAI. In this regard, respondents mostly said they disagreed about it (11 users). However, one of the users reported being "indifferent", another said "Partially Agree" and other "Totally Agree".

The students disagreed finding unnecessarily complex the ELAI, however it is understandable the return submitted by users that found too complicated, since the agent is inserted into a three-dimensional environment with many objects and resources, which for them is still the new and requires some time for proper adaptation. In addition, the interaction with the agent is through a communication channel that uses the expression "/ 16" expression is used for the ELAI understand that the student is communicating with him/her and not with another avatar of another student and also part of this dialogue is held in English, which may have led to greater difficulty to students.

The fact that the ELAI dialogue with students in some topics, in English, was due to the internationalization project of Institute, which also provides on-site classes taught in English. Allied to this, the authors aim to make it available internationally, making it available in online databases and repositories for wide use and occasionally getting the feedback from users so that they can be observed in future work.

These difficulties possibly damaged the interaction with ELAI, respond partly the third question (Q3), where users were asked whether they found it easy to use ELAI. In this regard three users indicated "disagree Partially" option, and against the ease of use of the agent, which leads to believe that had the same obstacles highlighted in question two. However, the majority opinion prevailed in consent with the claim, finding the understandable and uncomplicated agent.

In reference to the fourth question (Q4), respondents were questioned about whether would require technical support to use the ELAI. Regarding this question, four students showed "Totally disagree" and four "Partly disagree", but six users indicated "Agree Partially."

In this question, it was observed that there was heterogeneity as the opinions given by the students. This is because many of the students had never had contact with an immersive virtual environment, which should be turned to education, it is different from everyday environments to everyday users, e.g. video games and other games, and a new experience, with objectives and different inspirations. Nevertheless, most disagreed in need technical support to use the ELAI again, but the authors point out that this view may have been influenced by the mere fact of users have made the route stipulated being supervised by the authors of this work, in order to clarify their possible doubts.

Regarding the question five (Q5), the authors asked the users what they thought about the features of ELAI and if these were well integrated. In this question, four users showed "Indifferent", five "agreed Partially" and five "Totally agreed".

For that question, there was a positive weight to the work presented, as students were presented to all features of ELAI. Given that they communicated with the ELAI, and were aware that it was a chatterbot connected to the agent and, moreover, were notified that the ELAI was able to identify their level

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

of expertise and tailor the conversation to the level indicated. Thus avoiding frustrating the user with a very advanced topic or upset him/her with a very basic issue. Finally, students were also able to view how satisfying was the integration of ELAI with the TCN5 environment, which is very important in conducting this chapter.

However, in this case four students noted the "Not important", this fact also refers positively, as this choice may have occurred because the "invisibility" of these features that are intrinsic to the work, since the same work is in the background as support. If the main concept of ubiquitous computing: making human-computer interaction invisible. Moreover, all processes are performed by encoding scripts, what little help the user in view of the integration of features.

Regarding the issue six (Q6), students were asked whether they had too many inconsistencies in ELAI. Faced with this question, prevailed the opinion of eight members who disagreed, four marked the option "Strongly Disagree" and four "Partly disagree". But there were also five users who chose to remain impartial, noting the "indifferent" and a user only checked the "Agree Partially".

In the question about inconsistencies, many factors may have been considered by users. In general, the views of users were in favor of the submitted work, sometimes disagreeing. These views were possibly caused as a result of the chatterbot, it has part of their knowledge base in Portuguese and partly in English, often favoring the inconsistency or lack of understanding both of the questions made by users and the answers prepared by the agent, causing the latter offer answers on other subjects or even asking the student to provide more information on the topic addressed in order to deepen their search in the knowledge base.

Regarding the seventh question (Q7), students were asked if in their opinion most people learn to quickly use the ELAI. Therefore, almost all chose to assert positively that users learn easy how to use it, and seven students marked "Agree", five "Partly agree", one opted for impartiality while noting "Indifferent" and other "Disagree Partly".

The positive predominance evidenced by the responses of users, can be due to the support provided by the authors of this work during the tests. And may also have impacted on the opinion of the users and caused the students' sensation, which are understanding both environment as well as the functioning of the ELAI. However, one of the students pointed out negatively and other impartial probably these users faced with a complex environment with many new features and details to be acquired quickly, as it was the test.

In the eighth question (Q8), students were asked if they thought it ELAI was not trivial to use. So in this regard seven students responded positively, ie two marked "Strongly Disagree" and five "Disagree Partly", stating that the use of ELAI is simple and usual. Among the remaining respondents, three marked "Indifferent" other three "Partly agree" and "Agree". The opposing views can be sustained because it is a virtual world in three dimensions, which many users had never had experience or its size in length or content presented on Computer Networks and also because of the interaction with an agent in English.

In reference to question number nine (Q9), the evaluators were asked whether in their opinion felt confident to use the ELAI. In this sense, most users proved to be optimistic when use the agent, as five students pointed out the "Agree" and five "Partly agree". However, four users have chosen to point out the alternative "Indifferent".

This preference pierces even positive most students feel confident, some users remained neutral, which is an indication that the same could not be confident to use the agent. This distrust may be a result of lack of user experience in virtual 3D immersive environments or in interactions with intelligent agents, or also the large number of educational resources, with a pedagogical route that to be kept.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Since the question of number ten (Q10), respondents were asked whether need to learn a lot before to use the ELAI. Faced with this question, the opinion of the students prevailed to disagree, stating that there lacked learn a lot before use. So four students indicated "Strongly Disagree", seven showed "Disagree Partly", one remained neutral claiming to be "indifferent" and two marked "Agree Partially".

However, even the vast majority disagreeing, two users believed to be necessary to learn a lot before using it. This is due to the large amount of information that users received when performing the test, thus causing the sensation of need to learn more to be able to use their own agent. In addition, the environment itself often, because it is something new and 3D, causes uncertainty and insecurity among users regarding their use, based on the fact that most students had no experience with virtual worlds, which in this case has several alternative teaching and 3D objects for interaction.

This section sought to detail the assessment made by SUS usability scale, presenting the specificities of each question and relating to the work presented in this chapter. Therefore, the results obtained from the SUS was equivalent to 73.04. According to the mean score given by Bangor et al. (2009), corresponds to the "Good" concept. From this conception of students about the work, it was possible to identify indications that the intelligent agent ELAI and your features, contemplated the needs and preferences related to the question usability. Therefore, is possible indicate that the ELAI met the users' wishes on issues related to usability, which had the objective of estimating how much users were satisfied or dissatisfied relative ease and practicality in use the proposed work.

Among the analyzed answers, stood out the feedback on the complexity that involves the use of ELAI, where some students found difficulty in using and even said they would need technical support to other interactions. Since, users who have been assigned not believe that most people learn to use it quickly, also indicated that they found the trivial ELAI to be used. Therefore, it is necessary to correct the aspects considered troublesome by students, including reprogramming certain specific points in intelligent agent ELAI to provide an appropriate and complementary support to students, as transparent as possible to the understanding of users.

FUTURE RESEARCH DIRECTIONS

Throughout the development of this chapter, it identified some characteristics with the potential to be worked in the future, which could not be included in this study due to the intrinsic complexity of each case and it requires time more than expected in the stipulated schedule. So follow the possibilities for future work:

- Assign the ELAI the ability to cause imbalances in the construction of knowledge of students, establishing cognitive conflicts in order to develop or transform the mental schemes already developed by the learner;
- Implement with the ELAI, the ability to advise the student that remade his/her cognitive tests that define the style and level of expertise in order to maintain the agent with current contextual information about learners;
- Include to ELAI the features shown in studies of affection, relating these features with existing ones, including the more realistic characteristics, e.g. talking and gesturing;
- Develop to ELAI the ability to learn during their interactions with students or to seek external information when asked about a new concept.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

After being carried a literature review on the topics inherent in this study and in-depth analysis on them, it was found that several studies relate the concepts covered in this chapter. In both surveys, the authors discuss the use of intelligent agents and link the existing potential in this tool and its working methods, which is capable of adding significant improvement in virtual learning environments, especially in environments whose purpose is to provide a closer relationship with students, in which case, the three-dimensional virtual environments.

In support of intelligent agent presented in this work is the ability of being located adjacent to an environment with a high degree of immersion the user, having the shape of an avatar, in the same way that the student is represented in the environment, which increases proximity between the user and the artificial character. Another benefit found in intelligent agents and presented by Russel and Norvig (1995), deals with the agent feature perceive the environment in which it finds through its sensors, and take some action to certain situations, through its actuators and is establishing its actions according to their perceptions to achieve a certain goal. Adept to these features, the ELAI features, including the ability to identify, through its sensor, the closest student and check the level of experience that it hold, so that it can establish dialogue with the student and answer it on its possible doubts regarding the contents that meet with the topics in three-dimensional environment.

Allied to the contributions imposed by intelligent agents and their actions, which are defined through the use of artificial intelligence and specific languages, e.g. AIML, other factors that enrich and justify the development of both of these artificial characters as the development of three-dimensional virtual environments consist of the ability to engage users through the use of these immersive environments, as nurture and aim to improve user interaction with the content displayed on them.

Enabling even extend the practices and experimentation that are classroom views, into a controlled and highly realistic three-dimensional environment in which the student have at your fingertips the ability to repeat them as often as he/she feel necessary, in time, when and where they feel is appropriated; favoring communication between colleagues, encouraging the collaboration of both to solve problems; the aid provided by an intelligent agent that offers the student a communication channel, where it can take your questions concerning the area addressed in the environment; the authorship of capacity available to the user; and also an alternative environment that offers lower costs for implementation and non-hazardous over its use. Regarding the positive aspects provided by the three-dimensional virtual environment and also the feeling of past presence to those users, according Dalgarno et al. (2002), it is known as immersion, and occurs as a result of representation loyalty of the high degree of interaction and user control.

This chapter presented the prototyping and development of an intelligent agent called ELAI, which aimed to assist students in their activities and answer questions related to the topic of Computer Networks, but may also act as a colleague in three-dimensional collaborative environments or facilitator of learning, interacting with students at different times, the guiding, explaining tasks by asking questions or providing feedback.

In the opinion of the assessors students, the ELAI is a viable alternative that offers support in teaching and learning process thereof, and encourage interaction with 3D content and ready objects in immersive virtual environment TCN5 and also in carrying out activities. Furthermore, immersive experiences alone, seen in this work, tend to engage students further with the proposed objectives and added to this, the student actually have the assistance of an intelligent agent, encourage both to achieve these goals as and gain experience interacting with the contents disposed in the environment, making them more productive, and dynamic consolidated.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Aguiar, E.V.B., Tarouco, L.M.R., & Reategui, E.B. (2011). A construção do conhecimento matemático com engajamento e aprimoramento de habilidades cognitivas apoiada por um agente conversacional. *Revista Latinoamericana de Tecnología Educativa, 10*(2), 21–35.

Ai-Lim Lee, E. (2011). *An investigation into the effectiveness of virtual reality-based learning*. Murdoch University.

ALICE. (2015). *Official web page of ALICE*. Retrieved from http://alice.pandorabots.com/

Apple. (2015). *Official web page of Apple's Siri*. Retrieved from https://www.apple.com/ios/siri/

Bainbridge, W. S. (Ed.). (2010). Online Worlds: Convergence of the Real and the Virtual. Springer-Verlag.

Baldauf, M., Dustdar, S., & Rosenberg, F. (2007). A survey on context-aware systems. *International Journal of Ad Hoc and Ubiquitous Computing*, *2*(4), 263–277. doi:10.1504/IJAHUC.2007.014070

Bangor, A., Kortum, P., & Miller, J. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*, *4*(3), 114–123.

Bassani, PBS, & Behar, PA (2005). interROODA: mapeamento das interações individuais e interindividuais no ambiente virtual de aprendizagem ROODA. *Revista Novas Tecnologias Na Educação, 3*(2), 1–10.

Brooke, J. (1996). SUS - A quick and dirty usability scale. In A. L. M. P. W. Jordan, B. Thomas, & B. A. Weerdmeester (Eds.), *Usability evaluation in industry* (pp. 189–194). London: Taylor & Francis.

Callaghan, M., McCusker, K., Losada, J. L., Harkin, J., & Wilson, S. (2009). Integrating virtual worlds & virtual learning environments for online education. *2009 International IEEE Consumer Electronics Society's Games Innovations Conference*, (pp. 54–63). 10.1109/ICEGIC.2009.5293576

Childs, M., & Peachey, A. (2013). *Understanding Learning in Virtual Worlds*. London: Springer-Verlag London. doi:10.1007/978-1-4471-5370-2

Comarella, R. L., & Café, L. M. A. (2008). Chatterbot: Conceito, características, tipologia e construção. *Informação & Sociedade: Estudos*, *18*(2), 55–67.

Cunha, L. S., & Giraffa, L. M. M. (2001). Um estudo sobre o uso de agentes em jogos computadorizados interativos. In *Relatório Técnico Nº 017/2001* (Vol. 7, p. 77). Porto Alegre, Brasil: Campus Global - FACIN - PUCRS.

Da Silva, J. C. T. (2004). *Um modelo para avaliação de aprendizagem no uso de ferramentas síncronas em ensino mediado pela Web. Pontifícia Universidade Católica do Rio de Janeiro*. PUC-RJ.

Dalgarno, B., Hedberg, J., & Harper, B. (2002). The contribution of 3D environments to conceptual understanding. *Proceedings of the 19th Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education (ASCILITE)*.

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, *41*(1), 10–32. doi:10.1111/j.1467-8535.2009.01038.x

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

De Lucia, A., Francese, R., Passero, I., & Tortora, G. (2009). Development and evaluation of a virtual campus on Second Life: The case of SecondDMI. *Computers & Education*, *52*(1), 220–233. doi:10.1016/j.compedu.2008.08.001

Dey, A. K. (2001). Understanding and Using Context. *Personal and Ubiquitous Computing*, *5*(1), 4–7. doi:10.1007007790170019

Dos Santos, N. (2003). Agentes de Software em Ambientes Educacionais Mediados por Computador. *Revista Brasileira de Informática Na Educação*, *11*(1), 23.

Elektra. (2015). *Official web page of Elektra*. Retrieved from http://penta3.ufrgs.br:2002/

Eliza. (2015). *Official web page of Eliza*. Retrieved from http://nlp-addiction.com/eliza/

Esteves, M., Fonseca, B., Morgado, L., & Martins, P. (2010). Improving teaching and learning of computer programming through the use of the Second Life virtual world. *British Journal of Educational Technology*, *42*(4), 624–637. doi:10.1111/j.1467-8535.2010.01056.x

Franklin, S., & Graesser, A. (1996). Is it an Agent, or just a Program? A Taxonomy for Autonomous Agents. *ECAI '96 Proceedings of the Workshop on Intelligent Agents III, Agent Theories, Architectures, and Languages.* Springer-Verlag.

Garrido, P., Martinez, F., Guetl, C., & Plaza, I. (2010). Enchancing Intelligent Pedagogical Agents in Virtual Worlds. *Proceedings of the 18th International Conference on Computers in Education*, *18*, 1–8.

GenMyModel. (2015). *Official web page of GenMyModel*. Retrieved from http://www.genmymodel.com/

Griol, D., Molina, J., & Callejas, Z. (2014). An approach to develop intelligent learning environments by means of immersive virtual worlds. *Journal of Ambient Intelligence and Smart Environments*, *6*(2), 237–255.

Hayes-Roth, B. (1995). Artificial Intelligence An architecture for adaptive intelligent. *Artificial Intelligence*, *72*, 329–365. doi:10.1016/0004-3702(94)00004-K

Herpich, F., Ribeiro, R., Nunes, F. B., Voss, G. B., Fontoura, L. M., & Medina, R. D. (2014a). Virtual Lab: An Immersive Tool to Assist in the Teaching of Software Engineering. *XVI Symposium on Virtual and Augmented Reality (SVR)*. 10.1109/SVR.2014.36

Herpich, F., Voss, G. B., Nunes, F. B., Jardim, R. R., & Medina, R. D. (2014b). Immersive Virtual Environment and Artificial Intelligence : A proposal of Context Aware Virtual Environment. *UBICOMM 2014 - The Eighth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, (pp. 68–71).

Hwang, G. J., Yang, T. C., Tsai, C. C., & Yang, S. J. H. (2009). A context-aware ubiquitous learning environment for conducting complex science experiments. *Computers & Education*, *53*(2), 402–413. doi:10.1016/j.compedu.2009.02.016

IDEOne. (2015). *Official web page of IDEOne*. Retrieved from http://ideone.com/

Iglesias, A., & Luengo, F. (2004). Intelligent agents in virtual worlds. *Proceedings of the 2004 International Conference on Cyberworlds (CW04)*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Knappmeyer, M., Kiani, S. L., Reetz, E. S., Baker, N., & Ralf, T. (2013). Survey of Context Provisioning Middleware. *IEEE Communications Surveys and Tutorials*, *15*(3), 1492–1515. doi:10.1109/SURV.2013.010413.00207

Likert, R. (1932). A technique for the measurerement of attiudes. *Archives de Psychologie*, *22*(140), 1–55.

Liu, Q., Diao, L., & Tu, G. (2010). The Application of Artificial Intelligence in Mobile Learning. *2010 International Conference on System Science, Engineering Design and Manufacturing Informatization*, (pp. 80–83). 10.1109/ICSEM.2010.28

Medina, R. D. (2004). *ASTERIX – Aprendizagem Significativa e Tecnologias aplicadas no Ensino de Redes de computadores: Integrando e eXplorando possibilidades.* UFRGS.

Nelson, B. C., & Erlandson, B. E. (2012). *Design for Learning in Virtual Worlds (Interdisciplinary Approaches to Educational Technology)*. New York: Routledge - Taylor & Francis Group.

Nunes, F. B., Herpich, F., Voss, G. B., Medina, R. D., De Lima, J. V., & Tarouco, L. M. R. (2014). Laboratório Virtual de Química: uma ferramenta de estímulo à prática de exercícios baseada no Mundo Virtual OpenSim. *III Congresso Brasileiro de Informática Na Educação - XXV Simpósio Brasileiro de Informática Na Educação*, (pp. 712–721).

Nunes, F. B., Voss, G. B., Amaral, É. M. H. D., Herpich, F., Medina, R. D., & Tarouco, L. M. R. (2013). Integrating Virtual Worlds and Virtual Learning Environments through Sloodle: From theory to practice in a case of study for teaching of algorithms. *Nuevas Ideas En Informática Educativa TISE*, *9*, 598–601.

PandoraBots. (2015). *Official web page of PandoraBots*. Retrieved from http://pandorabots.com/

Peachey, A., Gillen, J., Livingstone, D., & Smith-Robbins, S. (2010). *Researching Learning in Virtual Worlds*. London: Springer London. doi:10.1007/978-1-84996-047-2

Pilastri, A. L., & Brega, J. R. F. (2009). *Chatterbot com Interatividade ao Avatar Encapsulado no Ambiente Virtual Second Life usando a base de conhecimento em AIML. Workshop de Realidade Virtual E Aumentada (WRVA)*.

Portella, F. A. (2008). *Um serviço de captura e acesso para espaços ativos. Pontifícia Universidade Católica do Rio de Janeiro*. PUC-RJ.

Possobom, C. C. (2014). *SistEX – Um sistema dinâmico para detectar a experiência do aluno.* UFSM.

Pozzebon, E. (2003). *Tutor Inteligente Adaptável Conforme as Preferências do Aprendiz.* UFSC.

Richards, D., & Taylor, M. (2015). A Comparison of learning gains when using a 2D simulation tool versus a 3D virtual world: An experiment to find the right representation involving the Marginal Value Theorem. *Computers & Education*, *86*, 157–171. doi:10.1016/j.compedu.2015.03.009

Rissoli, VRV, & Santos, GA (2013). O Agente Pedagógico Animado MInA. *XXIV Simpósio Brasileiro de Informática Na Educação (SBIE 2013) - II CBIE*, (Cbie), 828–837.

Russel, S. J., & Norvig, P. (1995). *Artificial Intelligence: A Modern Approach.* Englewood Cliffs, NJ: Prentice-Hall, Inc.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sakamura, K., & Koshizuka, N. (2005). Ubiquitous Computing Technologies for Ubiquitous Learning. *Proceedings of the 2005 IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'05)*, (pp. 11–20).

Sganderla, R.B., Ferrari, D.N., & Geyer, C.F.R. (2003). BonoBOT: Um Chatterbot para Interação com Usuários em um Sistema Tutor Inteligente. *XIV Simpósio Brasileiro de Informática Na Educação (SBIE)*, (pp. 435–444).

Sgobbi, F. S., Nunes, F. B., Bos, A. S., Bernardi, G., & Tarouco, L. M. R. (2014). Interação com artefatos e personagens artificiais em mundos virtuais. III Congresso Brasileiro de Informática Na Educação (CBIE) - XXV Simpósio Brasileiro de Informática Na Educação (SBIE), (pp. 642–651).

Soliman, M., & Guetl, C. (2010). Intelligent pedagogical agents in immersive virtual learning environments: A review. *MIPRO, 2010 Proceedings of the 33rd International Convention, 33*, 827–832.

Sung, J. (2009). U-Learning Model Design Based on Ubiquitous Environment. *International Journal of Advanced Science and Technology*, *13*, 77–88.

Tyugu, E. (2011). Artificial Intelligence in Cyber Defense. *3rd International Conference on Cyber Conflict (ICCC)*, *3*, 1–11.

Valente, C., & Mattar, J. (2007). *Second Life e web 2.0 na educação: o potencial revolucionário das novas tecnologias*. São Paulo, Brasil: Novatec Publisher.

Vecino, C., Lorenzo, C., Lezcano, L., Mesa, L. I., Franch, J. J., & Sicilia, M. Á. (2012). SLRoute: aprendiendo español en entornos inmersivos a través del Camino de Santiago. *Proceedings of the SP-DECE-2012. Ninth Multidisciplinary Symposium on the Design and Evaluation of Digital Content for Education*, (pp. 359–364).

Vilela, A. C. (2011). *Aplicação de agentes inteligentes capazes de desempenhar o papel de membros na execução de trabalhos em equipa, em ambiente virtual 3D OpenSimulator*. Universidade de Trás-os-Montes e Alto Douro.

Voss, G. B. (2014). *TCN 5 - Teaching computer networks in a free immersive virtual environment.* UFSM.

Weiser, M. (1991). The computer for the 21st century. *Scientific American*, *265*(3), 94–104. doi:10.1038cientificamerican0991-94

Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: Theory and practice. *The Knowledge Engineering Review*, *10*(2), 115–152. doi:10.1017/S0269888900008122

Yahya, S., Ahmad, E. A., & Jalil, K. A. (2010). The definition and characteristics of ubiquitous learning: A discussion. *International Journal of Education and Development Using Information and Communication Technology*, *6*(1), 117–127.

Yang, S. J. H., Huang, A. F. M., Chen, R., Tseng, S., & Shen, Y. (2006). Context Model and Context Acquisition for Ubiquitous Content Access in U-Learning Environments. *Proceedings of the IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC'06)*, *2*, 78–83. 10.1109/SUTC.2006.47

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

Chatterbots: Are artificial entities capable of conversational interaction, also known as conversational agents, robots of conversation and also chatbot.

Computer Network: Is the area of computing who cares to investigate aspects related to the interconnection of machines and the transmission of information.

ELAI: Is an intelligent agent adaptive to the level of expertise of students, which uses the intelligent agent characteristics, npc and chatterbot to promote responses about computer networks.

Inteligent Agent: Can be defined as those responsible for adoption of the best action, among many possible, to accomplish something about a certain situation.

Non-Player Character: This is a character that is not played by user, is autonomous and controlled by computer.

Ubiquitous Computing: Is the area of research that investigates the integration of computing the actions, attitudes and behaviors inherent to human beings, with the aim of making the human-computer interaction invisible.

Virtual Worlds: Are tools that simulate the real world in a three dimensional environment, providing users the sense of immersion in a controlled environment with many opportunities and experiences.

This research was previously published in the Handbook of Research on 3-D Virtual Environments and Hypermedia for Ubiquitous Learning edited by Francisco Milton Mendes Neto, Rafael de Souza, and Alex Sandro Gomes; pages 510-536; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 45
Modelling Propagation of Public Opinions on Microblogging Big Data Using Sentiment Analysis and Compartmental Models

Youjia Fang
Virginia Tech, USA

Zheng Song
Virginia Tech, USA

Xin Chen
Virginia Tech, USA

Tianzi Wang
Virginia Tech, USA

Yang Cao
Virginia Tech, USA

ABSTRACT

Compartmental models have been used to model information diffusion on social media. However, there have been few studies on modelling positive and negative public opinions using compartmental models. This study aimed for using sentiment analysis and compartmental model to model the propagation of positive and negative opinions on microblogging big media. The authors studied the news propagation of seven popular social topics on China's Sina Weibo microblogging platform. Natural language processing and sentiment analysis were used to identify public opinions from microblogging big data. Then two existing (SIZ and SEIZ) models and a newly developed (SE2IZ) model were implemented to model the news propagation and evaluate the trends of public opinions on selected social topics. Simulation study was used to check model fitting performance. The results show that the new SE2IZ model has a better model fitting performance than existing models. This study sheds some new light on using social media for public opinion estimation and prediction.

DOI: 10.4018/978-1-7998-0951-7.ch045

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

With the fast development of Internet, last decade has witnessed information explosion over social media. Major microblogging platforms such as Twitter in the United States and Sina Weibo in China provide the instant access to rapid-changing news and allow users to generate self-organized networks on popular social topics and celebrities (Gao, Abel, Houben, & Yu, 2012). As microblogging platforms have millions of users and billions of new messages every day, they have become an important channel to evaluate and predict the trends of public opinions on high-impact news and social events. But the high volume, velocity and complexity of microblogging messages also impose challenges on the processing and analysis under big data context.

The mechanism of news propagation is similar to the epidemiology of infectious disease. Disease diffusion starts with a small group of infected, and transmits along the paths of daily contact and interaction. Compartmental models in epidemiology are a state-of-practice approach to model the dynamics of disease propagation. The population is classified into several compartments, such as susceptible (*S*), exposed (*E*), infected (*I*) and recovered (*R*), to represent different epidemic states (Bettencourt, Cintrón-Arias, Kaiser, & Castillo-Chávez, 2006). People in susceptible compartment are healthy and have the opportunity to be infected; the susceptible will be infected if they have sufficient contact with the infected (Hethcote, 2000). There is a prolonged incubation period (i.e. exposed compartment) for some diseases, where people who have contracted the disease need a longer time to develop symptoms and become contagious. For some other diseases, the infected can fully recover and obtain immunity. Classical compartmental models for biological epidemiology include SI, SIS, SIR, and SEIR models. Different models are suitable for diseases with different characteristics.

Information diffusion has been widely studied by previous literatures (Canneyt, Schockaert, & Dhoedt, 2015; Escamilla et al., 2016; García-Silva, Rodríguez-Doncel, & Corch., 2013; Gruhl, Guha, Liben-Nowell, & Tomkins, 2004; Kwon & Lee, 2013; Lerman & Ghosh, 2010; Matsubara, Sakurai, Prakash, Li, & Faloutsos., 2012; Romero, Meeder, & Kleinberg., 2011; Shuai et al., 2012; Trpevski, Tang, & Kocarev, 2010; Weng, Lim, Jiang, & He, 2010; Jaewon Yang & Leskovec, 2010). Previous studies primarily focus on the source and mechanism of information diffusion in social media and its impact on public society. Epidemiological models have been used in some of these studies (Cha, Haddadi, Benevenuto, & Gummadi, 2010; Jin, Dougherty, Saraf, Cao, & Ramakrishnan, 2013; Jin et al., 2014; Khelil, Becker, Tian, & Rothermel, 2002; Newman, 2002; Xiong, Liu, Zhang, Zhu, & Zhang, 2012; Jiang Yang & Counts, 2010). Among all compartmental models in epidemiology, *SIZ* and *SEIZ* models are two most popular models for information diffusion, as they include a skeptical compartment not often seen in biological epidemiological process.

Some researchers have used *SEIZ* model to model news diffusion using Twitter microblogging data (Jin et al., 2013; Jin et al., 2014). They assumed that microblogging users who posted messages on popular news have a positive opinion on the news (i.e. being infected), and users who posted no message have a negative opinion on the news (i.e. being skeptical). With this treatment, the researchers only observed data of the infected compartment. However, the microblogging users can in fact post messages to express their both positive and negative opinions, so previous assumptions and treatments may not be completely reasonable.

The objective of this study is to simultaneously model the propagation of positive and negative public opinions using a large volume of Sina Weibo microblogging data. We used natural language processing and sentiment analysis to process the big microblogging data and extract the positive and negative

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

opinions, and then implemented two existing (*SIZ* and *SEIZ*) and one newly-developed (*SE2IZ*) models to model the dynamic of news propagation. Our analysis approach can deal with big data scenario due to its low computational cost. The methods in this study can help real-time data-driven decision making in political, business and other fields.

RELATED WORK

In news propagation context, the susceptible are people who have the opportunity to be exposed to the news, and the infected are people who perceived the news and posted messages supporting the news. The most basic compartmental models only involve two types of compartments: susceptible and infected. In SI model, the susceptible will be permanently infected and cannot recover. Thus the entire population will eventually be infected. In SIS model, the infected have the opportunity to recover and become vulnerable again (Bettencourt et al., 2006). These two models have some drawbacks on modeling information diffusion. Both models have too few compartments to account for sophisticated dynamics of news propagation. Moreover, SIS model may not be suitable because the news can fade out quickly and the recovered people will not be interested in same news again.

During news propagation, people may either agree (infected) or disagree (skeptical) with the news. This phenomenon is not commonly seen in biological epidemiology as the susceptible are only infected in that case. A skeptical (*Z*) compartment was developed to account for the skeptical state in news propagation (Bettencourt et al., 2006). In SIZ model, the susceptible have opportunities to either agree or disagree with the news.

In some situations, there is a prolonged incubation period for the susceptible perceiving the news to formulate a clear opinion on the news. An exposed (*E*) compartment was developed to represent the incubation period, leading to *SEIZ* model (Bettencourt et al., 2006; Hethcote, 2000; Jin et al., 2013; Jin et al., 2014). *SEIZ* model only has an exposed compartment before infected compartment, without a symmetric counterpart before skeptical compartment.

DATA AND METHODS

Sina Weibo Microblogging Data

This study used a large (~20 Gb) Sina Weibo data set shared by Data-Tang, a private online data sharing service that periodically extracts, processes, and shares publicly accessible microblogging data and covers the a large portion of active microblogging users (Datatang). We selected 13 most popular and high impact social event topics initiated in April to May 2014. We extracted the first 10-12 days' data since initiation of the event, resulting in over 84,000 Sina Weibo microblogging messages from over 64,000 Sina Weibo users. Each Sina Weibo message consists of the post date, time and message text in Chinese language.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

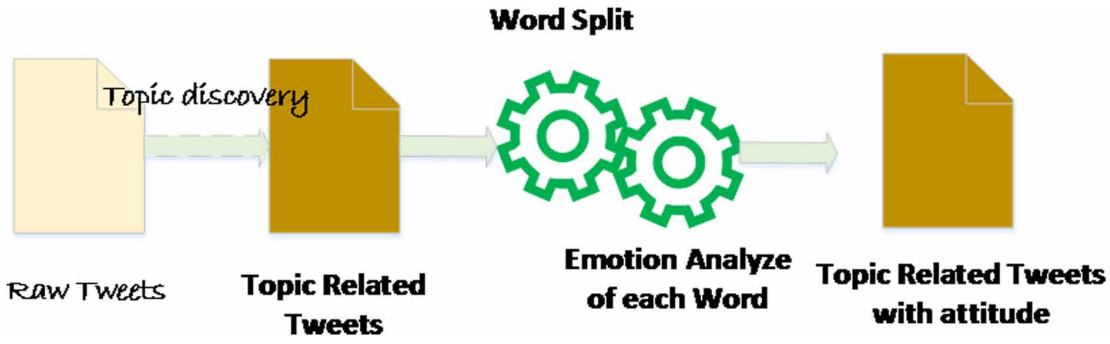
Natural Language Processing and Sentiment Analysis

Natural language processing and sentiment analysis were used to process the high volume of unstructured microblogging data and extract opinions of microblogging users on pre-selected social topics. The work flow of natural language processing and sentiment analysis is shown in Figure 1. First, we used Natural Language Processing & Information Retrieval (NLPIR) Chinese language processing system to split sentences in each message into meaningful phrases (NLPIR). Second, we matched the phrases with NTUSD Chinese sentiment dictionary which provides attributes of positive and negative emotions (Ku & Chen, 2007). For example, *support, like, great* are keywords with positive emotion, and *hate*, *dislike*, *against*, *protest* are keywords with negative emotion. Third, we assigned a value of 1 for each keyword with positive opinion, and -1 for each negative keyword with negative opinion. Forth, we sum up the values of all keywords within a message, and defined the message to represent a positive opinion on the news if the sum is positive, and to present a negative opinion on the news if the sum is negative. The message is considered be neutral if the sum is zero. As a neutral message does not provide useful information on whether the microblogging user is infected or skeptical, it was excluded from further analysis.

An example of natural language processing and sentiment analysis is demonstrated as follows. A microblogging message regarding Houston Rocket NBA basketball team is "In middle school I loved Rocket team so much, followed them closely, all because of Tracy McGrady" (originally as "初中超喜欢火箭。各种关注。都是因为麦蒂" in Chinese). We used NLPIR system to split the sentence (in Chinese) into "In middle school | I loved (so much) | Houston Rocket team | followed them | closely | all | because of | Tracy McGrady" ("初中|超喜欢|火箭。各种|关注。都是|因为|麦蒂" in Chinese). Among these phrases there is only one keyword ("love") that matches the NTUSD keyword dictionary with positive emotion. Therefore, this message has a sum of 1 and is considered to represent a positive opinion on the social topic of Houston Rocket team.

After natural language processing and sentiment analysis, seven popular social topics with sufficient number of opinion-available (i.e. non-neutral) messages were selected for further analysis (Table 1). For each topic, we calculated the number of messages with positive or negative opinions in daily basis from the initiation date of the topic. The amount of daily messages with positive opinion are observed data for infected compartment, and the amount of daily messages with negative opinion are the input for skeptical compartment.

Figure 1. Natural language processing and sentiment analysis



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. List of seven selected popular social topics

Social Topic	Event Detail
Evergrande FC	Chinese Guangzhou Evergrande Football Club (FC) won the 100th game with famous coach Marcello Lippi.
Air Pollution	Air pollution in north-eastern China was greatly alleviated during May 2014. Air quality in Beijing hit the best point over the past several years.
House Price	House price in China began to decline after years of rapid increase.
"My Old Classmate"	The movie "My Old Classmate" came on show in cinemas.
National Civil Service	Chinese National Civil Service Exam has ended.
Corruption	The criminal charge on corruption of some high-rank Chinese government officials has drawn much public attention.
GMO food	A famous Chinese TV host, Cui Yongyuan, initiated a public debate against genetically modified organisms (GMO) food.

Compartmental Models for Propagation of Public Opinions

This study focused on two compartmental models most suitable for news propagation, *SIZ* and *SEIZ* models. Both models have an infected compartment and a skeptical compartment. *SEIZ* model has an exposed compartment before the infected compartment. However, this paper is the first study to consider that messages with positive emotions represent positive opinions of the infected people on the news, and messages with negative emotions represent the negative opinions of the skeptical people on the news. This way, we can collect data for both *I* and *Z* compartments. Besides, we assumed that each message represents one data point, and thus multiple messages by the same user contribute to multiple data points even they occurred in the same day. Moreover, we developed a new *SE2IZ* model based on *SEIZ* model by adding an exposed compartment (E_z) before skeptical compartment as well as additional paths to skeptical.

Compartmental models use a system of Ordinary Differential Equations (ODEs) to model the transition quantities and rates of subjects between different compartments over time. Transition rates and transition probabilities are major parameters to estimate in the system of ODEs. In this study, the parameters were estimated by solving the system of ODEs numerically.

SIZ Model

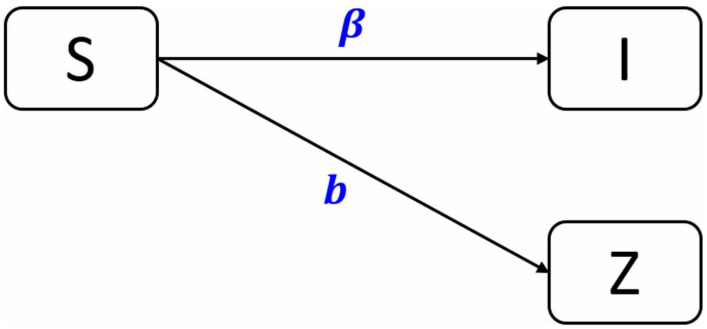
The framework of *SIZ* model is shown in Figure 2, where the blue-colored parameters are contact/transition rates. After the susceptible contact the infected or skeptical, they will move to infected or skeptical compartment with different transition rates.

The system of ODEs for the SIZ model is:

$$\begin{cases} \dfrac{dS}{dt} = -\beta S \dfrac{I}{N} - bS \dfrac{Z}{N} \\ \dfrac{dI}{dt} = \beta S \dfrac{I}{N} \\ \dfrac{dZ}{dt} = bS \dfrac{Z}{N} \end{cases}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. SIZ model

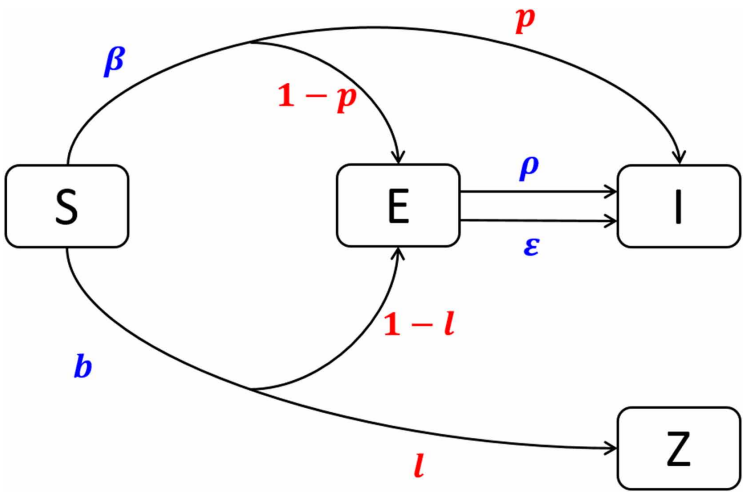


where *S*, *I*, *Z*, i.e. *S(t)*, *I(t)* and *Z(t)*, are the numbers of subjects in *S*, *I*, *Z* compartments at time *t*; *N* is the population size, *N*= *S(t)*+*I(t)*+*Z(t)*; β is per-capita *S*-*I* contact rate (i.e. number of *S* in contact with *I* per day); *b* is per-capita *S*-*Z* contact rate (number of *S* in contact with *Z* per day) (Bettencourt et al., 2006; Hethcote, 2000).

SEIZ Model

The framework of *SEIZ* model is shown in Figure 3, where the blue-colored parameters are contact/transition rates and red-colored parameters are transition probabilities. In *SEIZ* model, a proportion of susceptible enter the exposed compartment before they move to infected compartment (i.e. need a prolonged incubation period to generate positive opinions on the news). Meanwhile, it is possible for susceptible to be infected (but firstly go through exposed compartment) after they contact with the skeptical (i.e. the susceptible disagree with the skeptical and thus become infected).

Figure 3. SEIZ model



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The system of ODEs for the SEIZ model is:

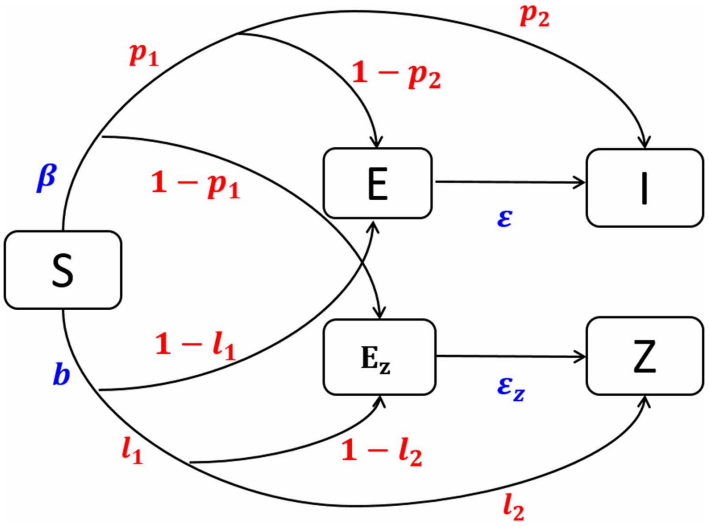
$$\begin{cases} \dfrac{dS}{dt} = -\beta S\dfrac{I}{N} - bS\dfrac{Z}{N} \\ \dfrac{dE}{dt} = (1-p)\beta S\dfrac{I}{N} + (1-l)bS\dfrac{Z}{N} - \rho E\dfrac{I}{N} - \varepsilon E \\ \dfrac{dI}{dt} = p\beta S\dfrac{I}{N} + \rho E\dfrac{I}{N} + \varepsilon E \\ \dfrac{dZ}{dt} = lbS\dfrac{Z}{N} \end{cases}$$

The term E, i.e. $E(t)$, is the number of subjects in E compartment at time t; $N= S(t)+I(t)+Z(t)+ E(t)$. The ε is E-to-I self-transition rate (i.e. no further contact with I needed), and $1/\varepsilon$ is the average incubation period; ρ is per-capita E-I contact rate (number of E in contact with I per day), representing a promotional effect of E-I contact; p is the S-to-I transition probability given the contact with I, and (1-p) is the S-to-E transition probability given the contact with I; l is the S-to-Z transition probability given the contact with Z, and (1-l) is the S-to-E transition probability given the contact with Z; β is per-capita S-I contact rate; b is per-capita S-Z contact rate.

SE2IZ Model

The framework of *SE2IZ* model is shown in Figure 4. There are three motivations for developing the new *SE2IZ* model. First, *SE2IZ* model includes an exposed compartment before skeptical compartment (E_z) in order to account for long incubation period before skeptical. Second, after the susceptible contact the infected, they still have a chance to be the skeptical; the E_z compartment provides a new path to skeptical compartment in addition to S-Z direct contact. Third, we excluded promotional contact rate ρ in order to reduce model complexity and parameter redundancy.

Figure 4. SE2IZ model



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The system of ODEs for *SE2IZ* model is:

$$\begin{cases} \dfrac{dS}{dt} = -\beta S\dfrac{I}{N} - bS\dfrac{Z}{N} \\ \dfrac{dE}{dt} = p_1\left(1-p_2\right)\beta S\dfrac{I}{N} + \left(1-l_1\right)bS\dfrac{Z}{N} - \varepsilon E \\ \dfrac{dE_z}{dt} = l_1\left(1-l_2\right)bS\dfrac{Z}{N} + \left(1-p_1\right)\beta S\dfrac{I}{N} - \varepsilon_z E_z \\ \dfrac{dI}{dt} = p_1 p_2\beta S\dfrac{I}{N} + \varepsilon E \\ \dfrac{dZ}{dt} = l_1 l_2 bS\dfrac{Z}{N} + \varepsilon_z E_z \end{cases}$$

The term E_z, i.e. $E_z(t)$, is the number of subjects in E_z compartment at time t; $N= S(t)+I(t)+Z(t)+ E(t) + E_z(t)$; ε and ε_z are the *E*-to-*I* and E_z-to-*Z* self-transition rates; p_1 is the *S*-to-*I* transition probability given a contact with *I*, $(1\text{-}p_1)$ is the *S*-to-E_z transition probability given a contact with *I*; p_2 is the secondary *S*-to-*I* transition probability after p_1, $(1\text{-}p_2)$ is the secondary *S*-to-*E* transition probability after p_1; l_1 is the *S*-to-*Z* transition probability given a contact with *Z*, $(1\text{-}l_1)$ is the *S*-to-*E* transition probability given a contact with *Z*; l_2 is the secondary *S*-to-*Z* transition probability after l_1, $(1\text{-}l_2)$ is the secondary *S*-to-E_z transition probability after l_1; β is per-capita *S*-*I* contact rate; *b* is per-capita *S*-*Z* contact rate.

SIMULATION STUDY ON MODEL FITTING PERFORMANCE

We used simulation study to investigate the model fitting performances of three compartmental models. The purpose of simulation study is to check whether the models can fit the simulated data as well as expected. Simulation study includes two steps. First, for each compartmental model, we generated a simulated data set of all compartments based on pre-specified values of model parameters. Second, we fit each model with on corresponding simulated data, and derived the fitted curve for each compartment over time. In the second step, we fit models using two versions of data: complete and partial versions. The complete version includes simulated data of all compartments; while the partial version only includes simulated data of *I* and *Z* compartments, which is similar to a real-world scenario.

The results of simulation study are shown in Figure 5, Figure 6, and Figure 7. The fitted curves match the simulated (i.e. true) data points very well. Moreover, the model fitting performances are similar when using complete and partial versions of data. These results suggest that, when models are correctly specified (i.e. the model matches the underlying process of the data), the compartmental models perform very well even when only *I* and *Z* compartments are observed.

MODELLING SINA WEIBO MICROBLOGGING DATA

After simulation study, we implemented three compartmental models on the seven selected topics of Sina Weibo microblogging data processed by natural language processing and sentiment analysis (Table 1). We adopted the modeling procedures similar to previous literatures (Bettencourt et al., 2006; Jin et al.,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. SIZ model fitting performance in simulation

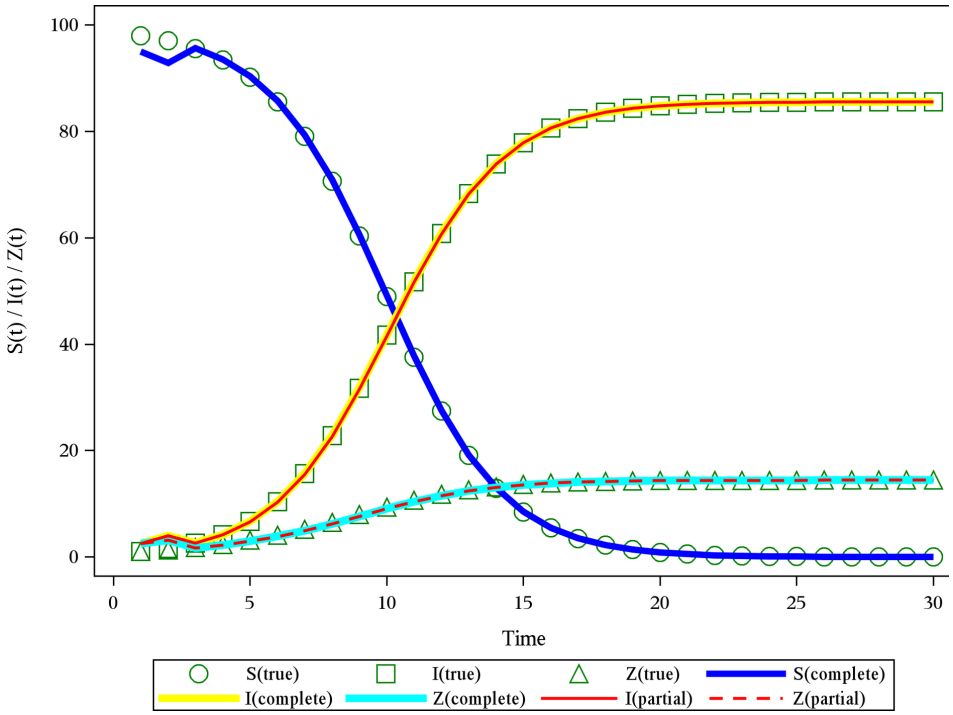
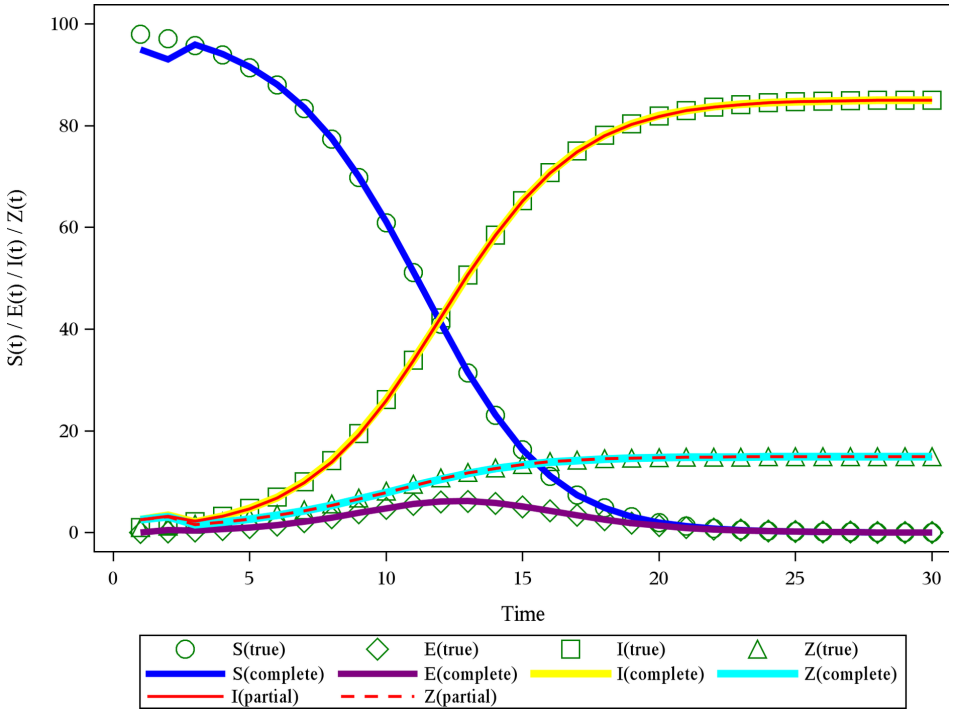
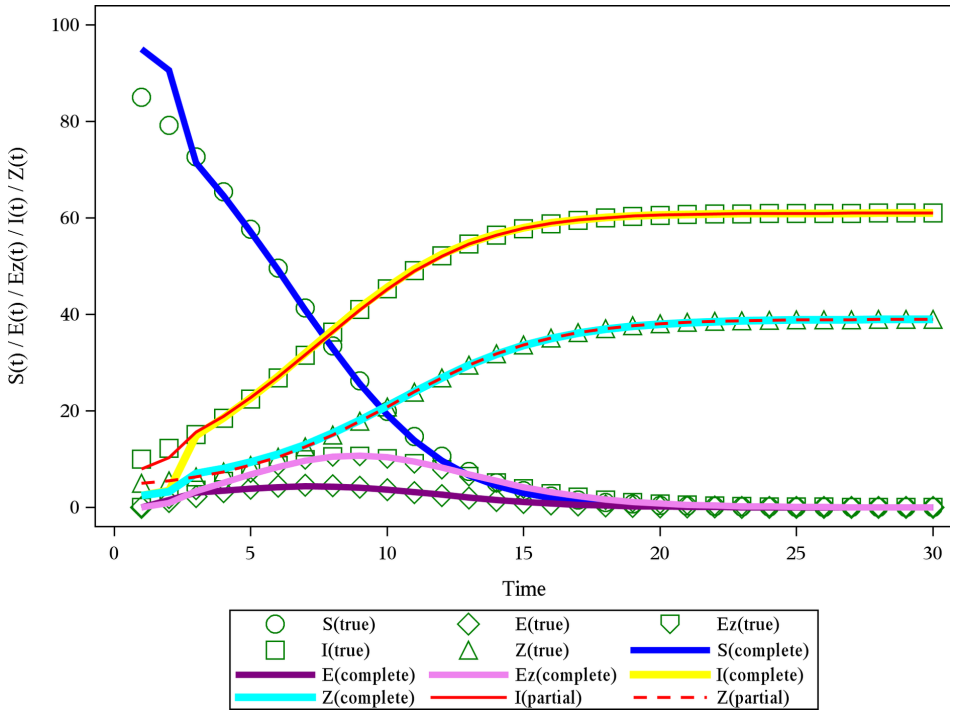


Figure 6. SEIZ model fitting performance in simulation



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. SE2IZ Model fitting performance in simulation



2013; Jin et al., 2014). We assumed that *I(0)* and *Z(0)* are observed in the data, but *S(0)*, *E(0)*, $E_z(0)$ and *N* as unknown parameters to estimate. Ordinary non-linear least squares method was used to estimate the model parameters. Model fitting performances of three models were compared both graphically (using fitted curves) and numerically (using MSE and MAD). We used two quantitative measures of overall model fitting errors: Mean Squared Error (MSE) and Mean Absolute Deviation (MAD), which are defined as:

$$MSE = \sum_{t=1}^{T} \left\{ \left[I_{fit}(t) - I_{obs}(t) \right]^2 + \left[Z_{fit}(t) - Z_{obs}(t) \right]^2 \right\} \Big/ T$$

$$MAD = \sum_{t=1}^{T} \left\{ \left| I_{fit}(t) - I_{obs}(t) \right| + \left| Z_{fit}(t) - Z_{obs}(t) \right| \right\} \Big/ T$$

The overall model fitting errors for 7 social topics are shown in Table 2. The results clearly show that the MSE and MAD for *SE2IZ* model are constantly lower than two existing models, while *SIZ* model has the highest model fitting error among 3 models. The existence of exposed compartment(s) significantly increases model flexibility for varying dynamics of news propagation. *SE2IZ* model has the best model fitting accuracy among three models due to its double exposed compartments and more transition paths.

The observed data and fitted curves for seven social topics are shown in Figure 8, Figure 9, Figure 10, Figure 11, Figure 12, Figure 13, and Figure 14. *SE2IZ* model fit the observed data very well, especially when *I* or *Z* compartment was very different from a classical S-shape (e.g. Air Pollution, House Price, "My Old Classmate" movie, and National Civil Service). These results agree with MSE and MAD findings that *SE2IZ* model is more flexible than *SEIZ* and *SIZ* models in modeling different dynamics of new propagation.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Overall model fitting errors for seven social topics

Topic	MSE			MAD		
	SIZ	SEIZ	SE2IZ	SIZ	SEIZ	SE2IZ
Evergrande FC	21,261	13,098	7,555	162	132	107
Air Pollution	54,232	50,977	2,367	281	262	57
House Price	118,064	98,565	6,049	396	322	82
"My Old Classmate"	192,059	159,990	7,752	538	452	93
National Civil Service	87,673	33,409	2,922	366	187	56
Corruption	15,563	15,340	3,325	144	139	56
GMO Food	10,594	8,182	1,689	113	96	46

We used the fitted curves to assess the trends of public opinions on related scocial topics. Sina Weibo microblogging users were more likely to have positive opinions on Evergrande FC (e.g. fans were proud of Evergrande FC's achievement in international club soccer competition). In contrast, microblogging users were more likely to have negative opinions on Air Pollution (e.g. complaining about the poor air quality), House Price (e.g. complaining about the rapidly inflating house price), Corruption (e.g. complaining about the adverse impact of government job corruption), and GMO food (e.g. concerning the adverse health impact of genetically modified food). Moreover, microblogging users were more likely to have neutral opinions on "My Old Classmate" movie (e.g. the movie had an averaged rating) and National Civil Service (e.g. some people were eager for a job in national civil service due to good job security and benefit, while some others complained that too many positions significantly increased the government's financial burden).

CONCLUSION AND DISCUSSION

This paper is the first study to simultaneously model the propagation of positive and negative public opinions using compartmental models. Natural language processing and sentiment analysis were used to identify public opinions on popular social topics using a large volume of Sina Weibo microblogging data. We developed a new *SE2IZ* model based on existing *SEIZ* and *SIZ* models and implemented all three models on the processed data. *SE2IZ* model has a better model fitting performance than *SEIZ* and *SIZ* models due to an additional exposed compartment before skeptical compartment and more paths to skeptical compartment.

This study sheds some new light on using microblogging social media data to perform public opinion prediction and data-driven decision making in political, business and other fields. The compartmental models can also be used in other fields such as biomedical, public health, epidemiology, business analytics, marketing, survey and poll, education, and transportation. All these fields have strong interest and demand on modeling event data (Antin et al., 2016; Brin, Motwani, Ullman, & Tsur, 1997; Chen, Fang, Guo, & Hanowski, 2016; Chen, Fang, & Lockee, 2015; Fitch et al., 2013; Guo & Fang, 2013; Guo, Fang, & Antin, 2014, 2015; Guo et al., 2016; Lebo & Norpoth, 2007; Ouimet et al., 2014; Waller & Gotway, 2004).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Observed data and fitted curves for Evergrande Football Club

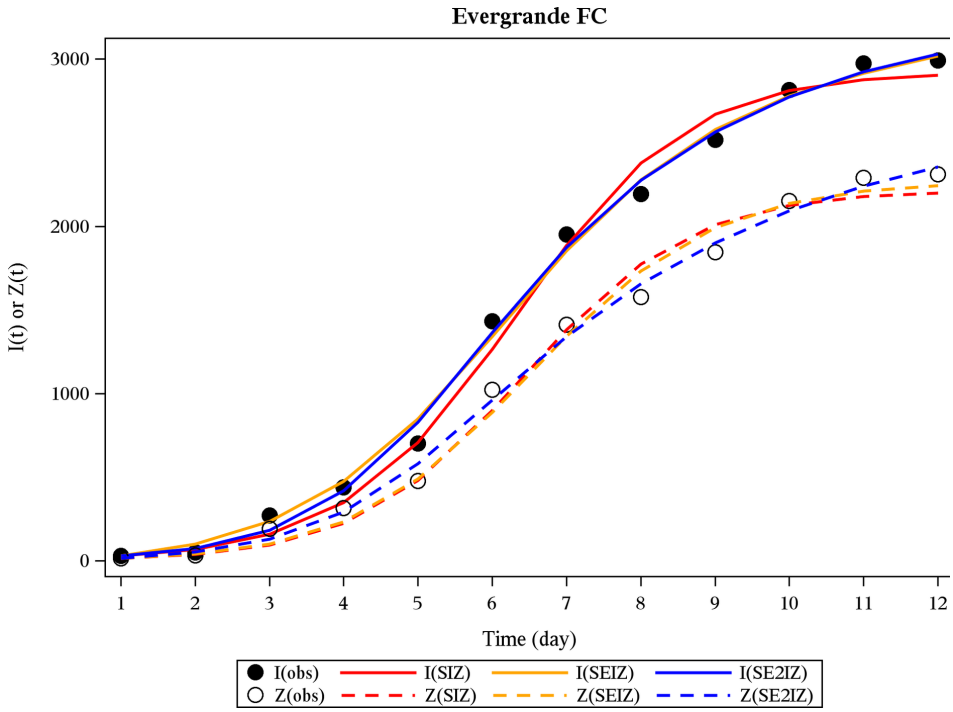
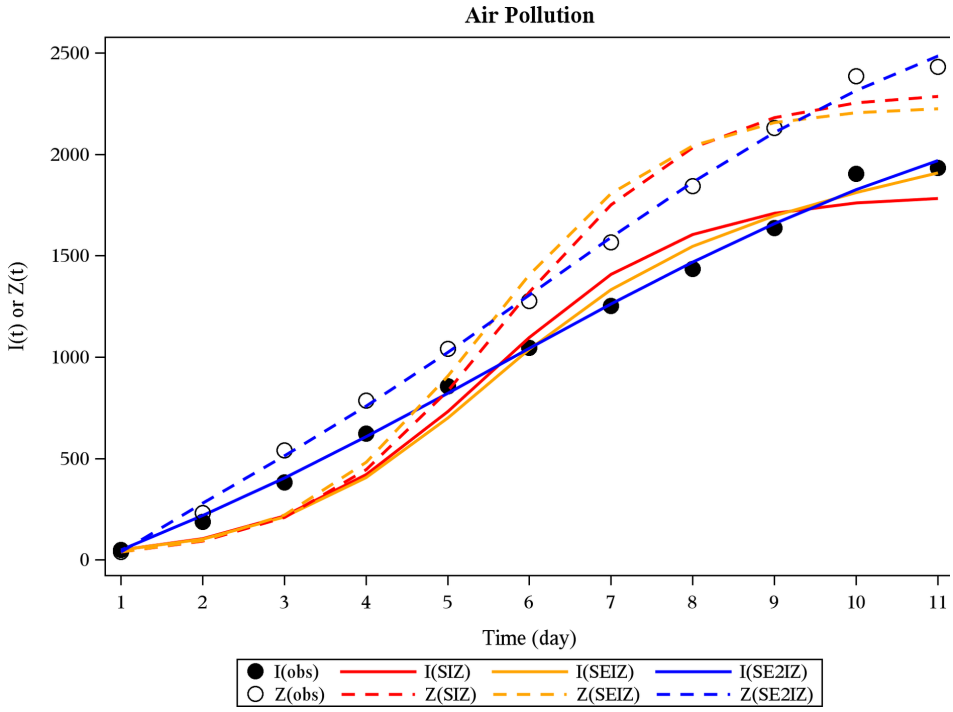


Figure 9. Observed data and fitted curves for air pollution



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. Observed data and fitted curves for house price

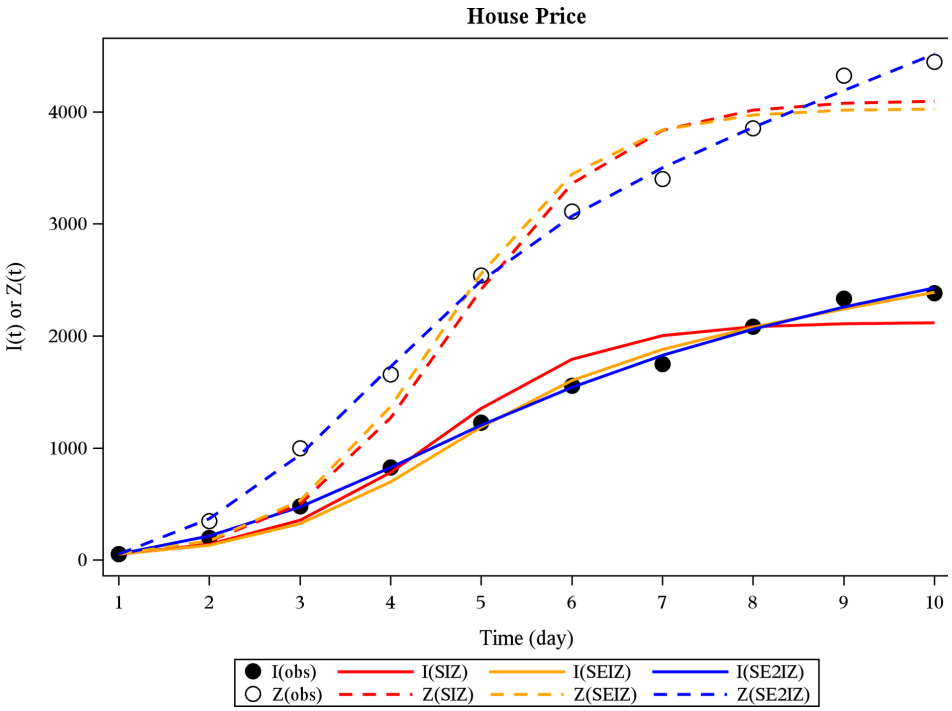
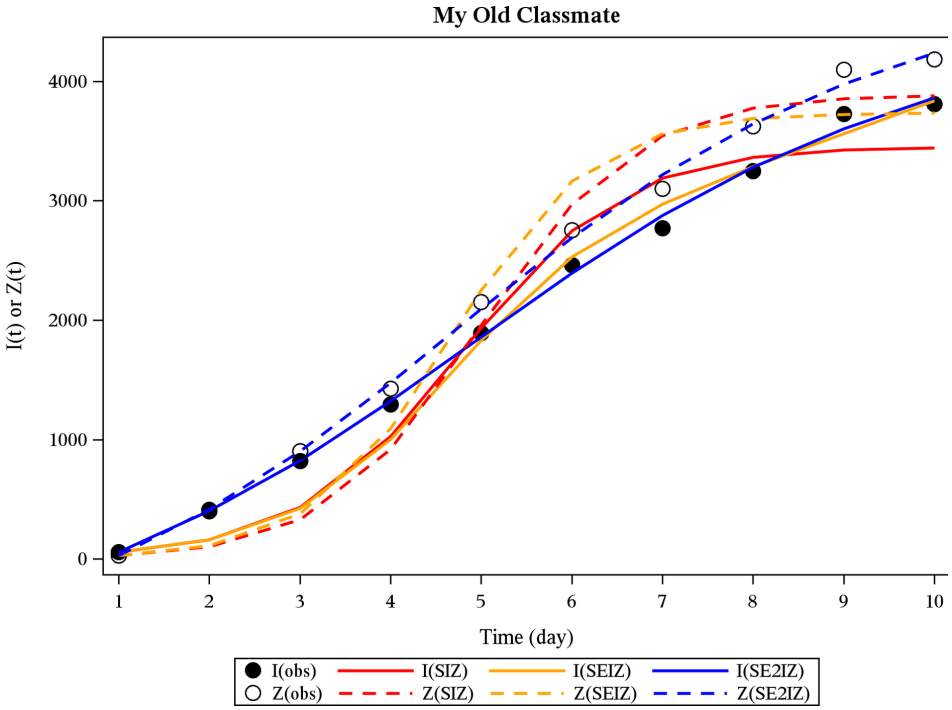


Figure 11. Observed data and fitted curves for "My Old Classmate" movie



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 12. Observed data and fitted curves for national civil service

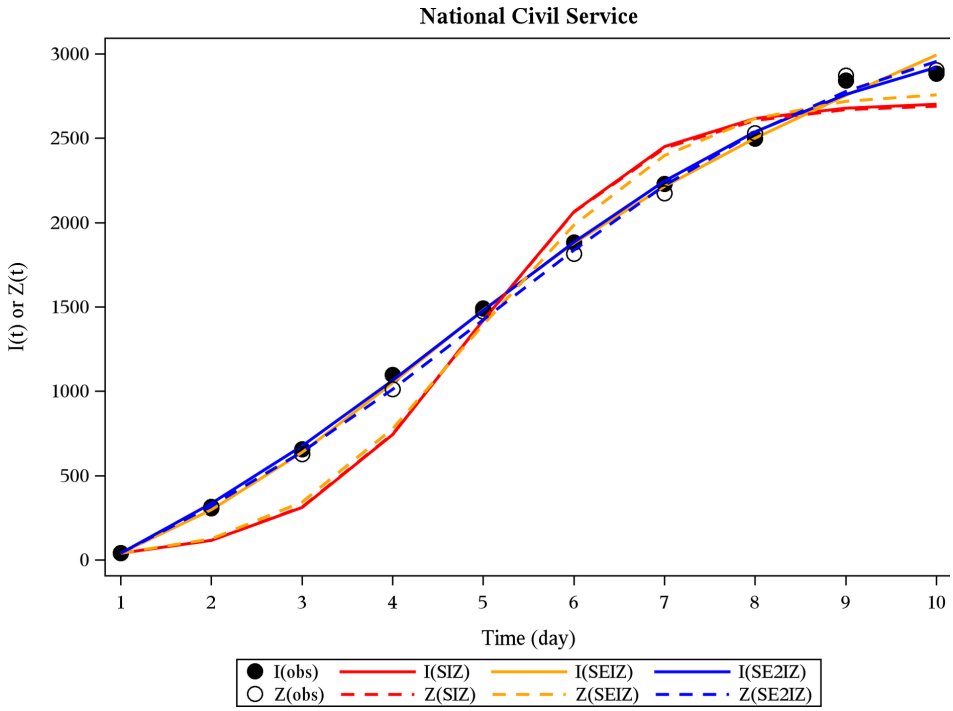
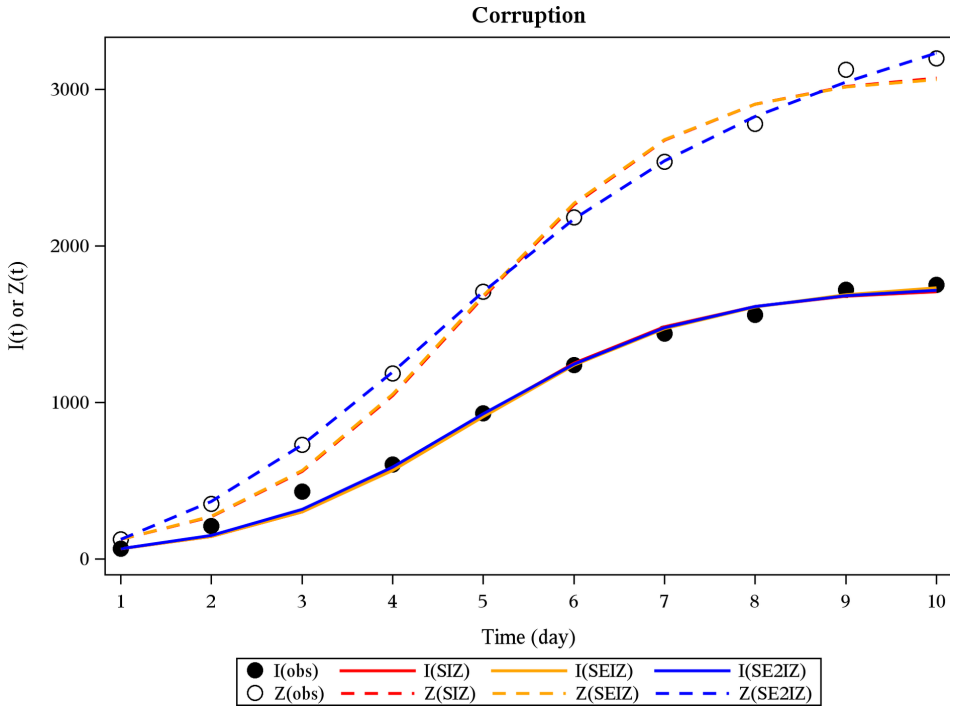
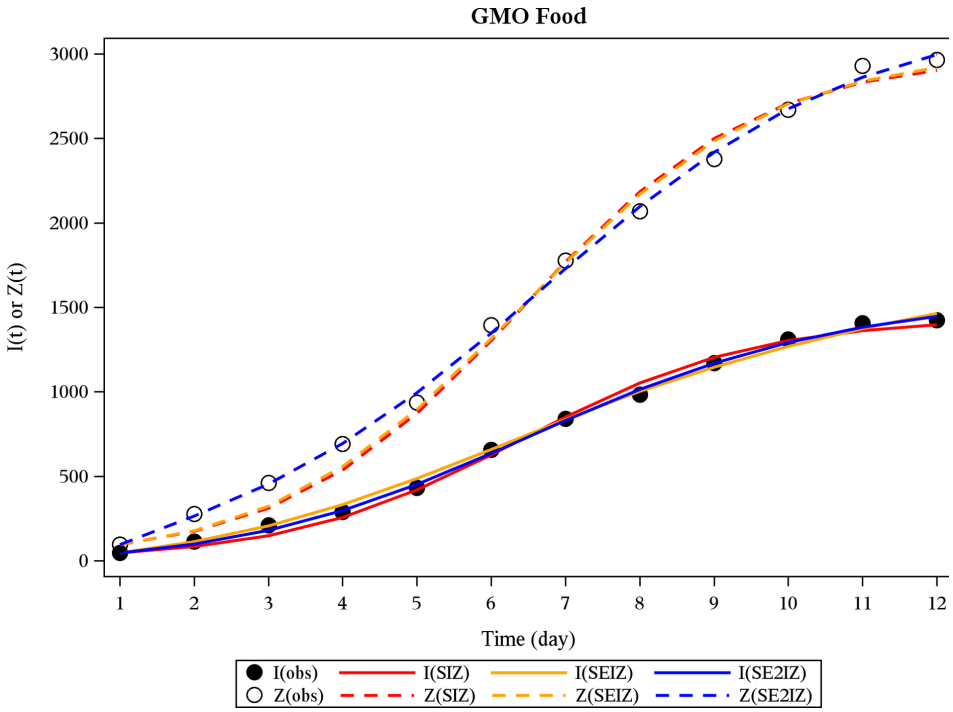


Figure 13. Observed data and fitted curves for corruption



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 14. Observed data and fitted curves for GMO food



There are two limitations in this study. First, natural language processing and sentiment analysis may identify some emotions incorrectly due to the tricks in human language. For example, "He will continue to be the coach of Houston Rocket Team for the next season, congratulations to the other 31 teams" is ironic about the poor performance of the head coach of Houston Rocket team. But sentiment analysis cannot recognize the language trick and incorrectly assigned a positive emotion to this sentence based on the positive keyword "congratulations". Second, this study did not distinguish the social impacts of different microblogging users. Public figures and celebrities have much higher social impacts than ordinary people, and thus their messages and opinions typically have much higher influences during propagation. Some promising directions for future study include: improving sentiment analysis to recognize human language tricks, assigning different weights to microblogging users based on their social influence, and incorporating other public-accessible data resources (e.g. other types of social media) to assess and predict the trends of public opinions with higher accuracy.

ACKNOWLEDGMENT

Youjia Fang is a corresponding author for this article and can be reached at youjia@vt.edu.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

REFERENCES

Antin, J., Guo, F., Fang, Y., Dingus, T. A., Perez, M., & Hankey, J. (2016). A validation of low mileage bias using Naturalistic Driving Study data. *Accident; Analysis and Prevention*.

Bettencourt, L. M., Cintrón-Arias, A., Kaiser, D. I., & Castillo-Chávez, C. (2006). The power of a good idea: Quantitative modeling of the spread of ideas from epidemiological models. *Physica A: Statistical Mechanics and its Applications, 364*, 513-536.

Brin, S., Motwani, R., Ullman, J. D., & Tsur, S. (1997). Dynamic itemset counting and implication rules for market basket data. *Paper presented at the ACM SIGMOD Record*. 10.1145/253260.253325

Canneyt, S. V., Schockaert, S., & Dhoedt, B. (2015). Discovering and characterizing places of interest using Flickr and Twitter. *International Journal on Semantic Web and Information Systems*, *9*(3), 77–104. doi:10.4018/ijswis.2013070105

Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, P. K. (2010). Measuring user influence in twitter: The million follower fallacy. *ICWSM*, *10*, 10–17.

Chen, G. X., Fang, Y., Guo, F., & Hanowski, R. J. (2016). The influence of daily sleep patterns of commercial truck drivers on driving performance. *Accident; Analysis and Prevention*, *91*, 55–63. doi:10.1016/j.aap.2016.02.027 PMID:26954762

Chen, X., Fang, Y., & Lockee, B. (2015). Integrative review of social presence in distance education: Issues and challenges. *Educational Research Review*, *10*(13), 1796–1806. doi:10.5897/ERR2015.2276

Datatang. (n. d.). www.datatang.com

Escamilla, I., Torres-Ruiz, M., Moreno-Ibarra, M., Quintero, R., Guzmán, G., & Luna-Soto, V. (2016). Geocoding tweets approach based on conceptual representations in the context of the knowledge society. *International Journal on Semantic Web and Information Systems*, *12*(1), 44–61. doi:10.4018/IJSWIS.2016010103

Fitch, G. M., Soccolich, S. A., Guo, F., McClafferty, J., Fang, Y., Olson, R. L., & ... . (2013). The impact of hand-held and hands-free cell phone use on driving performance and safety-critical event risk.

Gao, Q., Abel, F., Houben, G.-J., & Yu, Y. (2012). *A comprehensive study of users' microblogging behavior on Sina Weibo and Twitter User modeling, Adaptation, and Personalization*. Berlin, Heidelberg: Springer.

García-Silva, A., Rodríguez-Doncel, V., & Corch, O. (2013). Semantic Characterization of Tweets Using Topic Models: A Use Case in the Entertainment Domain. *International Journal on Semantic Web and Information Systems*, *9*(3), 1–13. doi:10.4018/ijswis.2013070101

Gruhl, D., Guha, R., Liben-Nowell, D., & Tomkins, A. (2004). Information diffusion through blogspace. *Proceedings of the 13th international conference on World Wide Web* (pp. 491-501).

Guo, F., & Fang, Y. (2013). Individual driver risk assessment using naturalistic driving data. *Accident; Analysis and Prevention*, *61*, 3–9. doi:10.1016/j.aap.2012.06.014 PMID:22785091

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Guo, F., Fang, Y., & Antin, J. F. (2014). *Evaluation of older driver fitness-to-drive metrics and driving risk using naturalistic driving study data.* National Surface Transportation Safety Center for Excellence.

Guo, F., Fang, Y., & Antin, J. F. (2015). Older driver fitness-to-drive evaluation using naturalistic driving data. *Journal of Safety Research*, *54*, 49–54. doi:10.1016/j.jsr.2015.06.013 PMID:26403900

Guo, F., Klauer, S. G., Fang, Y., Hankey, J. M., Antin, J. F., Perez, M. A., ... Dingus, T. A. (2016). The effects of age on crash risk associated with driver distraction. *International Journal of Epidemiology*, dyw234. doi:10.1093/ije/dyw234

Hethcote, H. W. (2000). The mathematics of infectious diseases. *SIAM Review*, *42*(4), 599–653. doi:10.1137/S0036144500371907

Jin, F., Dougherty, E., Saraf, P., Cao, Y., & Ramakrishnan, N. (2013). Epidemiological modeling of news and rumors on twitter. *Proceedings of the 7th Workshop on Social Network Mining and Analysis.* ACM. 10.1145/2501025.2501027

Jin, F., Wang, W., Zhao, L., Dougherty, E., Cao, Y., Lu, C.-T., & Ramakrishnan, N. (2014). Misinformation propagation in the age of twitter. *Computer*, *47*(12), 90–94. doi:10.1109/MC.2014.361

Khelil, A., Becker, C., Tian, J., & Rothermel, K. (2002). An epidemic model for information diffusion in MANETs. *Proceedings of the 5th ACM International Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems* (pp. 54-60). 10.1145/570758.570768

Ku, L. W., & Chen, H. H. (2007). Mining opinions from the Web: Beyond relevance retrieval. *Journal of the American Society for Information Science and Technology*, *58*(12), 1838–1850. doi:10.1002/asi.20630

Kwon, A.-R., & Lee, K.-S. (2013). Opinion bias detection with social preference learning in social data. *International Journal on Semantic Web and Information Systems*, *9*(3), 57–76. doi:10.4018/ijswis.2013070104

Lebo, M. J., & Norpoth, H. (2007). The PM and the pendulum: Dynamic forecasting of British elections. *British Journal of Political Science*, *37*(1), 71–87. doi:10.1017/S000712340700004X

Lerman, K., & Ghosh, R. (2010). Information contagion: an empirical study of the spread of news on Digg and Twitter social networks. *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media* (pp. 90-97).

Matsubara, Y., Sakurai, Y., Prakash, B. A., Li, L., & Faloutsos, C. (2012). Rise and fall patterns of information diffusion: model and implications. *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 6-14). 10.1145/2339530.2339537

Newman, M. E. (2002). Spread of epidemic disease on networks. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *66*(1), 016128. doi:10.1103/PhysRevE.66.016128 PMID:12241447

NLPIR. (n. d.). Retrieved from http://ictclas.nlpir.org/

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ouimet, M. C., Brown, T. G., Guo, F., Klauer, S. G., Simons-Morton, B. G., Fang, Y., & ... . (2014). Higher crash and near-crash rates in teenage drivers with lower cortisol reactivity: An 18-month longitudinal, naturalistic study. *JAMA Pediatrics*, *168*(6), 517–522. doi:10.1001/jamapediatrics.2013.5387 PMID:24710522

Romero, D. M., Meeder, B., & Kleinberg, J. (2011). Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. *Proceedings of the 20th International Conference on World Wide Web* (pp. 695-704). 10.1145/1963405.1963503

Shuai, X., Ding, Y., Busemeyer, J., Chen, S., Sun, Y., & Tang, J. (2012). Modeling indirect influence on twitter. *International Journal on Semantic Web and Information Systems*, *8*(4), 20–36. doi:10.4018/jswis.2012100102

Trpevski, D., Tang, W. K., & Kocarev, L. (2010). Model for rumor spreading over networks. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, *81*(5), 056102. doi:10.1103/PhysRevE.81.056102 PMID:20866292

Waller, L. A., & Gotway, C. A. (2004). *Applied spatial statistics for public health data*. John Wiley & Sons. doi:10.1002/0471662682

Weng, J., Lim, E.-P., Jiang, J., & He, Q. (2010). Twitterrank: finding topic-sensitive influential twitterers. *Proceedings of the Third ACM International Conference on Web Search and Data Mining* (pp. 261-270). 10.1145/1718487.1718520

Xiong, F., Liu, Y., Zhang, Z., Zhu, J., & Zhang, Y. (2012). An information diffusion model based on retweeting mechanism for online social media. *Physics Letters. [Part A]*, *376*(30-31), 2103–2108. doi:10.1016/j.physleta.2012.05.021

Yang, J., & Counts, S. (2010). Predicting the speed, scale, and range of information diffusion in twitter. *Proceedings of ICWSM* (pp. 355-358).

Yang, J., & Leskovec, J. (2010). Modeling information diffusion in implicit networks. *Proceedings of the 2010 IEEE 10th International Conference on Data Mining (ICDM)* (pp. 599-608). IEEE.

This research was previously published in the International Journal on Semantic Web and Information Systems (IJSWIS), 13(1); edited by Miltiadis D. Lytras; pages 11-27; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 46
Artificial Intelligence in Stochastic Multiple-Criteria Decision Making

Hanna Sawicka
Poznan University of Technology, Poland

ABSTRACT

This chapter presents the concept of stochastic multiple criteria decision making (MCDM) method to solve complex ranking decision problems. This approach is composed of three main areas of research, i.e. classical MCDM, probability theory and classification method. The most important steps of the idea are characterized and specific features of the applied methods are briefly presented. The application of Electre III combined with probability theory, and Promethee II combined with Bayes classifier are described in details. Two case studies of stochastic multiple criteria decision making are presented. The first one shows the distribution system of electrotechnical products, composed of 24 distribution centers (DC), while the core business of the second one is the production and warehousing of pharmaceutical products. Based on the application of presented stochastic MCDM method, different ways of improvements of these complex systems are proposed and the final i.e. the best paths of changes are recommended.

INTRODUCTION

Nowadays, an integral part of any organization is the application of practical tools and techniques that make changes to products, processes and services resulting in an introduction of something new in a market. Improving continuously operations is essential for a better performance of the organizations. Due to the growing impact of globalization, migration, technological and knowledge revolutions, it brings added value to customers and helps organizations to remain competitive. Many of them are forced to make changes in various areas, such as: technology, infrastructure, human resources etc. Moreover, to meet customers' expectations and needs, companies should also consider organization's manager interests, as well as the supplier's opinions, customers' point of view and the other stakeholders' preferences. It makes the decision situation very complex and the application of decision aiding methods seems crucial.

DOI: 10.4018/978-1-7998-0951-7.ch046

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

According to Vincke (1992) multiple criteria decision making (MCDM) is a field which aims at giving the decision maker (DM) some tools in order to enable him/her to solve a complex decision problem where several points of view must be taken into account. This methodology concentrates on suggesting "compromise solution", taking into consideration the trade-offs between criteria and the DM's preferences. The above mentioned compromise solution is selected from the family of variants. They are constructed in different ways. In some situations, it is assumed that the variants are exclusive and at least two of them cannot be implemented together. There are also real-world situations where two or more alternatives can be introduced conjointly.

The variants are evaluated by the set of criteria, which should be characterized by the following aspects (Roy, 1985):

- Completeness due to the decision-making aspects of the considered problem.
- Appropriate formation, taking into account the global preferences of the decision maker.
- Non-redundancy, i.e. a situation in which semantic ranges of criteria are not repeated.

Thanks to the criteria it is possible to compare variants, especially when the performances are expressed as deterministic values. However, in some cases the alternatives are modeled e.g. in a simulation tool and their performances are presented as stochastic values. In such circumstances the comparison process becomes complex. It is usually supported by stochastic MCDM methods, but most of them concentrate on decision maker's stochastic preferences. Based on the author's experience the methods dedicated to solve complex decision problems with stochastic criteria values are not efficient enough.

The procedure presented in this chapter shows that the combination of a classical group of MCDM methods aiming at ranking of variants, e.g. Electre III, Promethee II, with probability formula or a classical method of artificial intelligence aiming at classification of objects, e.g. Bayes classifier, could solve these complex problems.

Electre III method (Roy, 1985; Vincke, 1992) belongs to the European school of MCDM and it is based on the outranking relation. The method requires determination the model of DM's preferences by the indifference q, preference p, and veto v thresholds, as well as weights w for each criterion j. The aggregation procedure starts from the calculation of concordance and discordance indices. The first one measures the arguments in favor of the statement that alternative a outranks alternative b, while the second index represents the strengths of evidence against the above hypothesis. Based on these indices the outranking relation is calculated. The ranking of variants is based on two classification algorithms: descending and ascending distillations. Descending distillation procedure starts from choosing the best alternative i.e. the one with the highest value of qualification index $Q(a)$ and placing it at the top of the ranking. $Q(a)$ equals the difference between the number of alternatives, which are outranked by the alternative a and the number of alternatives that outrank a. In the consecutive steps, the best alternative from the remaining set of alternatives is selected and placed in the second position of the ranking. The procedure stops, when the set of alternatives is empty. Ascending distillation procedure starts from choosing the worst alternative and placing it at the bottom of the ranking. Then the worst alternative from the remaining set of alternatives is selected and placed in the second worst position of the ranking. The final graph corresponding to the outranking matrix is the intersection of the two distillations. It constitutes a graphical representation of indifference I, preference P and incomparability R relations between alternatives.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Promethee II method (Brans et al., 1986), similarly as Electre III method, represents the European school of MCDM based on the outranking relation. The method requires determination of the model of DM's preferences by the selection of a type of a generalized criteria function $\rho(x)$ for each criterion j and definition of this function characteristic parameters. There are distinguished six types of preference functions, widely described by Brans et al. (1986). They represent the level of preference intensity of alternative a compared with alternative b regarding each criterion j. The criteria function equals from 0 to 1 and its value increases when the difference between compared alternatives is getting higher. Next, the outranking relation is defined. It is calculated on the basis of the multicriteria preference index π of alternative a regarding alternative b. This index corresponds to concordance index in Electre III method. Finally, the positive $\Phi+$ and negative $\Phi-$ flows for each alternative are calculated. The positive value of $\Phi+$ represents the strength of this alternative over all the other alternatives, while the negative value of $\Phi-$ shows the strengths of all alternatives over considered alternative. The difference between $\Phi+$ and $\Phi-$ constitutes the net flow. The higher the net value, the better position in the ranking. Based on this information a graphical representation of indifference I and preference P relations between alternatives is obtained.

Classification theory (Smola & Vishwanatan, 2008; Demsar, 2006) is a quantitatively oriented methodology that helps the DM to classify and categorize objects into predefined classes. There is a wide variety of classification methods reported in the literature (Berger, 1985; Demsar, 2006; Mitchell, 1997), including:

- Classification by the induction of decision trees (Pisetta et al., 2010; Quinlan, 1986).
- K-NN – nearest-neighbor classification (Bremner et al., 2005; Hall et al., 2008).
- Bayesian classification (Berger, 1985; Mitchell, 1987; Smola & Vishwanatan, 2008).
- Rough sets theory (Pawlak, 1982; Yao, 2011).
- Neural networks (Cetiner, et al. 2010; Wei & Schonfeld, 1993).

Many authors (Berger, 1985; Mitchell, 1987) claim that one of the most efficient method to solve real-world classification problems is Bayesian classification. Its advantage is precision and effectiveness when solving problems associated with the manipulation on the large sets of data.

Mitchell (1997) states that Bayes classifier can predict class membership probabilities, such as the probability that an unknown target function $f: x(a) \rightarrow C(z)$, featured by the highest conditional a posteriori probability $P(C(z)|x(a))$ belongs to a particular class $C(z)$. This notation can be interpreted as follows: a certain object x described by the vector of description attributes a, has a chance (probability) $P(C(z)|x(a))$ to be assigned to a certain class $C(z)$.

The algorithm of Bayesian classification is composed of 3 phases, such as (Mitchell, 1997):

- The construction of learning model i.e. classifier.
- Testing of the learning model.
- Classification of the vector of observations to predefined classes.

In the proposed approach this machine learning method is utilized to classify the relations between variants and to generate the final stochastic ranking of variants.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In this chapter the combination of MCDM ranking method with Bayes classifier, and MCDM ranking method with classical probability formula are presented on real-world cases. The first one shows the distribution system composed of 24 distribution centers divided into 3 levels i.e. central, regional and local one. Different alternatives representing ways of changes are proposed. They are modeled in the simulation tool and evaluated by the consistent family of criteria. Their values are calculated on the basis of simulation experiments' results. They are presented as probability distributions. The decision problem is formulated as a multicriteria stochastic ranking problem. The above-described stochastic approach combining Promethee II method (Brans et al., 1986) and the probability formula is applied to solve it.

The second case presents pharmaceutical production company. Similarly, as in the previous case, the simulation model of the current state has been constructed and its redesign scenarios have been proposed and implemented in the simulation tool, as well. The set of evaluation criteria includes technical, organizational and social aspects. The values are presented as probability distributions. The decision problem is formulated as a multicriteria stochastic ranking problem. The stochastic approach combining Electre III method (Roy, 1985; Vincke, 1992) and Bayes classifier (Mitchell, 1997) is applied to solve it. In both cases, the computational results are completed by the sensitivity analysis.

BACKGROUND

Artificial intelligence (AI) is defined in two dimensions, i.e. the first one is a thought process and reasoning – an approach presented in the middle of 80's, while the second one is a behavior – many definitions were created at the beginning of 90's (Russel, & Norvig, 1995). Some of the researchers perceive this term as a field of study that seeks to explain and emulate intelligent behavior in terms of computational processes (Schalkoff, 1990).

Undoubtedly, since it beginning the goal of AI is to make machines do thing that would require intelligence if done by humans (Boden, 1977). In this context, intelligence can be defined as the ability to learn and understand, to solve problems and to make decisions (Negnevitsky, 2005). One of the topics of decision making widely described in the literature and applied in practice is multi criteria decision making (MCDM). Kahraman (2008) states that MCDM can be divided into two major approaches, i.e. multiple objective decision making (MODM) and multiple attribute decision making (MADM). In MODM problems the number of alternatives is usually large or sometimes it might be infinite. The alternatives are not given at the beginning of the decision process, but mathematical framework is provided for designing the set of them. Based on it the identified alternatives are evaluated and a level of multiple objective satisfactions is measured.

MADM problems, which are the most well-known within MCDM, are characterized by limited number of alternatives defined at the beginning of the decision process, by a number of decision criteria and additional information from decision maker or decision makers. The final solution of the MADM problem can be presented as a choice, sorting or ranking of alternatives. Based on this division, corresponding to the decision problem, MADM methods are classified into the following three groups:

- Choice methods, e.g. Genetic Algorithm (Goldberg, 1989), Pareto Simulated Annealing (Czyzak & Jaszkiewicz, 1998);
- Sorting methods, e.g. ELECTRE-TRI (Mousseau et al., 2000), UTADIS (Manshadi, 2015);

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Ranking methods, e.g. AHP (Saaty, 1980), Electre III (Roy, 1985; Vincke 1992), Promethee (Brans et al., 1986), UTA (Jacquet-Lagrèze & Siskos, 1982).

Many real-world problems, due to their complexity, are characterized by uncertainty. This phenomenon can be observed when the values of evaluation criteria are considered or the decision maker's preferences are collected. Thus, there are many modifications of the above-mentioned ranking methods, aiding the decision process. The examples of them are such as:

- SMAA-III (Tervonen & Figueira, 2008) – a modification of Electre III method.
- Stochastic PROMETHEE / GIS method (Marinoni, 2005) – a modification of Promethee method.
- Stochastic AHP method (Stam & Duarte Silva, 1997) – a modification of AHP method.
- Stochastic UTA method (Siskos, 1982) – a modification of UTA method.

SMAA-III

Stochastic Multicriteria Acceptability Analysis-III (Tervonen & Figueira, 2008) is a stochastic method of ranking alternatives based on outranking relation. It develops Electre III method by the non-deterministic components of the criteria values and decision maker's preferences. The values of criteria functions ξ_j corresponding to deterministic values $f_j(x)$ of density function $f\chi(\xi)$ in the space $\chi \subseteq R^{mxn}$, are presented as a probability distribution or expected value with its ranges of variations. Weights are also described by the probability distribution with the density function $f_W(w)$. Indifference q, preference p and veto v thresholds characteristic for Electre III method are defined by the decision maker and they are stochastic values represented by $\alpha_j(\cdot)$, $\beta(\cdot)$, $\gamma(\cdot)$ or by $\tau = (\alpha, \beta, \gamma)$. The issue of SMAA-III method is to aid the decision maker by computing the indices or paramters with the application of integrals of the area of stochastic values using Monte Carlo simulation method.

Based on the literature (Lahdelma et al., 1998; Tervonen & Figueira, 2008) there are presented three indices aiding the decisions, which are as follows: rank acceptability index, pair-wise winning index and incomparability index.

The computational procedure of SMAA-III method is composed of the three steps, presented below:

Step 1: Construction of the evaluation matrix and definition of the decision maker's preference model.
Step 2: Construction of the outranking relation based on the pair-wise winning index.
Step 3: Utilization of the outranking relation leading to the construction of the final ranking of alternatives based on the rank acceptability index and the incomparability index.

Stochastic Promethee

This method belongs to the group of MADM ranking methods (Marinoni, 2005). The values of criteria are presented as probability distributions. Marinoni (2005) suggests to use Monte Carlo simulation method to multiple sampling of the values of alternatives evaluated by each criterion, resulting in accumulated value ranging [0,1]. The number of iterations (samplings) should be as high as possible, i.e. at least 100 (Marinoni, 2005). The obtained results are the input data for further calculations carried out accordingly to Promethee method (Brans et al., 1986). They are presented as charts with alternatives'

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

values related to each final rank or as rankings of each alternative. The values of alternatives may differ and the selection of the compromise solution might be a difficult task. Thus, Marinoni (2005) suggests the application of MSR index - mean stochastic rank index, calculated for each alternative *i*. Then, the stochastic ranking index *SI* is computed (Marinoni, 2005).

The computational procedure of stochastic Promethee method is composed of the three steps, which are as follows:

Step 1: Construction of the evaluation matrix and definition of the decision maker's preference model.
Step 2: Construction of the outranking relation based on the stochastic ranking index.
Step 3: Utilization of the outranking relation leading to the construction of the final ranking of alternatives.

The above presented MADM methods are applied in many areas to solve problems connected with environmental issues (Delhaye et al., 1991; Malczewski et al., 1997; Salminen et al., 1998), industrial problems (Anand & Kodali, 2008; de Boer et al., 1998; Hafeez et al., 2002), business issues (Buchanan & Sheppard 1998; Halouani et al., 2009; Ngai, 2003), logistics problems (Badri, 1999; Behzadian et al., 2010; Sawicka et al., 2010).

They have many advantages, such as: precise modeling of decision maker's preferences (e.g. Electre III method), modeling of preferential function of each criterion (e.g. Promethee method), hierarchical representation of the problem giving the possibility to formalize the structure of the considered problem, very useful while the complex problems are considered (e.g. AHP method), possibility to solve the problem with a large number of alternatives i.e. more than 10 (e.g. UTA method) possibility to model the stochastic criteria values (e.g. SMAA-III method), possibility to model uncertain preferences of the decision maker (e.g. stochastic Promethee method, stochastic AHP method, stochastic UTA method).

However, these methods are not devoid of disadvantages, e.g. adding or removing of one alternative could change the power of preferences between remaining variants, resulting from the distillation procedure in Electre III method; solving very complex decision problems is time consuming and many information must be collected (e.g. SMAA-III method, AHP mathod) or many iterations must be carried out (e.g. stochastic AHP method), some methods do not provide modeling of subcriteria (e.g. Promethee and stochastic Promethee methods), to construct reference set of alternatives, their number must be higher than 10 (e.g. UTA method) and finally to define values of criteria the probability distributions of each of them must be calculated, which is time consuming (e.g. stochastic UTA method).

THE METHODOLOGY OF SOLVING STOCHASTIC RANKING MADM PROBLEMS

Solving complex multiple criteria decision problems is often very difficult. On the one hand, some of the MADM methods are not suitable for the considered issues e.g. the solution of stochastic problems with an application of deterministic methods would fail or simplify the question. However, many analysts or decision makers are very used to one or a group of such methods. This subjective perspective is a result of their wide experience and knowledge.

On the other hand, the multiple criteria methods dedicated to solve particular class of problems could not always cope with them due to their specific characteristics.

The author of this chapter proposes the MADM method to solve complex ranking decision problems with stochastic information of criteria values and/or stochastic preferential information. The general idea

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

is based on the transformation of the stochastic input information to deterministic values, which are then converted to stochastic final ranking of variants. The transformation encompasses the application of multiple criteria decision aiding methods and classification method or probability formula.

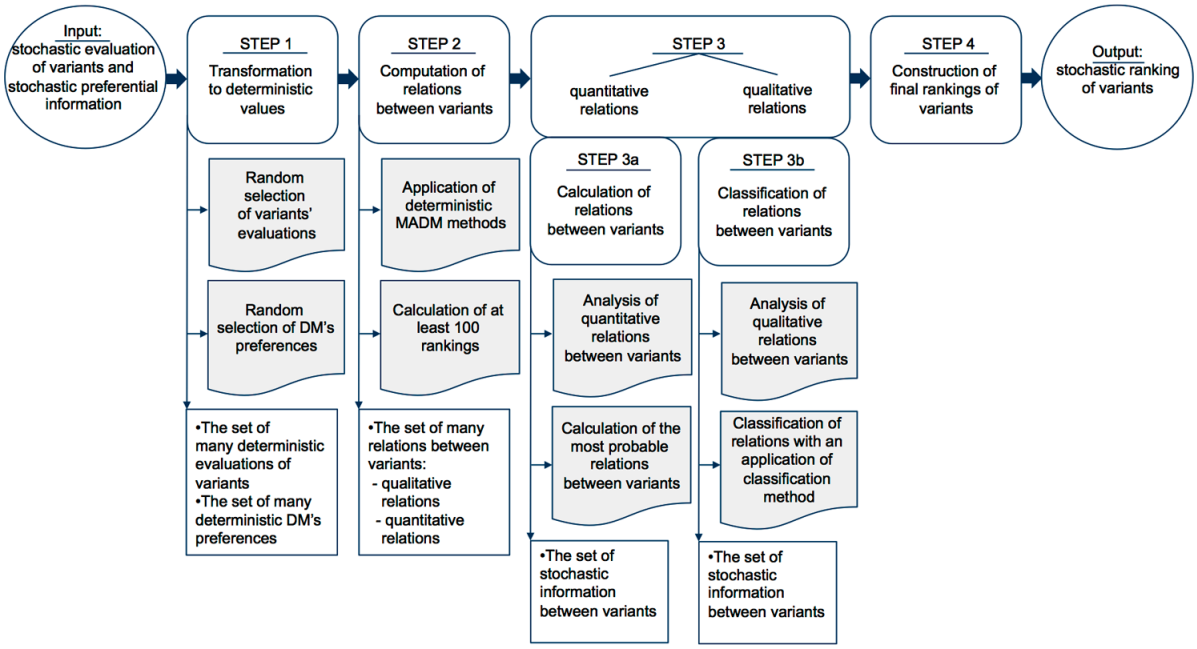
It is composed of 4 steps presented in Figure 1 and characterized below.

It is assumed that the character of the input information is fully stochastic or mixed - stochastic and deterministic. It usually happens when many aspects of the analyzed situation are considered and modeled in the simulation tool. As a result the evaluation of alternatives is stochastic and the decision maker expresses his/her hesitation of preferences. In such situations DM's preferences are defined as non-deterministic values.

The first step of the proposed procedure is the transformation of these values to deterministic information. If the analyzed problem is composed of stochastic evaluations of alternatives, then each of them is converted to at least 100 deterministic values using random number generator. One of the well-known is Monte Carlo method (Law & Kelton, 2000), which relies on repeated random sampling. The parallel operation can be done with stochastic preferential information, e.g. values of thresholds in Electre III method, pairwise comparisons between criteria, subcriteria and alternatives in AHP method. As a result, the set of many deterministic evaluations of alternatives and/or deterministic DM's preferences is obtained.

Based on them it is possible to calculate the relations between alternatives with an application of MADM ranking method – the second step. The number of rankings is at least 100. The type of relation between alternatives, i.e. quantitative or qualitative, depends on the applied method. Some of them, e.g. Promethee II, AHP methods provide precise distance between alternatives in the final rank, while other methods, e.g. Electre III, Promethee I present the qualitative relations, such as indifference, preference and incomparability.

Figure 1. Solution procedure of the stochastic multiple criteria decision problem
Source: Author's research based on Sawicka (2012)



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Thus, the third step is divided into two alternate parts *a* and *b*. In the step 3a the quantitative measures between alternatives are analyzed and the most probable relations are calculated. The step 3b is dedicated to qualitative relations between pairs of variants. Due to high number of results obtained in the previous step, the classification method is proposed. It helps to classify relations between alternatives to predefined classes (decision attributes). The result of both cases is the set of stochastic information between alternatives.

The fourth step is based on construction of the final ranking of variants. The probabilities of occurrence of particular relations between them are computed. The sensitivity analysis is advised. Finally, the stochastic ranking of alternatives is obtained.

SOLUTIONS AND RECOMMENDATIONS: APPLICATION OF STOCHASTIC MADM RANKING METHOD

Ranking of Distribution System Redesign Alternatives Using Promethee II Method and Probability Formula

Problem Definition

The first case study shows the distribution system of electrotechnical products. It is characterized by a high complexity. The system is composed of 24 distribution centers (DC) divided into 3 levels i.e. central with 1 DC, regional with 12 DC and a local one with 11 DC spread all over Poland. More than 75 suppliers supply these distribution centers and around 400 main customers are served. There are distributed various electrotechnical products, such as sockets, bulbs, extension cords, clusters, wires. There can be distinguished 38.000 assortments. The detailed analysis of the system reveals many strengths e.g. modern fleet and handling equipment, well developed promotion, highly qualified staff, as well as weaknesses i.e. high level of inventories, repetition of tasks in different organizational units, low utilization of the warehouse areas. Thus, different ways of changes based on system's redesign are proposed. Due to its complex character, the simulation model of a current state is constructed. It allows formalizing material and information flow in the system. Based on the simulation experiments the bottlenecks are recognized and quantitative analysis is carried out. The final information of the system is presented as a set of evaluation criteria. They represent various aspects i.e. technical, organizational, economical, social and environmental and the criteria are as follows:

C1: Time of delivery [days] – minimized criterion,
C2: Costs of distribution per month [m PLN] – minimized criterion,
C3: Utilization of in-company transportation means [%] – maximized criterion,
C4: Rotation level of inventory [days] – minimized criterion,
C5: Efficiency of system's employees (deviation from the optimum value) [%] – minimized criterion,
C6: Share of outsourced transportation orders [%] – maximized criterion,
C7: Difference between the levels of investments and divestments [m PLN] – minimized criterion.

Next, the distribution system's heuristically constructed 4 redesign scenarios are modeled in the simulation tool. They represent different changes resulting from the following aims:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Alternative A1: Reduction of distribution costs by eliminating the least effective distribution center; increasing the turnover, improving efficiency ratio;

Alternative A2: Reduction of the delivery time by shortening the distance between the distribution centers and customers, increasing the turnover, improving efficiency ratio connected with human and technical resources;

Alternative A3: Reduction of distribution costs by eliminating 11 local DC, reduction of the delivery time;

Alternative A4: Reduction of distribution costs by eliminating 11 local DC, reduction of inventory on regional level, reallocation of central DC, reduction of the delivery time.

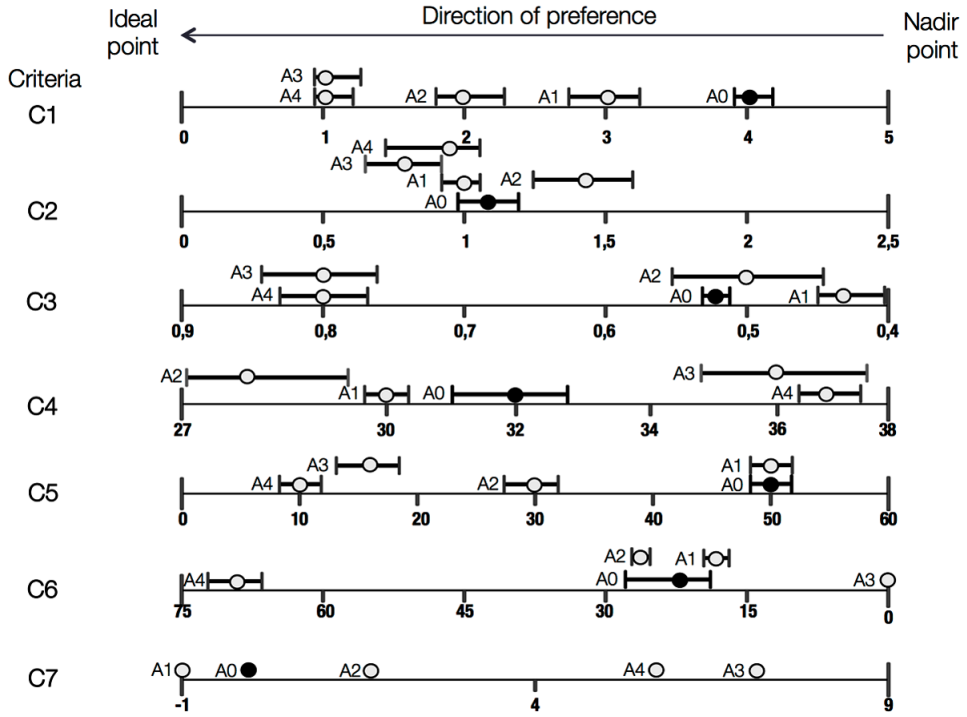
The consistent family of criteria evaluates the alternatives, including the current state represented by alternative A0. The criteria values are calculated on the basis of simulation experiments' results. They are presented as probability distributions. See Table 1.

Table 1. Performance matrix of the distribution system of electrotechnical products and its redesign scenarios

Criteria			Alternatives				
Name	**Direction of preferences**	**Value**	**A0**	**A1**	**A2**	**A3**	**A4**
C1	Min	Expected value	4	3	2	1	1
		Range of variation	[3,87; 4,13]	[2,76; 3,20]	[1,78; 2,30]	[0,95; 1,25]	[0,95; 1,13]
C2	Min	Expected value	1,1	1,0	1,4	0,8	0,9
		Range of variation	[0,95; 1,17]	[0,87; 1,05]	[1,24; 1,60]	[0,65; 0,87]	[0,72; 1,04]
C3	Max	Expected value	0,52	0,43	0,50	0,80	0,80
		Range of variation	[0,51; 0,53]	[0,41; 0,45]	[0,44; 0,57]	[0,76; 0,85]	[0,77; 0,83]
C4	Min	Expected value	32	30	28	36	37
		Range of variation	[31,33; 32,67]	[29,84; 30,24]	[27,19; 29,49]	[34,96; 37,44]	[36,68; 37,24]
C5	Min	Expected value	50	50	30	15	10
		Range of variation	[48,80; 52,00]	[48,80; 52,00]	[28,09; 31,51]	[12,57; 17,43]	[8,80; 12,00]
C6	Max	Expected value	21,2	18,6	22,7	0	70
		Range of variation	[19,14; 23,34]	[17,95; 19,17]	[22,54; 22,90]	-	[68,37; 71,71]
C7	Min	Average value	0	-1,0	2,7	7,5	5,8

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Alternatives' positions regarding evaluation criteria



The graphical representation of the alternatives' distance to the ideal point (represented by the most desirable criteria values) and the nadir point (characterized by the most undesirable criteria values) (Vincke, 1992) is shown in Figure 2.

Based on the information presented in Table 1 and Figure 2, it is hard to decide which alternative is the compromise solution. The following aspects have the most important influence on this situation:

- Ambiguity of alternatives' evaluation i.e. some of them have the best values on one criterion while on the other they are close to the worst position, e.g. alternative A4 has the best value on criterion C6 and its position with respect to criterion C4 is the worst;
- Uncertainty of the alternatives' mutual position regarding each criterion i.e. in some criteria their ranges of variations are almost evenly distributed between the ideal and nadir point e.g. positions of variants regarding criterion C1, while on the other criteria part of the set of alternatives is cumulated around one value e.g. alternatives A0, A1, A2 regarding criterion C3 and distant from the other part of variants e.g. A3 and A4;
- Difficulty in evaluating such a large collection of information i.e. many criteria and alternatives, many data records, different levels of criteria importance expressed by the DM.

Based on these ambiguities, uncertainties and difficulties with interpretation of data the multicriteria stochastic ranking problem has been formulated. The above-described stochastic approach combining Promethee II method and the probability formula is applied to solve it.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Computational Experiments

In the first step of the computational procedure, decision maker's preferences are modeled according to the principles of Promethee II method. The type 5 of the generalized criteria function ρ(x), which is characteristic for this method, is selected. The DM has assigned different weights to criteria, using the scale of 1 to 10 points, representing the least and the most important criteria, respectively. They are presented in Table 2. Based on the model of the DM's preferences one can conclude that the most important criterion is time of delivery – C1, while the least important one is the share of outsourced transportation orders – C6.

The definition of thresholds is a complex task. Thus, the analysis of difference between criteria values has been carried out based on the expected values and ranges of variation, too. The DM has decided to assign the values of q (indifference threshold) and p (preference threshold), using the information presented in Table 1.

The results indicate that the DM is highly sensitive to changes of some values of criteria e.g. C2 (one of the most important criterion), where the indifference threshold equals 0,05 and preference threshold equals 0,10; and his sensitivity to changes is rather low regarding other criteria e.g. C6 (the least important criterion), where the indifference threshold equals 3,00 and preference threshold equals 5,00.

Next, the random numbers of each criterion value have been generated. For this purpose, the Monte Carlo method has been applied. The number of iterations generating random numbers of criteria evaluating each alternative equals 150 and it provides large sample of random deterministic parameters. The example of results is presented in Table 3.

Based on the data records presented in Table 3 one can conclude that the generated values are randomly dispersed in the ranges of variations and some values equal the minimum or maximum value of these ranges.

The results of iterations are the input data for the computational experiments (second step) with an application of MADM ranking method Promethee II. Thus, the total number of calculations equals 150.

The exemplary 3 results are presented in Figure 3.

The additional information to final rankings is the net flow value in Promethee II method. The relations between alternatives are very similar. However, the distances between them represented by the net flows vary e.g. the highest difference in the analyzed net flows is between alternatives A1 in the first and third ranking. The difference between net flow values of alternative A3 in these rankings is very low. Thus, the distance between A1 and A3 in the first ranking is small, while in the third one is very high.

Table 2. The model of DM's preferences for the distribution system

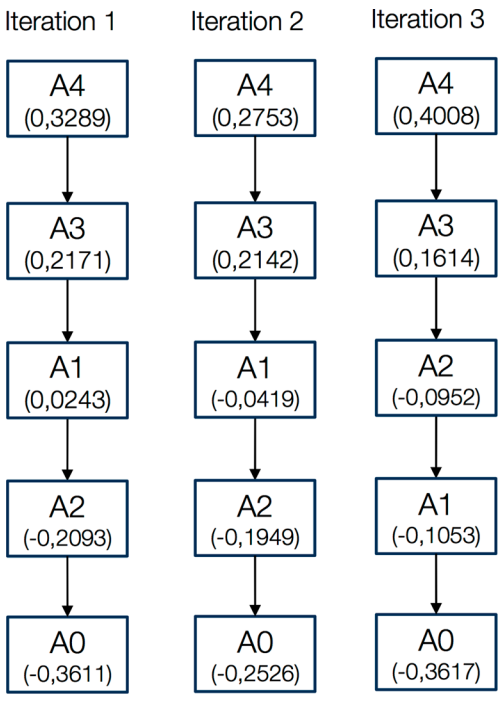
Criteria	Weights	Thresholds	
		Indifference q	Preference p
C1	10	0,60	1,10
C2	9	0,05	0,10
C3	5	0,10	0,20
C4	5	2,00	4,70
C5	6	6,00	14,00
C6	4	3,00	5,00
C7	7	0,50	1,00

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Deterministic values of criteria randomly generated with an application of Monte Carlo method – the distribution system

Criteria	Alternatives				
	A0	**A1**	**A2**	**A3**	**A4**
Iteration 1					
C1	4,06	3,02	2,05	1,07	1,13
C2	0,95	0,89	1,55	0,87	0,81
C3	0,52	0,43	0,44	0,82	0,80
C4	31,79	30,16	29,00	36,16	36,91
C5	50,43	48,96	30,39	14,87	11,80
C6	19,81	18,72	22,82	0,00	69,02
C7	0,0	-1,0	2,7	7,5	5,8
Iteration 2					
C1	4,07	2,77	2,13	1,05	1,07
C2	1,09	0,89	1,57	0,67	0,83
C3	0,51	0,45	0,44	0,82	0,78
C4	32,63	30,07	28,18	37,14	37,05
C5	50,96	51,45	31,32	13,49	11,31
C6	23,12	19,12	22,60	0,00	68,41
C7	0,0	-1,0	2,7	7,5	5,8

Figure 3. Three selected final rankings of alternatives resulting from calculations with an application of Promethee II method



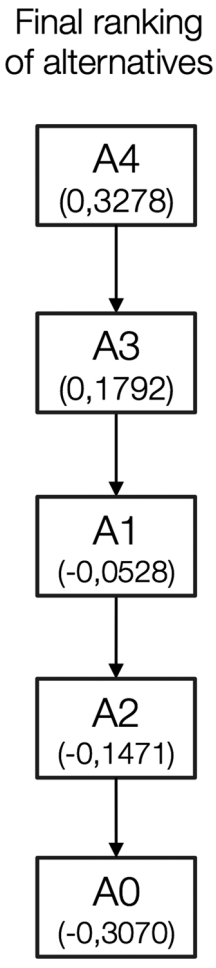
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In most of the rankings the best position is occupied by A4, the second location has A3 and the third one – A1. The next position has the alternative A1 and the last one – A0. In the last ranking the locations of alternatives A1 and A2 are switched.

Analyzing the total set of data composed of 150 rankings with the distance between alternatives (quantitative information), the author of this paper has decided to calculate the expected value of the relations between these variants. These computations constitute the third step of the presented stochastic MCDM method and the result is presented in Figure 4. Finally, in the fourth step the final ranking of alternatives is shown. See Figure 4.

The best scenario is the alternative A4. It represents the most radical changes in the analyzed distribution system. Next position in the ranking is occupied by the alternative A3. The changes modeled in this variant are significant, too. On the third position is the alternative A1, which is featured by slight changes in the existing distribution system of electrotechnical products. On the fourth location is the alternative A2 and the worst position is occupied by the current state of the system represented by the alternative A0.

Figure 4. Final ranking of alternatives with the expected values of net flows



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The computational experiments are finished by the sensitivity analysis. It is carried out for the minimum (Figure 5a), maximum (Figure 5b) and expected values (Figure 5c) of ranges of variations of each alternative evaluated by the set of criteria.

The results of the sensitivity analysis show that the rankings of alternatives are the same in terms of internal comparison and regarding the final ranking presented in Figure 4. The net flows are not the same. However, the difference between them is not significant. It is worth to mention that the expected value of the net flow (Figure 4) and the net value calculated for the expected values of the alternatives evaluated by the set of criteria (Figure 5c) are very close. The minimum difference between them equals 0,0056 - for the alternative A2 and the maximum difference is 0,0463 – for the alternative A0.

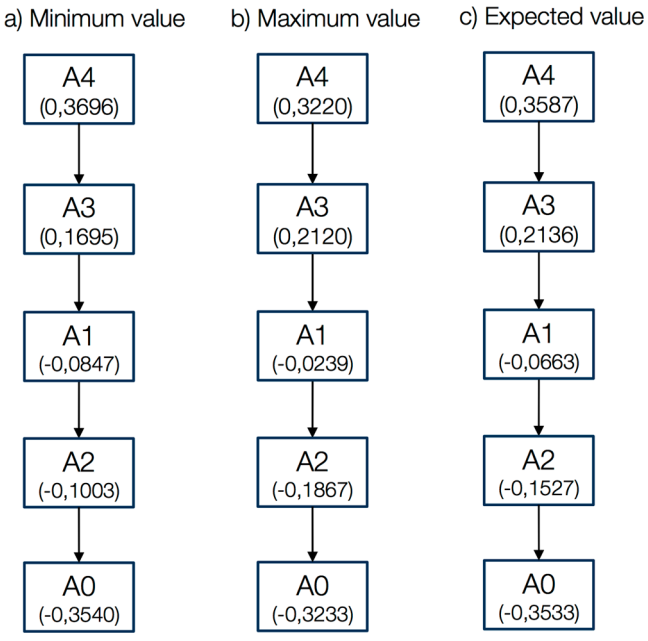
Based on the analysis of the final ranking of alternatives and the sensitivity analysis the author of this chapter recommends the most radical changes in the analyzed system represented by alternative A4. This step should radically improve the organization of the distribution system, while the current state is the worst scenario.

Ranking of Redesign Variants of Warehouse with Pharmaceutical Products Using Electre III Method and Bayes Classifier

Problem Definition

The second case presents company operating in Poland. Its core business is the production of pharmaceutical products. One of the important infrastructural objects of the company is a warehouse of finished goods. The composition of its rooms includes, among others: reception area, storage rooms, picking zone, edition zone and offices. In the area of admissions there is made an acceptance of the products

Figure 5. (a) Rankings of alternatives with the values of net flows for the minimum; (b) maximum; (c) expected values of ranges of variations



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

and their control (qualitative and quantitative). Storage facilities are divided into three zones, i.e. high storage (inventories are stored in 5-levels' stacks), low storage (replenished by goods from high storage zone, picking takes place there) and the cold room (temperature is maintained at 2-6° C). Based on the detailed analysis of the system its major drawbacks are recognized, such as: zone admission is often full, and vehicles with finished goods have to wait for unloading (sometimes several days). There are numerous overtimes associated with the picking process. Similarly, as in the previous case, the simulation model of the current state has been constructed and its redesign scenarios have been implemented in the simulation tool, as well. The set of evaluation criteria includes technical, organizational, economical and social aspects. The values are presented as probability distributions and they are as follows:

K1: Efficiency of warehouse employees (deviation from the optimum value) [-] – minimized criterion,
K2: Utilization of in-company transportation means [%] – maximized criterion,
K3: Efficiency of warehousing process [%] – maximized criterion,
K4: Difference between the levels of investments and divestments [PLN K] – minimized criterion.

The set of redesign scenarios is composed of 5 variants. They fulfill the following aims:

Variant V1: Increasing the employment in the warehouse leading to the improvement of employees efficiency ratio;
Variant V2: Reduction of the in-company transportation means, improving the efficiency ratio connected with technical resources;
Variant V3: Decreasing the queue of vehicles waiting for unloading, improving the efficiency ratio connected with human and technical resources;
Variant V4: Redesign of the warehouse area, improvement of the warehousing process;
Variant V5: Redesign of the warehouse area, reduction of the in-company transportation means, reallocation of tasks leading to the improvement of efficiency of human and technical resources and better organization of the warehouse.

The values of variants evaluated by different criteria are calculated on the basis of simulation experiments. They are presented in Table 4.

Table 4. Performance matrix of the warehouse with pharmaceutical products and its redesign scenarios

Criteria				Variants				
Name	**Direction ofpreferences**	**Value**	**V0**	**V1**	**V2**	**V3**	**V4**	**V5**
K1	Min	Expected value	15	1	4	5	8	5
		Range of variation	[14; 16]	[0; 2]	[0; 9]	[2; 13]	[6; 10]	[1; 11]
K2	Max	Expected value	43	39	84	82	78	88
		Range of variation	[25; 58]	[27; 58]	[73; 90]	[70; 87]	[68; 80]	[80; 92]
K3	Max	Expected value	89	91	90	97	100	98
		Range of variation	[80; 94]	[85; 96]	[84; 98]	[90; 99]	[99; 100]	[97; 100]
K4	Min	Expected value	0	0	-20	100	230	190
		Range of variation	-	-	[-35; -18]	[85; 120]	[200; 250]	[170; 210]

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The decision problem is formulated as a multicriteria stochastic ranking problem.

As in the previous case study, it is hard to decide which alternative is the compromise solution while the information presented in Table 4 is considered. Some variants have the best performance on particular criteria, such as V1 on criterion K1, and V2 on criterion K4, and V4 on criterion K3, and V5 on criterion K2, while on the other criteria their values are the worst, such as V1 on criterion K2 and V4 on criterion K4. Variant V3 hasn't got the best nor the worst value of its performance. However, on some criteria its performance is very good, e.g. criteria K2 and K3, while on the others is very poor, e.g. K4.

Thus, the stochastic approach combining Electre III method and Bayes classifier is applied to solve it.

Computational Experiments

The first step of the computational procedure is based on modeling of decision maker's preferences according to the principles of Electre III method. At the beginning, the DM has assigned different weights to criteria. He has used the scale of 1 to 10 points, where 1 represents the least important criterion and 10 – the most significant one. Next, the thresholds values i.e. q (indifference), p (preference) and v (veto) have been specified. Weights and thresholds are presented in Table 5.

Model of the DM's preferences indicates that the economical issues are the most important. They are represented by criterion K4. Very significant meaning has the general operation of the warehouse, i.e. criterion K3. Criteria K1 and K2 representing social and technical aspects are important, too. However, the weights assigned to them are the lowest in the total set of evaluations.

The DM is sensitive to changes. Some aspects of redesign scenarios make them incomparable with the current state, which is emphasized by the veto threshold, e.g. V2 vs. V0 regarding the criterion K1. Carrying out pairwise comparison of variants, some of them are weakly preferred, e.g. V3 vs. V1 regarding the criterion K3 or strongly preferred to each other, e.g. V3 vs. V5 regarding the criterion K4.

Next, computations within this step are based on the random numbers of each criterion value generation. For this purpose, the Monte Carlo method has been applied. The number of iterations equals 150 providing sufficient sample of deterministic parameters for the next step of calculations. The exemplary sets of 3 iterations are presented in Table 6.

The values presented in Table 6 are randomly dispersed in the ranges of variations. Some of them are very close to the minimum or maximum values of these ranges.

In the second step, based on the results of all iterations the deterministic values are the input data for the computations with an application of MADM ranking method Electre III. Similarly, as in Promethee II method, the total number of calculations equals 150.The exemplary 2 results are presented in Figure 6.

Table 5. The model of DM's preferences for the warehouse of pharmaceuticals

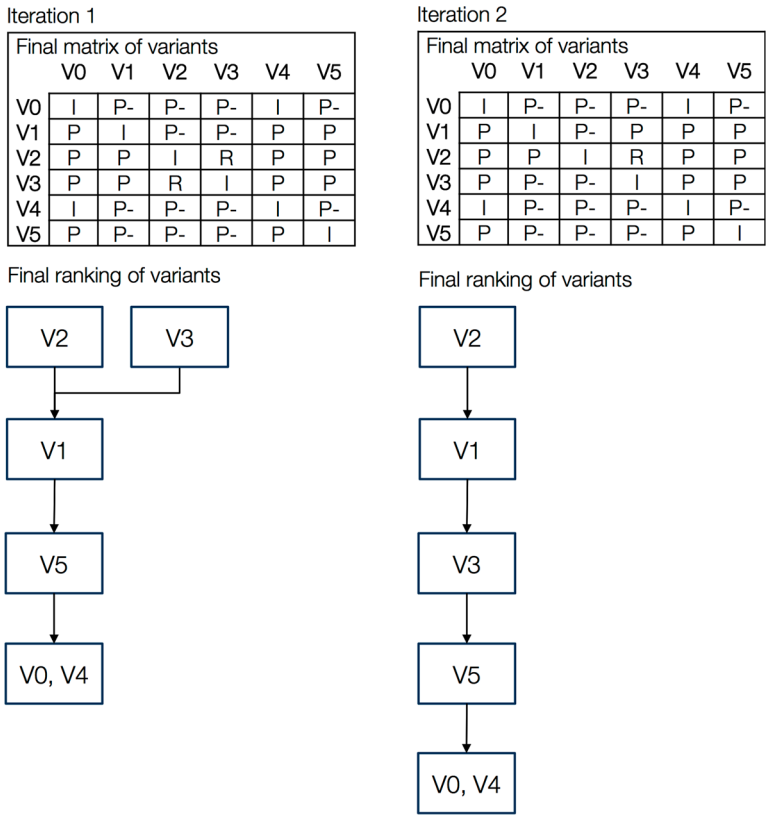
Criteria	Weights	Thresholds		
		Indifference q	Preference p	Veto v
K1	6	1	5	10
K2	5	10	20	40
K3	8	3	5	10
K4	9	20	50	100

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. Deterministic values of criteria randomly generated with an application of Monte Carlo method – an example with the warehouse of pharmaceuticals

Criteria		Variants				
	V0	V1	V2	V3	V4	V5
Iteration 1						
K1	14,95	1,62	8,29	2,19	7,12	4,89
K2	47,75	39,92	84,16	83,41	72,08	81,85
K3	84,87	94,92	92,87	96,47	99,64	97,79
K4	0,00	0,00	-21,04	94,06	230,49	186,27
Iteration 2						
K1	15,95	1,42	4,63	6,92	8,13	6,80
K2	53,74	38,53	86,37	80,40	76,71	90,58
K3	86,59	91,57	93,61	94,11	99,38	99,41
K4	0,00	0,00	-24,50	93,70	237,61	183,71

Figure 6. Two selected final matrices of variants and their graphical representation resulting from calculations with an application of Electre III method



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

These two matrices present the following relations: indifference I, preference P, reciprocal of preference P- and incomparability R. In both rankings, the indifference I relation is between V0 and V4 variants, the preference P relation is between V2 and V1, the reciprocal of preference P- relation is between variants V5 and V1, while the incomparability R relation can be observed between variants V2 and V3 in the first iteration.

The detailed analysis of the total set of data composed of 150 matrices, rankings and pairwise comparisons between variants' relations (qualitative information), led the author of this chapter to the conclusion that there are many differences between them. Based on that, it is hard to present the DM final recommendation regarding the changes in the analyzed warehouse. Thus, the classification of these relations and their allocation to predefined classes has been made. This allocation corresponds to the definition of decision attributes. The classification is also the third step of the presented stochastic MCDM method. Bayes classifier has been applied to solve this problem.

Initially, 150 matrices have been split into 15 sets with 10 matrices in each of them. For each set the relations between variants represented by the description attributes (i.e. indifference I, preference P, reciprocal of preference P- and incomparability R relations) have been recognized and the probabilities of occurrence of particular relationships in the whole set have been calculated. See Table 7.

The values of description attributes are in the range [0,1]. Based on them the decision attributes with the following rules have been constructed:

- If the probability of the description attribute for a relation between variants has the value greater than 0,5 then the decision attribute is the same as the description attribute
- If the probabilities of the description attributes P- and R for a relation between variants have the same values lower than 0,5 then the decision attribute is the incomparability relation (R)
- If the probabilities of description attributes I and P for a relation between variants have values 0,5 then the decision attribute is a weak preference relation (Q)
- If the probabilities of description attributes I and P- for a relation between variants have values 0,5 then the decision attribute is a reciprocal of weak preference relation (Q-)

Table 7. The results of computational experiments of the share of relations I, P, P-, and R between pairs of variants for a part of testing set and their assignment to decision attributes

Relation between variants	Description attribute				Decision attribute (Class)
	I	P	P-	R	
V0-V0	1	0	0	0	I
V0-V1	0	0	1	0	P-
V0-V2	0	0	1	0	P-
V0-V3	0	0	1	0	P-
V0-V4	0,8	0,1	0,1	0	I
…					
V1-V3	0	0,2	0,4	0,4	R
…					
V2-V3	0	0,7	0	0,3	P
…					

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

As a result, 6 decision classes (attributes) C(z) has been distinguished, such as:

C(I) = I: Indifference relation between variants
C(P) = P: Strong preference relation between variants
C(P-) = P-: Reciprocal of strong preference relation between variants
C(R) = R: Incomparability relation between variants
C(Q) = Q: Weak preference relation between variants
C(Q-) = Q-: Reciprocal of weak preference relation between variants

Next, the training set T composed of 10 sets of matrices has been randomly selected and its testing has been carried out. Based on the information collected in this set, the a priori probabilities P(C(z)) of an event that the classes C(z) occur as well as the conditional a priori probabilities P(x(a)|C(z)) have been calculated.

The remaining 5 sets of matrices (out of 15) have been utilized in the computational experiments. The conditional a posteriori probability has been calculated for each pair of variants. Based on it and based on the a priori probability of an event that the class C(z) occur, the conditional a posteriori probability P(C(z)|x(a)) that a certain relation between variants occurs has been calculated.

For each pair of variants from the testing set the classification error β has been calculated. It equals 0,06, which means that the training model has been constructed correctly. The credibility of selected classification method in solving the presented problem is high.

The vector of observations with a priori unknown decision attributes has been selected. This vector has been composed of 30 relations between pairs of variants and has been characterized by the description attributes. They are the basis to assign 30 decision attributes concerning the relations between variants. 14 pairs of variants have been classified to class P, 14 pairs to class P- and remaining 2 pairs – to class I. No pair of variants have been classified to the decision attributes Q, Q- and R.

In the fourth step the final matrix and ranking of variants have been generated. The results are presented in Figure 7.

The final relations between variants are presented in the matrix (Figure 7a) as bold text. The information of probabilities of occurrence of particular relations between variants is also demonstrated. For example, the final relation between variants V0 and V4 is I (indifference) and the probability of its occurrence equals 0,720. This stochastic ranking matrix is transformed into the final ranking of variants and presented in Figure 7b. The best solution is at the top of the ranking, while the worst one - at the bottom. Dotted rectangular box is drawn for two variants for which the probability of occurrence of preference P and indifference I relations are close to 0,5. The analyst recognizes this pair of variants as almost indifferent.

Next, the ranking has been analyzed. The leader is variant V2. It is featured by slight changes i.e. reduction of in-company transportation means. The second position has the variant V3. It represents more radical changes in the organization of warehouse, leading to a higher level of internal fleet utilization, improvement of human resources efficiency ratio and radical reduction of vehicles waiting for unloading.

The strength of preference relation between V2 and V3 is high (0,993), which means that the distance between them is long with a high degree of certainty. Fourth and fifth location in the final rank have variants V1 and V5, representing evolutionary and revolutionary changes, respectively. The probability of preference and indifference relation between them is very close, i.e. 0,527 (P) and 0,466 (I). Thus, they are marked by the dotted line. The worst position in the ranking have variants V0 and V4. The relation between them is indifference I (0,720) and they are presented in one box.

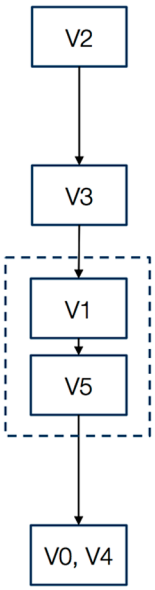
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. (a) The stochastic ranking matrix representing the mutual relations between variants and (b) the final stochastic ranking of variants showing the positions of variants in the graphical form

a) Stochastic ranking matrix

	V0	V1	V2	V3	V4	V5
V0	I (1,000)	P- (0,993) I (0,007)	P- (1,000)	P- (1,000)	I (0,720) P (0,267) P- (0,013)	P- (0,993) I (0,007)
V1	P (0,993) I (0,007)	I (1,000)	P- (1,000)	P- (0,720) R (0,267) P(0,013)	P (1,000)	P (0,527) I (0,466) P- (0,007)
V2	P (1,000)	P (1,000)	I (1,000)	P (0,993) R (0,007)	P (1,000)	P (1,000)
V3	P (1,000)	P (0,720) R (0,267) P- (0,013)	P- (0,993) R (0,007)	I (1,000)	P (1,000)	P (0,993) P- (0,007)
V4	I (0,720) P- (0,267) P (0,013)	P- (1,000)	P- (1,000)	P- (1,000)	I (1,000)	P- (1,000)
V5	P (0,993) I (0,007)	P- (0,527) I (0,466) P (0,007)	P- (1,000)	P- (0,993) I (0,007)	P (1,000)	I (1,000)

b) Final ranking of variants



Finally, the sensitivity analysis has been carried out. Similarly as in the previous case, the final rankings for minimum, maximum and expected values of criteria evaluating variants have been calculated. They are presented in Figure 8.

The results of the sensitivity analysis show that the rankings of variants are similar. All of them present variant V2 as the best scenario. Variants V0 and V4 are the worst redesign scenarios. However, their position is ambiguous – the following relations can be observed: I, P, P-. In two rankings the relation between variants V1 and V3 is R, while in one ranking it is P-. They are presented on the second or third location. In two rankings third position is occupied by the variant V5 and in one ranking this variant is on the second location.

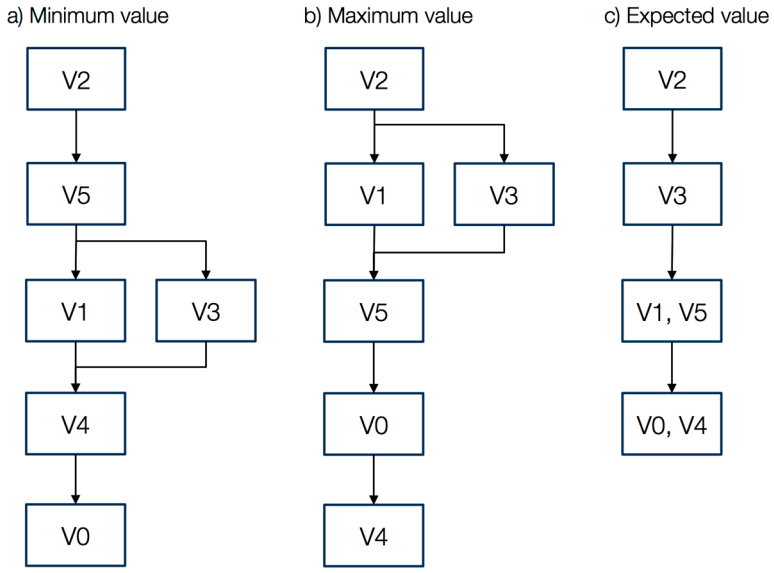
Based on the analysis of the final stochastic ranking of variants and the sensitivity analysis the author of this chapter recommends evolutionary changes in the analyzed warehouse. They should be a good start for more radical transformation of the system.

FUTURE RESEARCH DIRECTIONS

The consciousness of problems related to complex systems makes that the artificial intelligence becomes very eligible field of research. Intelligence combined with the application of selected MADM methods, probability theory and machine learning methods allows generating credible solutions, especially when dealing with uncertain and dynamically changing different systems' environments. They can be represented by the stochastic information bringing closer to the reality.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. (a) Ranking of alternatives for the minimum; (b) maximum; (c) expected values of ranges of variations calculated with an application of Electre III method



Thanks to the artificial intelligence it is possible to aid the decision process, i.e. to analyze and model unstructured decision situations, to take into account different points of view represented by the set of evaluation criteria, to collect preferential information, to generate reliable results and recommend satisfactory solutions. Based on these operations the general aim of systems operating on different markets, which is the improvement leading to the competitive advantage and constant development, can be achieved.

Considering systems presented in this chapter, it is worth mentioning about their complexity. This aspect has an influence on each step of their (and other systems) improvement. Systems' analysis and data collection has been the most time consuming process. At this stage, the artificial intelligence has been very helpful in finding, ordering the data and transforming them to the required form, as well. MADM methods have provided the bridge between the analytical part of the research and real-world information. They also have played an important role in coping with structuring process of the DM's preferences. The huge number of information the DM had to deal with during the analysis, raised doubts and sometimes disbelief of this process success. However, the step-wise procedure has helped to structure and clarify the situation. The probability analysis and Bayes classification method confirmed in the belief that the recommended solution is the best one taking into account the set of criteria and preferential information.

Undoubtedly, artificial intelligence plays an important role during the ordering of information, computational experiments, finding the best solutions. It also helps to cope with social problems, decision maker's doubts and uncertain information.

The next step in solving complex, stochastic decision problems would be concentrated on shortening the time of finding the best solution. Generally, the computational experiments with the application of stochastic MADM methods are time consuming and the algorithm of stochastic MADM method presented in this chapter, as well. Both cases proved that the computational procedure takes long hours and it is very laborious. One of the solutions is the implementation of this idea in computer – based decision support system developed to solve complex decision problems.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

In this chapter, some deterministic ranking methods like Electre III (Roy, 1985; Vincke, 1990), Promethee II (Brans et al., 1986) and stochastic ranking methods like SMAA-III (Tervonen & Figueira, 2008), stochastic Promethee (Marinoni, 2005) have been presented. Based on them it is possible to solve decision problems with different types of data, i.e. stochastic values of criteria evaluating the set of alternatives, uncertain DM's preferences, his/her hesitations. However, some decision situations are very complex and the application of existing methods would not result in a desirable solution.

Thus, the concept of stochastic multiple criteria decision making method to solve such decision problems has been presented. The proposed approach can be implemented to solve complex decision problems aiming at ranking of variants. It is based on a composition of classical methodologies, i.e. MADM, probability theory and classification method. Depending on the considered situation, this stochastic method gives the opportunity to solve problems with deterministic, stochastic and mixed information about the alternatives evaluated by the set of criteria, and uncertain preferences of the decision maker.

After successful implementations of the proposed approach, results show the most desirable direction of changes leading to high improvements of both analyzed systems.

It is proved that the combination of classical methods and tools would result in a reasonable solution. Undoubtedly, it is time consuming and the final application would make it user friendly.

Artificial intelligence is still evolving. The recent subjects of research interests within its one area i.e. multiple criteria decision aiding, constitute future directions. They are, among others, as follows:

- The selection of the most suitable MADM method, which has been the subject of research of Guitouni and Martel (1998), Sawicka (2012)
- Robustness analysis, widely described by Dias and Climaco (1999) and recently presented by Zavadskas et al. (2013);
- Multiple decision makers, including group decision making process, public participation etc., with general issues presented by Keeney (1976), Belton and Stewart (2002) and recently presented by Yue (2011)

REFERENCES

Anand, G., & Kodali, R. (2008). Selection of Lean Manufacturing Systems Using the PROMETHEE. *Journal of Modeling in Management*, *3*(1), 40–70. doi:10.1108/17465660810860372

Badri, M. (1999). Combining the AHP and GP Model for Global Facility Location-Allocation Problem. *International Journal of Production Economics*, *62*(3), 237–248. doi:10.1016/S0925-5273(98)00249-7

Behzadian, M., Kazemzadeh, R., Albadvi, A., & Aghdasi, M. (2010). PROMETHEE: A Comprehensive Literature Review on Methodologies and Applications. *European Journal of Operational Research*, *200*(1), 198–215. doi:10.1016/j.ejor.2009.01.021

Belton, V., & Stewart, T. J. (2002). *Multiple Criteria Decision Analysis. An Integrated Approach*. Dordrecht: Kluwer Academic Publishers. doi:10.1007/978-1-4615-1495-4

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Berger, J. (1985). *Statistical Decision Theory and Bayesian Analysis*. New York, NY: Springer-Verlag. doi:10.1007/978-1-4757-4286-2

Boden, M. A. (1977). *Artificial Intelligence and Natural Man*. New York, NY: Basic Books.

Brans, J. P., Vincke, P., & Mareschal, B. (1986). How to Select and How to Rank Projects: The PROMETHEE Method. *European Journal of Operational Research*, *24*(2), 228–238. doi:10.1016/0377-2217(86)90044-5

Bremner, D., Demaine, E., Erickson, J., Iacono, J., Langerman, S., Morin, P., & Toussaint, G. (2005). Output-Sensitive Algorithms for Computing Nearest-Neighbor Decision Boundaries. *Discrete & Computational Geometry*, *33*(4), 593–604. doi:10.100700454-004-1152-0

Buchanan, J., & Sheppard, P. (1998). Ranking Projects Using the Electre Method. In *Proceedings of the 33rd Annual Conference of Operational Research Society of New Zealand*.

Cetiner, B., Sari, M., & Borat, O. (2010). Neural Network Based Traffic-Flow Prediction Model. *Mathematical and Computational Applications*, *15*(2), 269–278. doi:10.3390/mca15020269

Czyzak, P., & Jaszkiewicz, A. (1998). Pareto Simulated Annealing – a Metaheuristic Technique for Multiple-Objective Combinatorial Optimization. *Journal of Multi-Criteria Decision Analysis*, *7*(1), 34–47. doi:10.1002/(SICI)1099-1360(199801)7:1<34::AID-MCDA161>3.0.CO;2-6

de Boer, L., van der Wegen, L., & Telgen, J. (1998). Outranking Methods in Support of Supplier Selection. *European Journal of Purchasing and Supply Management*, *4*(2/3), 109–118. doi:10.1016/S0969-7012(97)00034-8

Delhaye, C., Teghem, J., & Kunsch, P. (1991). Application of the ORESTE Method to a Nuclear Waste Management Problem. *International Journal of Production Economics*, *24*(1-2), 29–39. doi:10.1016/0925-5273(91)90150-R

Demsar, J. (2006). Statistical Comparisons of Classifiers over Multiple Data Sets. *Journal of Machine Learning Research*, *7*(Jan), 1–30.

Dias, L., & Clímaco, J. (1999). On Computing ELECTREs Credibility Indices under Partial Information. *Journal of Multi-Criteria Decision Analysis*, *8*(2), 74–92. doi:10.1002/(SICI)1099-1360(199903)8:2<74::AID-MCDA234>3.0.CO;2-7

Goldberg, D. (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Boston, MA: Addison-Wesley.

Guitouni, A., & Martel, J.-M. (1998). Tentative Guidelines to Help Choosing an Appropriate MCDA Method. *European Journal of Operational Research*, *109*(2), 501–521. doi:10.1016/S0377-2217(98)00073-3

Hafeez, K., Zhang, Y. B., & Malak, N. (2002). Determining Key Capabilities of a Firm Using Analytic Hierarchy Process. *International Journal of Production Economics*, *76*(1), 39–51. doi:10.1016/S0925-5273(01)00141-4

Hall, P., Park, B., & Samworth, R. (2008). Choice of Neighbor Order in Nearest-Neighbor Classification. *Annals of Statistics*, *36*(5), 2135–2152. doi:10.1214/07-AOS537

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Halouani, N., Chabchoub, H., & Martel, J.-M. (2009). PROMETHEE-MD-2T Method for Project Selection. *European Journal of Operational Research*, *195*(3), 841–895. doi:10.1016/j.ejor.2007.11.016

Jacquet-Lagrèze, E., & Siskos, Y. (1982). Assessing a Set of Additive Utility Functions for Multicriteria Decision Making: The UTA Method. *European Journal of Operational Research*, *10*(2), 151–164. doi:10.1016/0377-2217(82)90155-2

Kahraman, C. (Ed.). (2008). *Fuzzy Multi-Criteria Decision-Making. Theory and Applications with Recent Developments*. New York, NY: Springer. doi:10.1007/978-0-387-76813-7

Keeney, R. L. (1976). A Group Preference Axiomatization with Cardinal Utility. *Management Science*, *23*(2), 140–145. doi:10.1287/mnsc.23.2.140

Lahdelma, R., Hokkanen, J., & Salminen, P. (1998). SMAA - Stochastic Multiobjective Acceptability Analysis. *European Journal of Operational Research*, *106*(1), 137–143. doi:10.1016/S0377-2217(97)00163-X

Law, A. M., & Kelton, W. D. (2000). *Simulation, Modeling and Analysis*. New York, NY: McGraw Hill.

Malczewski, J., Moreno-Sanchez, R., Bojorquez-Tapia, L. A., & Ongay-Delhumeau, E. (1997). Multicriteria Group Decision-Making for Environmental Conflict Analysis in the Cape Region, Mexico. *Journal of Environmental Planning and Management*, *40*(3), 349–374. doi:10.1080/09640569712137

Manshadi, E. D., Mehregan, M. R., & Safari, H. (2015). Supplier Classification Using UTADIS Method Based on Performance Criteria. *International Journal of Academic Research in Business and Social Sciences*, *5*(2), 31–45. doi:10.6007/IJARBSS/v5-i2/1457

Marinoni, O. (2005). A Stochastic Spatial Decision Support System Based on PROMETHEE. *International Journal of Geographical Information Science*, *19*(1), 51–68. doi:10.1080/13658810412331280176

Mitchell, T. (1997). *Machine Learning*. New York, NY: McGraw-Hill.

Mousseau, V., Slowinski, R., & Zielniewicz, P. (2000). A User-Oriented Implementation of the ELECTRE-TRI Method Integrating Preference Elicitation Support. *Computers & Operations Research*, *27*(7-8), 757–777. doi:10.1016/S0305-0548(99)00117-3

Negnevitsky, M. (2005). *Artificial Intelligence. A Guide to Intelligent Systems*. Harlow: Addison-Wesley.

Ngai, E. W. T. (2003). Selection of Web Sites for Online Advertising Using AHP. *Information & Management*, *40*(4), 233–242. doi:10.1016/S0378-7206(02)00004-6

Pawlak, Z. (1982). Rough Sets. *International Journal of Computer and Information Science*, *11*(5), 341–356. doi:10.1007/BF01001956

Pisetta, V., Jouve, P.-E., & Zighed, D. (2010). Learning with Ensembles of Randomized Trees. In J. Balcazar, F. Bonchi, A. Gionis, & M. Sebag (Eds.), *Machine Learning and Knowledge Discovery in Databases* (pp. 67–82). Berlin: Springer-Verlag. doi:10.1007/978-3-642-15939-8_5

Quinlan, J. (1986). Induction of Decision Trees. *Machine Learning*, *1*(1), 81–106. doi:10.1007/BF00116251

Roy, B. (1985). *Methodologie Multicritere d'Aide a la Decision*. Paris: Economica.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Russel, S. J., & Norvig, P. (1995). *Artificial Intelligence. A Modern Approach*. Englewood Cliffs, NJ: Prentice Hall.

Saaty, T. L. (1980). *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. New York, NY: Mc-Graw Hill.

Salminen, P., Hokkanen, J., & Lahdelma, R. (1998). Comparing Multicriteria Methods in the Context of Environmental Problems. *European Journal of Operational Research*, *104*(3), 485–519. doi:10.1016/S0377-2217(96)00370-0

Sawicka, H. (2012). *The Method of Redesign the Distribution System* (Unpublished doctoral dissertation). Warsaw University of Technology, Warsaw, Poland.

Sawicka, H., Weglinski, S., & Witort, P. (2010). Application of Multiple Criteria Decision Aid Methods in Logistic Systems. *LogForum*, *6*(10), 99–110.

Schalkoff, R. I. (1990). *Artificial Intelligence: An Engineering Approach*. New York, NY: McGraw-Hill.

Siskos, J. (1982). A Way to Deal with Fuzzy Preferences in Multicriteria Decision Problems. *European Journal of Operational Research*, *10*(3), 314–324. doi:10.1016/0377-2217(82)90230-2

Smola, A., & Vishwanatan, S. (2008). *Introduction to Machine Learning*. Cambridge, UK: Cambridge University Press.

Stam, A., & Duarte Silva, A. P. (1997). Stochastic Judgments in the AHP: The Measurement of Rank Reversal Probabilities. *Decision Sciences*, *28*(3), 655–688. doi:10.1111/j.1540-5915.1997.tb01326.x

Tervonen, T., & Figueira, J. (2008). A Survey on Stochastic Multicriteria Acceptability Analysis Methods. *Journal of Multi-Criteria Decision Analysis*, *15*(1-2), 1–14. doi:10.1002/mcda.407

Vincke, P. (1992). *Multicriteria Decision-Aid*. Chichester, UK: John Wiley & Sons.

Wei, C., & Schonfeld, P. (1993). An Artificial Neural Network Approach for Evaluating Transportation Network Improvements. *Journal of Advanced Transportation*, *27*(2), 129–151. doi:10.1002/atr.5670270202

Yao, Y. (2011). Superiority of Three-Way Decisions in Probabilistic Rough Set Models. *Information Sciences*, *181*(6), 1080–1096. doi:10.1016/j.ins.2010.11.019

Yue, Z. (2011). A Method for Group Decision-Making Based on Determining Weights of Decision Makers using TOPSIS. *Applied Mathematical Modelling*, *35*(4), 1926–1936. doi:10.1016/j.apm.2010.11.001

Zavadskas, E. K., Antucheviciene, J., Saparauskas, J., & Turskis, Z. (2013). MCDM methods WASPAS and MULTIMOORA: Verification of Robustness of Methods when Assessing Alternative Solutions. *Economic Computation and Economic Cybernetics Studies and Research*, *47*(2), 5–20.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

KEY TERMS AND DEFINITIONS

Classification Theory: The research field of artificial intelligence concentrated on categorization of objects to predefined classes. It is assumed that these objects are characterized by quantitative information.

Decision Maker: A person, who expresses his/her preferences, evaluates the decision situation and the results of the computational experiments. This person makes the final decision regarding the considered problem, e.g. selects the best redesign scenario.

Decision Problem: A situation where the decision maker has to face the problem of selecting one of at least two alternatives. The decision problem is composed of the description of the decision situation, definition of constraints and evaluation criteria, identification of the decision maker and his/her role in the decision process and definition of stakeholders, as well.

MADM: An abbreviation of Multiple Attribute Decision Making. It is one of the Multiple Criteria Decision Making approaches referring to the situation of finite set of alternatives evaluated by the set of attributes leading to the selection of the best solution.

MCDM: An abbreviation of Multiple Criteria Decision Making. It is composed of two approaches, i.e. MADM (Multiple Attribute Decision Making) and MODM (Multiple Objective Decision Making). In MADM problems the number of alternatives is finite and the evaluation attributes are usually conflicting, while in MODM problems the set of alternatives is infinite and the evaluation criteria are described by continuous function.

Ranking Methods: A group of MCDM methods concentrated on finding the final hierarchy of alternatives, i.e. the rank. Generally, ranking methods are divided into two sets referring to different methodological backgrounds. The first one originates from European school and it is based on outranking relation, e.g. Electre III, Promethee II methods. The second one, belongs to American school and it is based on multiattribute utility theory, e.g. AHP, UTA methods.

Stochastic Information: Information represented by a random number, usually by a random probability distribution. It is connected with uncertainty about the values of parameters, expected input or output information. In the real-world systems the stochastic information is the result of unexpected disturbances, demand fluctuations etc.

This research was previously published in the Handbook of Research on Advanced Data Mining Techniques and Applications for Business Intelligence edited by Shubhamoy Dey, Shrawan Kumar Trivedi, Anil Kumar, and Tapan Kumar Panda; pages 315-340; copyright year 2017 by Business Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 47
Trends of Educational Informatics in Latin America

Adán A. Gómez
Universidad de Córdoba, Colombia

Angela M. Solano
Universidad de Córdoba, Colombia

Manuel F. Caro
Universidad de Córdoba, Colombia

Yina M. Vega
Universidad de Córdoba, Colombia

ABSTRACT

Educational Informatics is a multidisciplinary research area that uses Information and Communication Technology (ICT) in education. This article reports a wide overview about the new perspective of educational informatics that has been implemented in Latin America. This general view is the result of the different talks presented during the XIII Conference of Educational Informatics, organized by the Ibero-American Network of Educational Informatics RIBIE (Red Iberoamericana de Informática Educativa) in Colombia. These talks were proposed by international panelists from various parts of Latin America in the lines of ICT for Education and Peace, Thinking and Artificial Intelligence, Emergent Scenarios in Education and Mobiles Technologies and Apps for Entrepreneurship and Learning.

INTRODUCTION

Informatics is the science of information that studies the nature of information, its processing, and ways of transformation between information, matter and energy (Wang et al., 2011). Educational Informatics is a multidisciplinary research area that uses Information and Communication Technologies (ICT) in education (Hervatis, Loe, Barman, O'Donoghue, & Zary, 2015). Educational informatics is the integration of processing, conservation and transmission of information through automatic machines in the teaching learning process (Chiappe & Sánchez, 2014; García-Peñalvo & Conde González, 2013; Levy et al., 2003; Sierra-Rodríguez & García-Peñalvo, 2015; Srivastava, 2012). Also, educational informatics involves the development, use, and evaluation of digital systems that use pedagogical knowledge in order to support learning (Caro, Toscano, Hernández, & David, 2009; Levy et al., 2003; Sierra-Rodríguez & García-Peñalvo, 2015). According to (Chiappe & Sánchez, 2014; Srivastava, 2012) educational infor-

DOI: 10.4018/978-1-7998-0951-7.ch047

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

matics has two main areas of concern: i) to understand the effects on people of using digital information resources, services, systems, environments and communications media for learning and education; ii) to contribute to the development of practical knowledge of relevance to diverse forms of learning/education using ICT and digital resources.

The Ibero-American Network of Educational Informatics RIBIE (*Red Iberoamericana de Informática Educativa* — www.ribie.net) was created by the VII Subprogram of Electronics and Applied Informatics of CYTED (*Programa Iberoamericano de Cooperación en Ciencia y Tecnológica para el Desarrollo*). RIBIE is formed by institutions and groups that develop and apply ICT for the solution of educational problems through research and technological innovation. RIBIE also promotes communication and collaboration among its members, around the project management, the formulation of policies and the development of strategies for education improvement.

RIBIE has been an important stimulus for the development of educational informatics all over the Ibero-American region. RIBIE has organized scientific activities, courses, workshops, researches and development projects, involving groups from the Ibero-American countries. The International Conference of Educational Informatics is an example of it. This Conference is carried out every two years in a member country of RIBIE.

Colombia organized the XIII Conference of Educational Informatics — International Version in 2017. During this event, the most recent trends of educational informatics in Latin America were presented around the following topics: i) ICT for Education and Peace; ii) Thinking and Artificial Intelligence; iii) Emergent Scenarios in Education; and iv) Mobile Technologies and Apps for Entrepreneurship and Learning.

In the research area "ICT for Education and Peace" the talks were related to the following topics: ICT as a support of peace construction, online government, democracy, participation, digital literacy, empowerment, and teaching- learning processes inside and outside the classroom.

In the topic "Thinking and Artificial Intelligence" different projects and experiences in terms of psychology, philosophy, cognition, neuroscience, neural networks, evolutionary computation, artificial intelligence, linguistics and robotic were presented.

The topic named "Mobile Technologies and Apps for entrepreneurship and Learning" involved topics related to new proposals in order to promote self-learning, entrepreneurship and innovation using mobile technologies.

By last, in "Emerging Scenarios in Education" projects and experiences connected with the most recent technologies to enrich the learning processes in multiple scenarios at social, business, cultural and educational levels were tackled.

The main objective of this paper is to present an overview about innovative trends in educational informatics, taking into account the 73 talks performed in the event mentioned above. This paper is a summary about the talks developed showing the research areas proposed through the following chapters; Chapter II shows briefly the researches and experiences in the topic called ICT for Education and Peace. Chapter III describes the studies framed in the Thinking and Artificial Intelligence area. Chapter IV tackles the different emerging scenarios that are being used in Latin America education. Chapter V takes into account the wire range of mobile technologies and Apps for entrepreneurship and learning. Finally, Chapter VI presents a set of conclusions and a proposal about future works in this area of knowledge.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ICT FOR EDUCATION AND PEACE

Decades ago certain Latin American countries had been working on the solution of their armed conflict. One of these countries is Colombia, in which the peace process is being carried out currently. The education and the implementation of ICT is becoming the starting point for this conflict solution of the nation. As an example of this, the use of ICT is taken as a tool for re-socialization in areas of conflict (Zarate & Medina, 2017). Workshops on digital literacy, human rights and ethical values are being promoted In order to influence the life project of people who have been victims of the violence and conflict in Colombia; (Prieto & Nájar, 2017).

This research area in Colombia also proposes to redesign the existing curricula, pursuing the participation of the students in topics related to citizenship (Rangel & Silva, 2017), culture of peace and Chair for Peace (Yagama & Sierra, 2017). This "chair for peace" is based on a law that guides implementation in schools around the country, where teachers design plans and pedagogical strategies, which promote students' participation in order to develop social skills and ethical values (Ley1732, 2014).

In this context, the role of the teacher is to lead processes that prevent, monitor and control problems through classroom projects, focusing on the development of values (Sandoval, 2017).

THINKING AND ARTIFICIAL INTELLIGENCE

The main goal of this research was to present a general scenario about the techniques implemented on the thinking modeling and the use of cognitive informatics and cognitive computing (CI & CC) in Education.

In this event, researches were conducted in different areas of knowledge, which aim to respond, in one or another way the following question: How does the natural intelligence process information? This question takes us along the path of cognitive sciences (Wang, 2002), which are dedicated to discovering the representational and computational capacities of the human mind (Miller, 2003). Any emerging disciplinary such as CI & CC shares this common root problem question (Wang, 2002). Studies in these disciplines allow representing both cognitive and metacognitive processes, which help us to develop educational resources in terms of hardware and software in order to enrich the teaching - learning process.

Other researches show a high tendency in the use and personalization of learning environments mediated by technologies to promote the development of thinking skills (Bejarano, 2017; Díaz, 2017; Martínez, Acevedo, & Correa, 2017). Such as: virtual learning and multimedia learning computer environments, which are pedagogically and technologically designed (Palomino, Barrera, & Culchac, 2017).

Furthermore, a methodology for the creation of robotic educational toys was presented (Suarez & Garces, 2017). As well as the use of cognitive robots as teaching - learning tools in the classroom. These new learning scenarios allow students to develop a systemic, structured, logical and formal thinking through problem-solving processes.

The experiences and researches presented in this event will mark a future trend in Latin America to the design of environments mediated by technology. It is observed that artificial intelligence facilitates the role of the teacher inside the classroom as the administrator of knowledge and students to be the key of the teaching learning process

Finally, It is evident through all the exposed in the thinking and Artificial Intelligence topic that in Latin America exists initiatives in the developing of intelligent systems based on cognitive and metacognitive processes simulations that facilitate building of new learning scenarios.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

EMERGING SCENARIOS IN EDUCATION

Trying to integrate ICT in the educational context is a complex task (Jimoyiannis, 2009; Tondeur, Coenders, van Braak, ten Brummelhuis, & Vanderlinde, 2009). This process implies a substantial change in the way the teaching and learning process is developed (Bingimlas, 2009). However, thanks to the potential that the use of ICT has within the classroom, teachers can implement new methodologies of teaching to respond to these changes (García-Valcárcel, Basilotta, & López García, 2014; John & Sutherland, 2004).

These new educational scenarios were evidenced in the XIII International Conference of Educational Informatics, where different researches presented the impact of new methodologies mediated by ICT upon education. In this research area, the talks were focused on two main topics. On the one hand, some relevant experiences in relation to the implementation of educational video games in different areas of knowledge, were socialized (Hernández, Hernández, & Sanabria, 2017; Patiño, 2017). These types of scenarios have allowed the formulation of classroom projects that pursuit not only to reinforce ICT skills but also to generate new learning spaces through Gamification (Ramírez & Solano, 2017; Sandoval, 2017).

On the other hand, the use and appropriation of ICT in the school facilitates the interdisciplinary as a determining characteristic in teaching learning process (Bulla & Solano, 2017). According to Kandiko (2012) and Klein (2010) interdisciplinary allows to identify skills and necessary knowledge to conduct researches and projects of innovation (Kandiko, 2012; Klein, 2010). This could be observed in experiences that try to change the ideas that young learners have about the use of ICT because they commonly associate these tools with entertainment and leisure activities (Patiño, 2017). That is to say, interdisciplinary facilitates the generation of new knowledge (Van Rijnsoever & Hessels, 2011).

Emerging scenarios in education mark the path of new experiences that transform the way of learning and teaching. This path leads the use of innovative tools for students used to deal with a society full of information and knowledge.

MOBILE TECHNOLOGIES AND APPS FOR ENTREPRENEURSHIP AND LEARNING

The use of apps and mobile technology has become an increasingly tendency that promotes entrepreneurship and learning. This scenario represents a great potential for development and implementation of classroom and innovative projects. In this conference it could be observed researches, which contribute to the solution of social problems through autonomous learning, entrepreneurship and innovation (Barroso, 2017). Within the framework of this event, a series of proposals that respond to this tendency and that seek to promote virtual scenarios for the construction of a more equitable society were socialized (Tanos, De Castro, & Gaviria, 2017).

Another trend in this area, is the one that seeks to promote tourism as a strategy for the development of communities, through the promotion of people's cultural values, customs and traditions (Brito & Cortés, 2017). The experiences framed in this tendency uses technology as tools to raise awareness of the cultural richness of communities barely visible, with the purpose of favoring tourism as an economic and social activity.

The use of mobile applications has led researchers to propose experiences that contribute to improve the quality of people's life, and solve social and cultural problems presented in their environment. This type of applications offers the possibility of new opportunities to the most vulnerable communities

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

and those with less access to technology. Apps also play an important role in the education, when they represent alternative solutions associated with teaching strategies, the development of didactic activities and the implementation of the new assessment approaches (Espitia, 2017).

CONCLUSION

This paper has summarized a wide overview about the new perspective of educational informatics in Latin America presented in the XIII Conference of Educational Informatics. A wide range of theoretical breakthroughs and computing applications have been reported in the areas of *ICT for Education and Peace*, *Thinking and Artificial Intelligence*, *Emergent Scenarios in Education* and *Mobiles Technologies and Apps for Entrepreneurship and Learning*.

Researches in the area of *ICT for Education and Peace,* presented the use of ICT as a tool for re-socialization in conflict zones, the incorporation of technologies through social reintegration processes and the curricula redesign in terms of citizenship and peace culture. In this context, the role of the teacher is to lead processes that prevent, monitor and control educational problems through classroom projects, focusing on the development of values.

In *Thinking and Artificial Intelligence area,* the tendency was: the personalization of learning environments mediated by technologies; development of thinking skills; methodologies for the creation of robotic educational toys; cognitive robots; e-learning environments and the study of metacognitive processes that help developing educational resources in hardware and software.

Emerging Scenarios in Education area, researches were conducted on two main topics: *(i)* the implementation of educational video games in different areas of knowledge. Here the research process is focused on the formulation of classroom projects to reinforce ICT skills and generate new learning spaces through Gamification; *(ii)* Formulation of interdisciplinary works to identify skills and the required knowledge to lead researches and innovative projects.

Researchers in *Mobile Technologies and Apps for Entrepreneurship and Learning* area proposed: educational video games in different areas of knowledge, apps for classrooms and apps for solving social-cultural problems.

The future works in the field of educational informatics in Latin America are oriented to the implementation of cognitive informatics and cognitive computing in the development of new scenarios for teaching and learning, especially in the simulation of cognitive and metacognitive processes.

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

Barroso, D. (2017). Vale para Todos. In *Proceedings of XIII Congreso de Informática Educativa Versión Internacional,* Monteria (pp. 222–344). Retrieved from http://www.ribiecol.org/#memorias

Bejarano, G. (2017). Las Narraciones Digitales como Estrategia para promover la Expresión Libre y Creativa. In R. I. B. I. E. In (Ed.), *Proceedings of* XIII Conferencia Internacional en Informática Educativa, Monteria (pp. 298–308). Retrieved from http://www.ribiecol.org/#memorias

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bingimlas, K. A. (2009). Barriers to the Successful Integration of ICT in Teaching and Learning Environments: A Review of the Literature. *Eurasia Journal of Mathematics*, *Science & Technology Education*, *5*(3), 235–245. doi:10.1103/PhysRevA.71.032327

Brito, A., & Cortés, M. (2017). Diseño De Un Portal Web, Como Estrategia Publicitaria Para Incentivar El Turismo Del Balneario "El Pozo", En El Municipio De Hatonuevo Departamento De La Guajira. In *Proceedings of XIII Congreso de Informática Educativa Versión Internacional,* Monteria (pp. 632–638). Retrieved from http://www.ribiecol.org/#memorias

Bulla, J., & Solano, J. (2017). Uso y Apropiación de Tecnologías Digitales para el Aprendizaje de Lengua Castellana en Estudiantes de Grado Once. In *Proceedings of XIII Congreso de Informática Educativa Versión Internacional,* Monteria (pp. 662–673). Retrieved from http://www.ribiecol.org/#memorias

Caro, M., Toscano, R., Hernández, F., & David, M. (2009). Diseño de software educativo basado en competencias. *Ciencia E Ingenieria Neogranadina, 19*(1), 71-98. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=fua&AN=46984112&lang=es&site=ehost-live

Chiappe, A., & Sánchez, J. O. (2014). Informática educativa: Naturaleza y perspectivas de una interdisciplina. *Revista Electrónica de Investigación Educativa*, *16*(2), 135–151.

Díaz, A. (2017). Talleres Autodirigidos para Dinamizar el Aprendizaje Autonomo en el Área de Tecnologia e Informática. In *Proceedings of* XIII Conferencia Internacional en Informática Educativa, Monteria (pp. 322–331). Retrieved from http://www.ribiecol.org/#memorias

Espitia, N. (2017). Impacto del Uso de Entornos Tecnológicos Móviles en el Aprendizaje de las Matemáticas. In *Proceedings of XIII Congreso de Informática Educativa Versión Internacional,* Monteria (pp. 618–631). Retrieved from http://www.ribiecol.org/#memorias

García-Peñalvo, F. J., & Conde González, M. Á. (2013). Informática Educativa: Avances y Experiencias. *RED Revista de Educación a Distancia, XI*. Retrieved from http://www.um.es/ead/red/36/

García-Valcárcel, A., Basilotta, V., & López García, C. (2014). ICT in collaborative learning in the classrooms of primary and secondary education. *Comunicar*, *21*(42), 65–74. doi:10.3916/C42-2014-06

Hernández, J., Hernández, W., & Sanabria, L. (2017). Diseño de un Ambiente de Aprendizaje basado en Videojuego que articula Habilidades Espaciales y Potencial Creativo según el Estilo Cognitivo del Estudiante. In *Proceedings of XIII Conferencia Internacional en Informática Educativa version internacinal* (pp. 642–654). Monteria. Retrieved from http://www.ribiecol.org/#memorias

Hervatis, V., Loe, A., Barman, L., O'Donoghue, J., & Zary, N. (2015). A Conceptual Analytics Model for an Outcome-Driven Quality Management Framework as Part of Professional Healthcare Education. *JMIR Medical Education*, *1*(2), e11. doi:10.2196/mededu.4789 PMID:27731840

Jimoyiannis, A. (2009). Factors Determining Teachers' Beliefs and Perceptions of ICT in Education. In Encyclopedia of Information Communication Technology (pp. 321–323). doi:10.4018/978-1-59904-845-1.ch043

John, P. D., & Sutherland, R. (2004). Teaching and learning with ICT: New technology, new pedagogy? *Education Communication and Information*, *4*(1), 101–107. doi:10.1080/1463631042000210971

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kandiko, C. B. (2012). Leadership and creativity in higher education: The role of interdisciplinarity. *London Review of Education*, *10*(2), 191–200. doi:10.1080/14748460.2012.691283

Klein, J. T. (2010). A taxonomy of interdisciplinarity. In *The Oxford Handbook of Interdisciplinarity* (pp. 15–30). Retrieved from http://www.oup.com/us/catalog/general/subject/GeneralScience/?view=usa&ci=9780199236916

Levy, P., Ford, N., Foster, J., Madden, A. D., Miller, D., Nunes, J. M. B., ... Webber, S. (2003). Educational informatics: An emerging research agenda. *Journal of Information Science*, *29*(4), 298–310. doi:10.1177/01655515030294006

Martínez, A., Acevedo, S., & Correa, A. (2017). In *La Radio Escolar como Estrategia de Inclusión Digital para la Reconstrucción del Tejido Social con Jóvenes de la Guajira* (pp. 369–379). Monteria: XIII Congreso de Informática Educativa Versión Internacional; Retrieved from http://www.ribiecol.org/#memorias

Miller, G. A. (2003). The cognitive revolution: A historical perspective. *Trends in Cognitive Sciences*, *7*(3), 141–144. doi:10.1016/S1364-6613(03)00029-9 PMID:12639696

Palomino, M., Barrera, J., & Culchac, A. (2017). Sistema Recomendador de Materiales Educativos Digitales basado en Estilos de Aprendizaje. In *XIII Congreso de Informática Educativa Versión Internacional* (pp. 93–603). Monteria. Retrieved from http://www.ribiecol.org/#memorias

Patiño, A. (2017). Uso de los Videojuegos en Educación Básica, en el área de Tecnología. In *XIII Congreso de Informática Educativa Versión Internacional* (pp. 655–661). Monteria. Retrieved from http://www.ribiecol.org/#memorias

Prieto, L., & Nájar, O. (2017). Derechos Humanos Aprendizajes de Paz y TIC : Herramientas para la Construcción de Las TIC como tema de inclusión social para la rehabilitación y resocialización de la población recluida en el área de reflexión de la Policía Nacional del Municipio de Fonse. In *Proceedings of XIII Conferencia Internacional en Informática Educativa* (pp. 425–430). Retrieved from http://www.ribiecol.org/#memorias

Ramírez, J., & Solano, J. (2017). Ambientes Digitales de Aprendizaje: Una Aproximación Pedagógica, Tecnológica y Comunicativa para la Educación a Distancia en la Universidad del Tolima. In *Proceedings of XIII Congreso de Informática Educativa* (pp. 674–681). Monteria. Retrieved from http://www.ribiecol.org/#memorias

Rangel, A., & Silva, C. (2017). Minuto para la paz. Propuesta radial para el desarrollo de competencias ciudadanas y construcción de paz en la era del Postacuerdo Derechos Humanos Aprendizajes de Paz y TIC : Herramientas para la Construcción de Las TIC como tema de inclusión social par. In *Proceedings of* XIII Conferencia Internacional en Informática Educativa, Monteria (pp. 264–273). Retrieved from http://www.ribiecol.org/#memorias

República de Colombia, Congreso de Colombia. (2014). Ley1732. Por la cual se establece la cátedra de la Paz en todas las Instituciones educativas del país, Pub. L. No. 1732 Retrieved from http://wsp.presidencia.gov.co/Normativa/Leyes/Documents/LEY 1732 DEL 01 DE SEPTIEMBRE DE 2014.pdf

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sandoval, A. (2017). Mediaciones Pedagógicas con TIC en los Proyectos de Aula de los Maestros en Formación de la LIT. In Proceedings of XIII Conferencia Internacional en Informática Educativa, Monteria (pp. 122–132). Monteria: XIII Conferencia Internacional en Informática Educativa; Retrieved from http://www.ribiecol.org/#memorias

Sierra-Rodríguez, J. L., & García-Peñalvo, F. J. (2015). Informática Educativa y Educación en Informática. *Computers in Education and Education in Computer Science.*, *16*(4), 25–31. doi:10.14201/eks20151642531

Srivastava, P. (2012). Educational informatics: An era in education. In *Proceedings of the 2012 IEEE International Conference on Technology Enhanced Education ICTEE '12*. doi:10.1109/ICTEE.2012.6208613

Suarez, J., & Garces, N. (2017). Diseño y Validación de una Metodología para el Desarrollo de Juguetes Robóticos Educativos. In *Proceedings of XIII Congreso de Informática Educativa Versión Internacional* (pp. 523–543).

Tanos, L., De Castro, L., & Gaviria, M. (2017). El Ambiente Virtual de Aprendizaje: Una Forma de Aprender sobre Sexualidad, Salud y Construcción de Ciudadanía. In *Proceedings of XIII Congreso de Informática Educativa Versión Internacional* (pp. 346–356). Retrieved from http://www.ribiecol.org/#memorias

Tondeur, J., Coenders, A., van Braak, J., ten Brummelhuis, A., & Vanderlinde, R. (2009). Using online tools to support technology integration in education. Handbook of Research on New Media Literacy at the K-12 Level: Issues and Challenges (Vol. 1). doi:10.4018/978-1-60566-120-9.ch025

Van Rijnsoever, F. J., & Hessels, L. K. (2011). Factors associated with disciplinary and interdisciplinary research collaboration. *Research Policy*, *40*(3), 463–472. doi:10.1016/j.respol.2010.11.001

Wang, Y. (2002). On cognitive informatics. In *Proceedings of the 1st IEEE International Conference on Cognitive Informatics, ICCI 2002* (pp. 34–42). doi:10.1109/COGINF.2002.1039280

Wang, Y., Berwick, R. C., Haykin, S., Pedrycz, W., Baciu, G., & Bhavsar, V. C., … Zhang, D. (2011). Cognitive informatics in year 10 and beyond: Summary of the plenary panel. In *Proceedings of the 10th IEEE International Conference on Cognitive Informatics and Cognitive Computing, ICCI*CC 2011* (pp. 11–22). 10.1109/COGINF.2011.6016117

Yagama, S., & Sierra, R. (2017). Consejos Estudiantiles mediadores en la Cátedra De Paz. In Proceedings of XIII Conferencia Internacional en Informática Educativa, Monteria (pp. 85–96). Retrieved from http://www.ribiecol.org/#memorias

Zarate, E., & Medina, D. (2017). Las TIC como tema de inclusión social para la rehabilitación y resocialización de la población recluida en el área de reflexión de la Policía Nacional del Municipio de Fonseca La Guajira. In *Proceedings of XIII Conferencia Internacional en Informática Educativa* (pp. 425–430). Monteria: RIBIE. Retrieved from http://www.ribiecol.org/#memorias

This research was previously published in the International Journal of Software Science and Computational Intelligence (IJSSCI), 10(1); edited by Yingxu Wang; pages 80-87; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 48
Patient Data De-Identification: A Conditional Random-Field-Based Supervised Approach

Shweta Yadav
Indian Institute of Technology Patna, India

Sriparna Saha
Indian Institute of Technology Patna, India

Asif Ekbal
Indian Institute of Technology Patna, India

Parth S Pathak
ezDI, LLC, India

Pushpak Bhattacharyya
Indian Institute of Technology Patna, India

ABSTRACT

With the rapid increment in the clinical text, de-identification of patient Protected Health Information (PHI) has drawn significant attention in recent past. This aims for automatic identification and removal of the patient Protected Health Information from medical records. This paper proposes a supervised machine learning technique for solving the problem of patient data de- identification. In the current paper, we provide an insight into the de-identification task, its major challenges, techniques to address challenges, detailed analysis of the results and direction of future improvement. We extract several features by studying the properties of the datasets and the domain. We build our model based on the 2014 i2b2 (Informatics for Integrating Biology to the Bedside) de-identification challenge. Experiments show that the proposed system is highly accurate in de-identification of the medical records. The system achieves the final recall, precision and F-score of 95.69%, 99.31%, and 97.46%, respectively.

INTRODUCTION

With the start of the golden era in the medical interpretation, the vast amount of information in the clinical domain is increasing at a rapid rate. In the past decade, with the development of the health information technology and health data documentation, there has been progress in how heath care is performed (Berner et al., 2005).

DOI: 10.4018/978-1-7998-0951-7.ch048

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

With the widespread use of health information technology, there has been huge pace in the increment of clinical data in addition to the fast adoption of the Electronic Clinical Records and with the conversion of narrative data to the electronic form. The amount of information can be improved further with the minimization of the medical error. This requires the development of some sophisticated tools for Medical Language Processing (MLP). Most medical records are in the narrative forms which are formed as the result of transcription of dictations, direct entry by providers, or use of speech recognition applications. However, their use in this form is restricted to any organization or research, as medical records have a sufficient number of personal health information or protected health information (PHI). According to Health Insurance Portability and Accountability Act (HIPAA), 1996, the PHI terms need to be enclosed and protected. This has lead to de- identification problem. Paragraph 164.514 of the Administrative Simplification Regulations promulgated under the Health Insurance Portability and Accountability Act (HIPAA) states that for data to be treated as de-identified, it must clear one of two hurdles (HIPPA ACT 1996).

1. An expert must determine and document "that the risk is very small that the information could be used, alone or in combination with other reasonably available information, by an anticipated recipient to identify an individual who is a subject of the information."
2. Or, the data must be purged from a specified list of seventeen categories of possible identifiers relating to the patient or relatives, household members and employers, and any other information that may make it possible to identify the individual.

Studies showed that there was a significant drop in the patient consent request reducing the participation rate and also, this is quite infeasible for the huge population. Even, in the case when a patient provides the permission, documents must be tracked to stop any unauthorized disclosure. This emerging problem of consent, waiver, and tracking can be effectively handled if the patient personal health information is properly de-identified facilitating the clinical NLP research (Wolf & Bennett, 2006).

De-identification task is more specifically defined as the step where the private information is removed or replaced while keeping the record as it is (Stubbs et al., 2015). De-identification is a type of traditional named entity recognition (NER) problem, with the property of defining a term to be PHI type or not. The main aim of de-identification challenge as pointed out earlier is to remove the PHI terms maintaining data integrity as much as possible. Every record is enclosed in the RECORD_ tags and is provided a unique ID which is randomly generated. Figure 1 shows Sample Discharge Summary Excerpt; a sample discharge summary from the training dataset where the goal is to identify the PHI (private health information) terms. In this summary, some the PHI terms are doctors' name ("Dr. Do Little"), patient name ("John Doe") and hospital name ("ABHG", "SBHG"). A TEXT_ tag encloses the text of different records. Each PHI instance is enclosed within PHI_ tags and the PHI TYPE represents the category of the PHI term as shown in Figure 1.

As shown in Figure 1, the task is to enclose the PHI terms such as ("Dr. Do Little", "John Doe", "ABHG", "SBHG"). This task can be seen as the typical sequence labeling task where for e.g. "Dr. Do Little" should be labeled as whole with Doctor. Here, we present an example where input forms the sequence of words and output is the label sequence and its corresponding de-identified sentence. "BIO" notation was followed to label the NE where "B" represents the beginning of label sequence, "I" denotes intermediate of label sequence and "O" represents others.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Sample discharge summary excerpt

```
<RECORD ID="641">
<TEXT>
<PHI TYPE="HOSPITAL">SBHG</PHI>
HISTORY OF PRESENT ILLNESS:
Mr. <PHI TYPE="PATIENT">Doe</PHI> is a XX-year-old white male with a history of
diabetes mellitus , inferior myocardial infarction , who underwent open repair of his
increased diverticulum <PHI TYPE="DATE">January 1st</PHI> at <PHI
TYPE="HOSPITAL"> ABHG</PHI> .
No cardiopulmonary resuscitation was performed as per the patient and family wishes
DISCHARGE SUMMARY
ARF32 FA
DISCHARGE SUMMARY NAME:
UNIT NUMBER:
<PHI TYPE="PATIENT">JOHN DOE</PHI>
ADMISSION DATE:
<PHI TYPE="DATE">01/01/1900 </PHI>
DISCHARGE DATE:
<PHI TYPE="DATE">01/02/1900 </PHI>
DD:
<PHI TYPE="DOCTOR">Dr. Do little</PHI>, M.D.
</RECORD>
```

Table 1. Example of PDI task following BIO notation

Sentence	Mr.	John	D	Doe	is	xx-year-old	white	male
Named Identity	O	B-Doctor	I-Doctor	I-Doctor	O	O	O	O
De-Identified Sentence	Mr.	XYZ_Doctor			is	xx-year-old	White	male

In the past few years, de-identification of the clinical record has become the booming research area. Sixth Message Understanding Conference (MUC) had defined the Named Entity Recognition (NER) problem with an objective to identify the relevant entities from the text. NER forms an important task in NLP. Similarly, in the medical domain, NER is very crucial in the identification of the clinical terms by which we can get insights to the problem, medication, procedure, and examination.

Major research works for entity extraction in the medical domain had used machine learning and rule-based concepts. MedLEE (Friedman et al., 1996) is one such system in clinical domain proposed by Carol Friedman et al. at Columbia University. Aronson (2001) developed MetaMap system at National Library of Medicine. Another system cTAKES (Savova et al., 2010), developed using Unstructured Information Management Architecture (UIMA) framework and OpenNLP, a widely accepted system.

Recently, the Medical Natural Language Processing Community had organized a number of challenges for NER task. The earliest shared task is organized by Center of Informatics for Integrating Biology and the Bedside (i2b2) in 2009 on medical name recognition (Uzuner et al., 2010). In 2010, i2b2 shared task was conducted with the aim to identify the medical problems, test entities and various treatments from the clinical records (Uzuner et al., 2011). Since then, Conference and Labs of the Evaluation Forum (CLEF) had organized a challenge in 2013 with the aim of solving disorder mention recognition and normalization (Goeuriot et al., 2013). Very recently, i2b2 in 2014 (Stubbs et al., 2015) had organized a challenge for patient data de-identification problem and CUI (concept unique identifier) detection. Semantic Evaluation (SemEval) 2014 also came up with the disorder mention recognition and the normalization. With a large number of challenges organized almost every year, it had drawn the attention of many researchers to solve these real-life challenging problems.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Different approaches have been used to solve this problem ranging from machine learning techniques (Stubbs et al., 2015; Yang et al., 2015), rule-based system (Beckwith et al., 2006; Berman et al., 2003) to the hybrid system. Rule-based systems are designed over dictionaries with the rules that are designed manually to recognize the PHI terms from the text. Building a system using rules lacks generalization capability as well as it consumes more time and skill for developing rules. Rule-based systems are seen to provide good performance for the PHI terms those appear less often. This has lead to the development of some machine learning techniques for automatic extraction of PHI terms from the text which makes the process more generalized. The only drawback of machine learning based approach is the requirement of having large annotated corpus which can be developed manually.

Generally, the NER problem is the sequence labeling problem where the aim is to identify the proper output sequence of the entities. Thus, for the given input sequence of words, we need to obtain the best labeled- sequence. Researchers have primarily used supervised machine learning algorithms to solve this kind of medical NER problem. The algorithms used for extracting medical entities are Conditional Random Fields (CRFs) (Lafferty et al., 2001), Maximum Entropy (ME) (McCallum et al., 2000), and Structural Support Vector Machines (SSVMs) (Xue et al., 2008). Among these different classifiers, Conditional Random Field had drawn a huge attention for solving the sequence labeling task. Top best system (Yang et al., 2015) of de-identification task had used CRF as the underlying classifier because of the property of CRF to properly map the relationships between their surrounding tokens. Supervised machine learning techniques are able to capture the handcrafted features very effectively. Some of the prominent features which have been used are orthographic features that capture the word structures, syntactic features like Part of Speech (PoS), and semantic features as well as the discourse features. These features are typically used for solving any NER problems. Many other domain specific features could be generated for further improving the performance of the system.

In the current work, we develop a supervised machine learning based classification framework for solving the patient de-identification problem. The proposed method is evaluated on the i2b2 2014 dataset. It concerned identifying four different PHI categories, namely PATIENT, DOCTOR, HOSPITAL names and DATE from the medical records. Taking into account the effectiveness of applying CRF in solving the NER problem, CRF was adopted to train the model. Diverse features are extracted and CRF based model was trained using this huge collection of features. Several experiments were carried out with different cross folds. Results of the proposed system are shown for different cross-validation settings like 5,6,7,8,9,10 fold cross validations.

This task suffers from various challenges as similar to any NLP task. One such challenge is the ambiguity problem where we found that due to the lexical similarity PHI terms mostly overlap with the non-PHI terms. An example includes Brown (Doctor name) which is a PHI term vs. brown (Colour name) which is a non-PHI term. Another major challenge is the occurrence of some out-of-vocabulary PHI terms which are not present in the dictionaries. These words act like the foreign words which can be misspelled. Gazetteers list, dictionaries are the mostly used for solving any NER problem but the contribution of these types of words lowers the performance and creates some errors. Results obtained prove that the proposed system outperforms the existing systems in identifying four different PHI terms.

The remaining paper is organized as follows. Next immediate section discusses the background of de-identification task and its significance. This is followed by a section which provides the insight to the related work. Next section discussed the methodology adopted for the de-identification which is followed by the details of experiments and their analysis. Finally, we conclude in the last section.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

BACKGROUND AND SIGNIFICANCE

Most of the works done in biomedical Natural Language Processing (NLP) domain can be distinguished into two different types on the basis of the text, biomedical text and clinical text (Meystre et al., 2010). Biomedical texts are the texts that appear in the journals, articles, posters etc. On the other hand, the clinical texts are mostly the texts written by the doctors in the form of prescriptions for the patients. This text comprises of the patient information like his/her name, age, gender, the procedure that was applied to him/her and what are his/her medical histories. These texts also provide insights to the findings made during procedure and the medication provided to them.

The clinical text is the whole range of narrative found in the patient records. These texts vary from the short-segment which gives the chief complaints to the long-text comprising of the history of the patients. There is another class of text which is discussed in the literature very often (Coorevits et al., 2013). While some of the texts are similar to that of biomedical texts for example internal research reports, the others appear to be like clinical texts for example patient notes. As these narratives are rarely exposed outside the organizations, there is a very sparse study on them.

There is one obvious question that what distinguishes clinical text from the biomedical text and why processing of these poses a challenge to the NLP society. Firstly, in the clinical texts, a large number of grammatical errors can be found which may have occurred at the telegraphic phrase. Other segments in clinical documents which create some problem are discharged summaries and the lab test reports. Moreover, clinical documents suffer from the problems like abbreviations and acronyms. These lexical structures are often overloaded i.e., there is a multiple rendering for the same word set. Literature shows that almost 33% of the times, acronyms are overloaded which makes it prone to the ambiguity problem (Liu et al., 2001). There is a huge problem due to the spelling mistake, which is observed most commonly in the notes lacking spelling support. US Veterans Administration (VA) has one of the largest EHR systems with only the basic text support. Even in their corpus, there are lots of abbreviations and acronym problem. Furthermore, the clinical data consists of different types of characters which appear anywhere within the text. Another problem that is raised is to change the unstructured narrative text into the structured; templates along with pseudo-tables are in general used in the method of conversion (Meystre et al., 2010). Template having a narrative structure for patient history or the discharge summary that is common among the medical society provides benefits to NLP. Explicit templates are formatted text with the fields specific to the institution are filled by the user. These different variations in template make this domain quite challenging.

There are many methodologies adopted to cover different types of questions related to clinical research domain (Dehghan et al., 2015), (Liu et al., 2015), (He et al., 2015). Although, the pharmaceutical industry covers the clinical trails and therapy but nowadays, these industries are spreading the horizon of solving the medication issues like drug interaction, disease prevention through the proper investigation of diseases. These include the course of the disease and various other factors like, what were the different criteria for diagnosis, how the patient education affected his/her health and his/her surveillance. With the rapid and booming research interest in the clinical domain, the studies have also spread to obtain the role of different gene pathways in the identification of their relations in the development of diseases. Some clinical research focuses on improving the health care with lowering the health care cost, thus providing the benefits to the mass. As such, it is highly important to have a proper study of the clinical records in solving the major unsolved and challenging problems of the society with the NLP system.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

De-identification task forms one such challenge which aims to identify the noun phrases or part of a noun phrase automatically from the text. In the traditional named entity recognition task, originated by the MUC challenge, the focus was to identify the person, organization, location names from the newspaper type electronic text. Generally, de-identification task is the task of identifying and enclosing patient personal health information like not just identifying the persons but also identifying the type of person (Patient, Doctor), Organizations (Hospital), Location (Streets, Country), ages, dates but not in the term of years. Apart from these, structures of the text too vary a lot. While in newswire or journalistic domain, lots of clues can be obtained from their titles, abstract and cases while in the patient records, due to its semi-structured form the task becomes quite challenging.

LITERATURE SURVEY

De-identification task is a type of the traditional Named Entity Recognition (NER) problem. Message Understanding Conference had been the center for the evaluation of NER system in several domains. MUC had organized several challenges to identify several named entities like person, location, organization, time, dates etc. For the MUC-6 NER task, a total of 15 teams participated in the challenge submitting 20 different systems. The best F-Measure value reported was 96.42% (Sundheim, 1997). In the MUC-7 challenge, 14 systems were submitted from 12 different teams obtaining the highest F-Measure value of 93.39% (Chinchor, 2001). National Institute of Standards and Technology (NIST) organized two different tasks, Information Extraction Named Entity task and the Automatic Content Extraction (ACE) tasks. Both of these tasks were evaluated on the newswire domain.

In the biomedical domain, Critical Assessment of Information Extraction Systems in Biology (BioCreAtIvE) (Hirschman et al.,2005), organized several biomedical challenges including the entity extraction task, where the task is to identify different entities like gene name, protein name from the given biomedical text derived from the PubMed, PMC. PubMed and PMC are huge repositories consisting of almost 3.5 million biomedical literature. For the evaluation, TREC Genomics Track was considered. The BioNLP 2009 challenge (Kim et al., 2009) focuses on event extraction primarily on the protein and gene event extractions. The same challenge was also focussed on BioNLP 11(Pyysalo et al., 2011) and BioNLP 13 shared task (Nedellec et al., 2013), where data was used from MEDLINE. Another task BioASQ was organized for entity extraction using data from PubMed. Several challenges have been organized for entity extraction. Recently, i2b2 had organized the task for the de-identification of the clinical terms in the years 2006 and 2014 (Stubbs et al., 2015). Here, the challenge was to identify the certain PHI terms.

In the 2006 shared task, a total of 889 patient records were used. This challenge got sixteen submissions from seven different teams. The challenge had used longitudinal records of each patient rather than the individual records. It was observed that longitudinal records contain more personal information of patients rather than an individual record. Completely identifying the patient information from individual records is a problematic task. This makes the de-identification using the longitudinal records a more complicated task. The submitted systems identified the task as a token classification task (using Support Vector Machine) while some distinguished it as the sequence labeling task (using Hidden Markov Model, Conditional Random Field).

Wellner et al. (2007) in i2b2-2006 challenge built the best system based on CRF. They used some lexical and semantics features. They adapted two important NER systems, one based on the CRF, Carafe, and other Lingpipe, HMM-based implementation. The carafe was seen to be highly effective in capturing

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

the PHI terms. Szarvas et al. (2007) adopted an iterative method for named-entity de-identification. They treat de-identification as a classification task and use machine learning based decision tree with local features and dictionaries. Aramaki et al. (2006) also used CRF as their classifier with the features such as local, external and global features. The global feature includes the features which capture position, the length of a sentence and also several tokens in the previous sentence. External features include the dictionary for location, person name, and dates. They performed two-phase learning wherein the first phase, they learned their system using local features and in the second phase, they include first phase features set in addition to the most frequent label. They identified that sentence feature is very informative in identifying dates, IDs and patient because of the appearance of this PHI in beginning of the records while the external feature was useful in identifying dates, locations, doctors and patients.

While many systems have used CRF based classifier, some submitted systems also used SVM. Hara (2006) developed a hybrid system making use of rules. The rules identified the patterns for the PHI terms such as phone numbers, IDs, and dates. They train SVM on the global and local features which were found to be very helpful in identifying the PHI terms such as a hospital, patient, location, doctors, and age. In addition to these features, their system included the section heading which is closed to the target, category of the sentence, the surface form of target, the root, PoS tag and orthography features. The system used two types of the ordered tree, dependency tree and "n-grams trees" making use of the Boosting Algorithm for Classification Trees (BACT). With this information and in addition of chunking technique, they identified the position and the text span of PHI terms such as a patient, doctor, location, and hospital. The results show that performance is degraded with the incorporation of sentence feature.

In 2014 patient de-identification task, machine learning technique was more prominent. According to the 2014 i2b2/UTHealth shared task, the challenge was to de-identify the data in a stricter way as compared to the HIPAA standards in the form of additional categories. The task was to identify the 25 PHI categories with the associated subcategories from the clinical records like professions, time and dates, facilities etc. These tasks also consider the longitudinal records. Yang et al. (2015) submitted the system in the i2b2-2014 challenge and they reported the accuracy of 93.60%. This was the best performing system. They developed a hybrid model which used both the machine learning and the rule-based technique to extract the PHI terms. They exploited the various features like linguistic features, syntactic feature and different word surface oriented features which were further improved by using the task-specific features. They also used the regular expression pattern to identify the PHI terms like date, ID.

The second best system (Dehghan et al., 2015) from Manchester used a combined system developed using the knowledge-based technique as well as the rule-based system. The rules were defined to obtain the orthographic features, different patterns. They used CRF as the underlined classifier to obtain the PHI terms. For some categories like PATIENT, DOCTOR, HOSPITAL, the initial annotation was extracted at the patient level and dictionary was created for the patient. At the last stage, the outputs of the rule-based system, as well as the CRF based model, were integrated on the basis of some rules. These systems were highly efficient in obtaining the PHI terms such as the DATE, DOCTORS, and PATIENT.

System (Liu et al., 2015) from Harbin Institute of Technology Shenzhen Graduate School in i2b2 challenge used two different variations of the CRF. They define one CRF system working at the token level while another CRF system that works at the character level. They performed feature engineering and used the features such as bag-of-words, section relevant features, lexical features, semantic features. Apart from the two different CRF outputs, the third output was obtained by developing the rules using the Rule-based system. The final outputs of all the three systems were merged on the basis of the output

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

overlapping. If there was no overlapping i.e., if all the three systems obtained the same PHI tag, then that tag was accepted else the tag was obtained on the basis of the preference. They give the first preference to the rule based system, then the character-level system and lastly to the token-level based CRF system.

Another submitted system from Harbin Institute of Technology (He et al., 2015) used OpenNLP system to pre-process their data, sentence splitting, and tokenizer. They trained their system using CRF making use of features such as lexical, syntactic and orthographic features. They did not use external features such as dictionaries to identify the terms. The system developed by Kaiser Permanente (Torii et al., 2014) adapted the MIST tool. They include some additional rules that used lexicon for Location and Profession and different regular expressions for the categories such as Phone Numbers, ZIP and Organization. In addition to MIST system, the team learned NER model using Stanford NER system. The system then merges the overlapped identified entities. They reported the F-Measure value of 92.21%.

Through the literature survey, we found that hybrid systems (both the rule-based system as well as the machine learning based) perform well when compared with the individual techniques. It is reported that among top ten ranked systems, seven systems have adopted the rule-based technique. MedEx (Xu et al., 2010) forms one such system proposed at Vanderbilt University that utilizes the biomedical dictionary to identify the important medical concepts.

Overall, the best systems report high recall and precision, often greater than 90%, and sometimes as high as 99%. Nevertheless, no study has evaluated the performance of automated de-identification for all PHI classes (Meystre et al., 2010). Only a handful of studies provide details on over-scrubbing (non-PHI wrongly identified as PHI) and none of them investigate the effect of de-identification on subsequent IE tasks. It is indeed possible that de-identification has an adverse effect on IE accuracy (Meystre et al., 2010). Over-scrubbing errors could overlap with useful information, for example, if a disease name is erroneously recognized as a person name, it will be removed and lost to subsequent IE applications. Second, NLP techniques such as part-of-speech tagging and parsing may be less effective on modified texts.

It can be concluded that though a lot of systems have evolved in the recent past, but still there is a need to develop a more accurate automated de-identification system.

METHODS

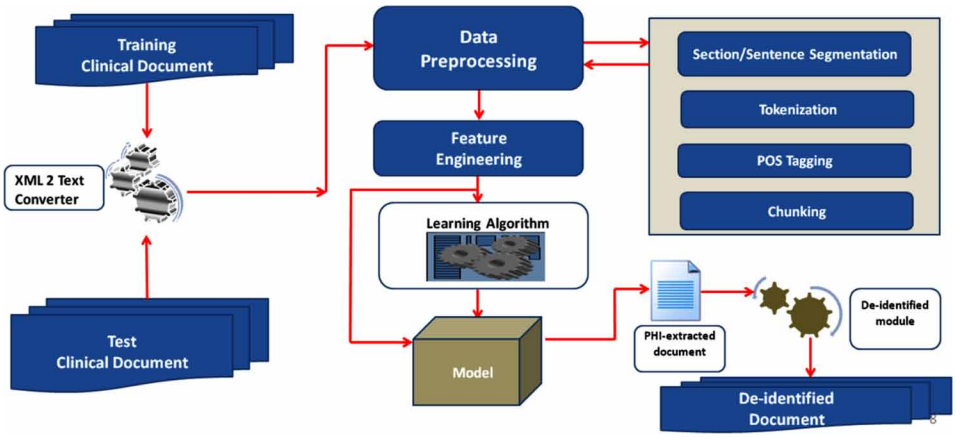
Supervised machine learning based classification frameworks have been developed to detect the PHI tokens from the medical records. Similar to any entity extraction technique, de-identification task is also considered to be the sequence labeling task. The proposed system architecture is divided into four different modules. The first module deals with the pre-processing of data. The second module represents the feature engineering part where features are generated. Training of the model through Conditional Random Field (CRF) is done in the third module. The fourth module represents the de-identification of the PHI terms. Detailed descriptions of the methodology adopted for the system are given below. Figure 2 represents the Proposed System Architecture.

Pre-Processing of the Medical Records

Datasets provided with the i2b2 de-identification challenge were untokenized; this necessitates the use of the pre-processing module for the process of tokenization. The pre-processing was performed in three different stages. The first stage of the pre-processing module deals with the sentence splitting. The second

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Proposed system architecture



module focuses on the tokenization of the medical records. The last and final stages are used to identify the rules for tokens such that when they are divided, they should match some patterns.

Apart from this process, pre-processing steps also include shallow parsing and Part of Speech (PoS) tagging to obtain the word lemma, chunk and part-of-speech of the corresponding words. Additionally, we extracted some of the document- level features, for example obtaining the different sections (HISTORY OF PATIENT ILLNESS, MEDICATIONS etc.) and position of the sentence in the text (beginning of the text, end of the text) using some of the hand-crafted rules.

Feature Generation

This module focuses on developing some prominent features. The features were developed, such as the linguistic, syntactic features as well as word surface-oriented features in order to capture the semantics of the PHI terms. The features were further enhanced by embedding the feature set by the task-specific features. The list of the extracted features for the PHI recognition task is given below:

- **Token Features:** Some features are generated during the pre-processing step like word lemma, word chunk and the Part-of-Speech tag of the current word.
- **Bag-of-Word Feature:** It includes the unigrams, bigrams, trigrams of the target token within the window size of [-2, 2]. An *n*-gram is a contiguous sequence of *n* items from a given sequence of text or speech. An *n*-gram of size 1 is referred to as an unigram, size 2 is a bigram and size 3 is a trigram.
- **Contextual Features:** This feature is the combined feature which includes the lemma, chunk, POS tag of the surrounding tokens. We consider the window size of [-3, 3]. These features give the context knowledge for each word.
- **Combined POS-Token Feature:** The combination of POS tag within the context of [-1, 1] and the current token $\{w_0p_{-1}, w_0p_0, w_0p_1\}$ is formed, where w_0 represents the current word, and p_{-1}, p_0 and p_1 represent the previous POS tag, current POS tag and the next POS tags respectively.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Some of the orthographic features with their examples in uses

Feature	Example	Feature	Example
InitCap	David	CapMixAlpha	EpoR
InCap	mAb	DigitSpecial	12-4
DigitOnly	12,23	AlphaDigitAlpha	IL34S,EIB
DigitAlpha	28BC	CapLowAlpha	Srs, Ras, Epo
Hyphen	_	DigitCommaDigit	2,34
CapsAndDigits	28Houston	LowMixAlpha	mAB,mRNA
StopWord	at, on	AllCaps	DHC
AlphaDigit	s72	EndwithDot	Dr., Mr.

- **Affixes:** This feature includes prefix and suffix of the word. It refers to the fixed length characters stripped either from the left most part or the rightmost part of the token. For example, the suffix and prefix of word 'health' are h, he, hea, heal and h, th, lth, alth respectively. Here, we set the length to size 4.
- **Sentence Information Feature:** This feature includes the structure of the sentence. It includes the information regarding length of the sentence, whether the sentence has any delimiters like '.', '?', '!', whether there is any bracket unmatched.
- **Orthographic Feature:** Several orthographic features were developed to capture the capitalization and digit information. These are very useful information to capture the entity extraction. Symbol such as the ',' is found to be very crucial in detecting the boundaries of the NEs. Table 2 lists some of the orthographic features with their examples in uses.
- **Word Normalization:** Two different variants of this feature were developed. The first feature is used to capture the words which are in plural form, verb, alphanumeric letters, hyphen, and digit. This feature actually converts the words to its root form. Another feature represents the orthographic construction of the words. This feature is termed as the word shape feature where each word is mapped to its respective class. Words are normalized by converting every uppercase character by 'A', lowercase character is changed to 'a' and every digit by '0'. For example, if the token is 'Medical-49', its normalized form would be 'Aaaaaaa-00'.
- **Word Class Feature:** This feature is generated motivated by the fact that similar entities reside within the similar class. This feature is the enhancement of the word normalization feature. After the word conversion, the squeezing operation is performed where the consecutive sequence of words is converted into the single character. For example for the word 'Medical-49', its word class form would be 'Aa-0'.
- **Section Feature:** This binary feature was generated after extracting the sections from the training set manually. Top 16 sections were extracted with their probabilities of carrying the relevant PHI terms. These sections were generated manually and form very effective feature in the identification of the PHI terms.
- **Gazetteer List:** For the identification of the PHI terms, task related term list helps in improving the performance of the system. These lists include the headings which give very strong remarks about the occurrence of the PHI terms. Example: 'Transcribed by', 'Dictated by'. These sections carry information about the occurrence of the DOCTOR name. Apart from this, the dictionary

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

was used for the names like the common names across countries, full name and the abbreviation of the common US locations for e.g. "New York" and "NY". We used the list of names for the days, months.

In addition, some lexical clues were identified for each individual PHI category. These are the types of the trigger words that give an indication regarding the presence of the particular PHI terms for example 'Dr.', 'M.D' in the case of DOCTOR, 'Mr.', 'Mrs' in the case of the patient. These lexical clues are obtained from the neighboring contexts of the target PHI terms. These are filtered on the basis of occurrence and their importance with the target PHI categories.

- **General Entity Information:** The Stanford Named Entity Recognizer (http://nlp.stanford.edu/software/CRF-NER.shtml) was used to obtain the named entity tag of the corresponding token.
- **Discourse Features:** This feature plays a very important role in the identification of the terms that are very much related to section headings and their positions of occurrence in the sentence. For example the strong contextual clues words like'CC:' and 'Dictated by'.

Classification: CRF Classifier

We trained our system using the Conditional Random Field algorithm. This algorithm is very popular in the domain of entity recognition because of its ability to tackle the sequence labeling task very efficiently. CRF is categorized under the class of statistical modeling. They are mainly applied for the structural prediction which is formed by the combination of classification and graphical modeling methods. CRF integrates the performance of the graphical models to compactly model multivariate data and classification methods to perform prediction using a large set of input features.

The CRF-classifier works on the Markov Assumption that given the observation sequence of the token $x_1, x_2, x_3 \ldots\ldots\ldots x_n$, the corresponding label-sequence $y_1, y_2, y_3 \ldots\ldots\ldots y_n$ is predicted by the CRF. It maximizes the conditional probability for every y in the possible set of label sequence. Mathematically, CRF can be represented as

$$\exp\left(\sum_{j} \gamma_j t_j(y_{(i-1)}, y_i, x, i) + \sum_{k} \mu_k s_k(y_i, x, i)\right) \quad (1)$$

Here, $t_j(y_{i-1}, y_i, x, i)$ represents the transition feature function for the whole set of observation sequence. i and $(i - 1)$ denote the position of the labels in the label sequence. Here $s_k(y_i, x, i)$ represents the state feature function at the label position i and observation sequence. γ_j and μ_k are the parameters estimated from the training dataset. While defining the feature function, the real valued feature set of the observation is constructed as $b(x,i)$. It expresses empirical distribution characteristics of the training data in order to hold the model distribution.

There are various other algorithms that can handle the sequence labeling task like Maximum Entropy Markov Model (MEMM) and Hidden Markov Model (HMM). The main advantage of CRF over HMM is that it is conditional in nature, which helps in relaxation of the independence assumption as needed by the HMM. Apart from that, CRF is very much efficient in handling the label bias problem. Label bias problem is due to the directed graphical model that is exhibited by MEMM and HMM. It was observed that CRF outperforms the other models in a number of tasks such as speech recognition, bioinformatics.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

De-identification task targets to automatically extract the PHI terms with the proper boundary identification. As CRF being the better performer than other models, this work used CRF as an underline base classifier. The open source toolkit CRF++ was used to carry out our experiments. BIO notation was followed to label the PHI terms which occur in sequence. Here 'B' identifies the beginning of the PHI token, 'I' represents that the PHI token is within the sequence. 'O' is used to identify that token belongs to none of the PHI term categories.

For the classifier learning, whole feature set discussed in the previous section was used. CRF modeling is done by considering all the PHI categories through learning by the training phase and predicting the output of the model built during the test phase.

De-Identification of PHI Terms

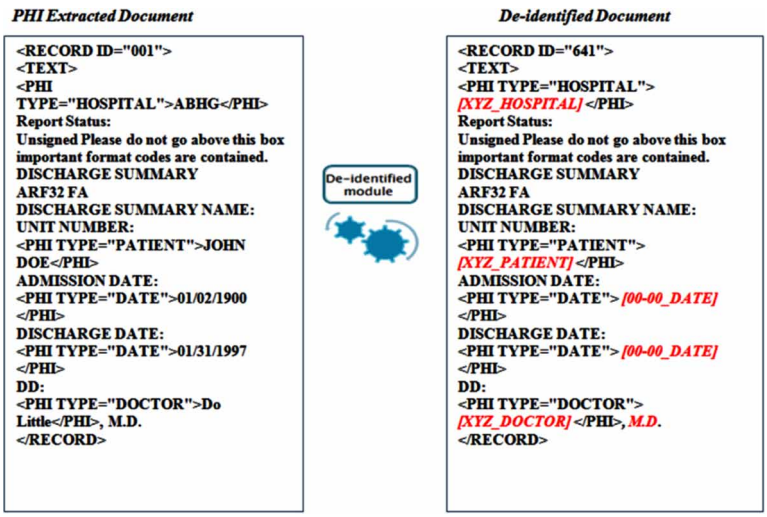
In order to hide the patient personal information, de- identification of the PHI terms is highly necessary. A template was adopted through which we convert all PATIENT, HOSPITAL, DOCTOR names present into a generic format like XYZ_PATIENT, XYZ_HOSPITAL, XYZ_DOCTOR, respectively and all the dates into the format 00-00_DATE. Thus, capturing the information required without compromising the personal details. We represent this below in Figure 3 by considering the original document and the de-identified document. Figure 3 shows De-Identification Process.

DATASET AND EXPERIMENTS

Dataset

The 2014 i2b2 challenge dataset was used in the current paper, which was taken from the Research Patient Data Repository of Partners' Healthcare. A total of 1304 medical records basically of longitudinal nature were manually annotated. These records were obtained from 297 patients following the annotation

Figure 3. De-identification process



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

guidelines. The dataset was divided into two parts where the first part consists of 790 records of 188 patients that were used as the training dataset and 514 records of 109 patients were test dataset.

i2b2-2014 challenge dataset compromises of 17,045 PHI relevant instances in the training dataset, while the test dataset consists of total 11,462 PHI instances. To ensure the patient confidentiality as much as possible, the challenge aims to identify HIPAA- PHI categories firstly with the added subcategories. They were annotated using seven main PHI categories with the twenty- five associated subcategories. While our experiments cover the seven main PHI categories, i2b2 challenge covers almost all HIPAA defined categories and subcategories. The list of categories, as well as subcategories that were needed to be identified, is as follows:

- NAME (subtypes: PATIENT, DOCTOR,USERNAME)
- PROFESSION
- LOCATION (subtypes: HOSPITAL, DEPARTMENT, ORGANIZATION, ROOM, STREET, CITY, STATE, COUNTRY, ZIP)
- AGE
- DATE
- CONTACT(subtypes: PHONE,FAX,EMAIL,URL,IPADDRESS)
- Ids(SUBTYPES: MEDICAL RECORD NUMBER, HEALTH PLAN NUMBER, SOCIAL SECURITY NUMBER, ACCOUNT NUMBER, VECHILE ID, DEVICE ID, LICENCE NUMBER, BIOMETRIC ID)

In this work, the aim was to identify four different PHI subtypes; PATIENT, DOCTOR, HOSPITAL name and DATE from the above-defined categories. The experiment was performed with only the training dataset.

Evaluation

The system was evaluated by using 10 fold cross-validation on the training data. For evaluating the classifier's performance, we have used the most popular evaluation metrics; recall, precision, and F-Measure. Recall also known as sensitivity is defined as the ratio of total number retrieved relevant documents to the number of relevant records. Precision represents the positive prediction and is the ratio of retrieved relevant records to the total number of documents.

Recall and precision are represented as follows

$$recall = \frac{Total\ number\ of\ correctly\ predicted\ NE}{Total\ number\ of\ actual\ NE\ present\ in\ the\ gold\ data} \tag{2}$$

$$precision = \frac{Total\ number\ of\ correctly\ retrieved\ NE}{Total\ number\ of\ NE\ retrieved\ by\ the\ system} \tag{3}$$

F-Measure considers both the precision and recall. It is defined as the harmonic mean of recall and precision. We defined F-Measure value as follows:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$f_{\alpha} = \frac{(1+\alpha^2)(recall * precision)}{\alpha^2 * precision + recall} \tag{4}$$

Here we set the value of α=1 .

Experimental Results

Table 3. reports the extensive results of 10-fold cross validation: Recall, Precision and F-score values for individual entities. The system was developed using the handcrafted feature set trained on CRF. It was an interesting observation that with the set of whole features learned using CRF, the system performs quite well. The system was evaluated using 5,6,7,8,9,10 fold cross validation. Here, the performance of the proposed system with10 fold cross validation is reported.

Table 3 shows that overall system has achieved good recall values as compared to precision values. But there are some interesting observations regarding the recognition of individual entities. If the comparison is done on the performance among 3 important PHI terms; Hospital name, Doctor Name & Patient name then Doctor class is having the highest precision and recall values of 99.70% and 93.80% respectively. Precision and recall values for Date are quite good, but the performance can be further improved by introducing some more date based patterns. The recall value for the hospital is low as compared to all the PHI terms introduced here; the one possible reason behind this is to use abbreviation instead of full hospital names.

The report shows that the system is accurate in predicting the PHI terms with the high degree of precision. The detailed error analysis is performed on the obtained results.

Table 4 shows different error cases. Here, X or Y represents any class like PATIENT, DOCTOR, HOSPITAL, and DATE. “B_X” or “B_Y” represents the beginning of the token, “I_X” or “I_Y” represents the intermediate of the token and “O” denotes others.

Table 5 mentions about case 1: missed entity error. Here the system was unable to identify the entity. In the following e.g. ‘DO’ which in actual represents the Doctor name was predicted as others. The reason for the error is the short words or the abbreviated words. ‘DO’ is the abbreviated form of doctor name ‘DO D LITTLE’ which is missed by the system.

Table 6 presents the case 2: actual entity error. Here the entity was predicted but was not able to classify it in the proper class. For example in the following sentence, ‘FOR’ was captured as the B-DOCTOR while it should be others.

Table 3. Results of 10-fold cross validation: Recall, precision & F-score values for individual entities

Class	Recall	Precision	F-Score
Overall	0.9569	0.9931	0.9746
Hospital	0.9056	0.9769	0.9399
Doctor	0.9380	0.9970	0.9666
Patient	0.9355	0.9831	0.9587
Date	0.9905	0.9984	0.9944

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Different error cases

Case	Actual Labels	Classified Labels
MISSED ENTITY	B_X I_X	O
ACTUAL ENTITY	B_X I_X	B_Y I_Y
WRONG BOUNDARY	B_X I_X	B_X O
OVER PREDICT	O	B_X I_X

Table 5. CASE 1: Missed entity error

Words	DR.	DO	2	.	DO	D	LITTLE	,	M.D	.
Actual	O	B-DOCTOR	O	O	B-DOCTOR	I-DOCTOR	I-DOCTOR	O	O	O
Predicted	O	O	O	O	B-DOCTOR	I-DOCTOR	I-DOCTOR	O	O	O

Table 67. CASE 2: Actual entity error

Words	Dictating	For	DO	LITTLE	,	M.D	.
Actual	O	O	B-DOCTOR	I-DOCTOR	O	O	O
Predicted	O	B-DOCTOR	I-DOCTOR	I-DOCTOR	O	O	O

Table 7 represents the case 3: wrong boundary error. This case represents the boundary detection problem, where the system was able to capture the token properly but was not able to properly identify the boundary of the token. In the below example 'LITTLE' was not identified as the boundary of the system.

Table 8 represents the case 4: over predicted error. Last error occurred due to the over prediction where non-entity was categorized as the entity. For example here 'DO LITTLE' was categorized as the entity and was tagged by the DOCTOR. It might be because our system is unable to properly identify the context.

Table 7. CASE 3: Wrong boundary error

Words	HE	WILL	FOLLOW-UP	WITH	DR.	DO	LITTLE
Actual	O	O	O	O	O	B-DOCTOR	I-DOCTOR
Predicted	O	O	O	O	O	B-DOCTOR	O

Table 8. CASE 4: Over predict error

Words	DO	LITTLE	THORACOSCOPIC	EXAM	BX	DRAINAGE	.
Actual	O	O	O	O	O	O	O
Predicted	B-DOCTOR	I-DOCTOR	O	O	O	O	O

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

From the confusion matrix presented in Table 3, it was identified that mainly case 1 type of the error was more dominated.

- DOCTOR was tagged as O where there are the following forms: < D1 > M.D., < D2 > M.D., < D3> M.D.
- Misclassification has occurred when Hospital name is written in short for example CMC, FIH, OLH, HGMC etc.
- Most of the misclassification errors occurred when the patient name was preceded by the Mr., Mrs., Ms.
- Dates of the forms like 11th of April, January 25, Veterans Day were not captured properly.

CONCLUSION

In this paper, a clinical de-identification system is developed that aims to identify the Protected Health Information and to classify them into the predefined categories present in the clinical records. The proposed approach for de-identification is purely based on a popular machine learning technique. A variety of domain independent features, semantic, lexical features were designed for the system to be domain adaptable. Various others features extracted from localization dictionaries were combined in order to further enhance the performance of the system. The proposed system achieves encouraging performance. The system obtained the overall F-Measure of 97.46%, in identifying PATIENT, HOSPITAL, DOCTOR names and DATE.

However, a lot of issues are yet to be resolved. One such issue is to resolve abbreviation problem which is identified as a non-PHI term. In addition, it would be interesting to observe the performance of the proposed system on other type of PHI mentions.

FUTURE WORK

In future, this work can be extended using deep neural network architecture to identify the PHI terms from the medical records. Another major challenge is to deal ambiguity and abbreviation problem to properly distinguish PHI terms with non-PHI terms which can be another future work.

REFERENCES

Aramaki, E., Imai, T., Miyo, K., & Ohe, K. (2006, November). Automatic deidentification by using sentence features and label consistency. In *i2b2 Workshop on Challenges in Natural Language Processing for Clinical Data* (pp. 10-11).

Aronson, A. R. (2001). Effective mapping of biomedical text to the UMLS Metathesaurus: the MetaMap program. In *Proceedings of the AMIA Symposium* (p.17). American Medical Informatics Association.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Beck, T., Gollapudi, S., Brunak, S., Graf, N., Lemke, H. U., Dash, D., ... Brookes, A. J. (2012). Knowledge engineering for health: A new discipline required to bridge the ICT gap between research and healthcare. *Human Mutation*, *33*(5), 797–802. doi:10.1002/humu.22066 PMID:22392843

Beckwith, B. A., Mahaadevan, R., Balis, U. J., & Kuo, F. (2006). Development and evaluation of an open source software tool for deidentification of pathology reports. *BMC Medical Informatics and Decision Making*, *6*(1), 1. doi:10.1186/1472-6947-6-12 PMID:16515714

Berner, E. S., Detmer, D. E., & Simborg, D. (2005). Will the wave finally break? A brief view of the adoption of electronic medical records in the United States. *Journal of the American Medical Informatics Association*, *12*(1), 3–7. doi:10.1197/jamia.M1664 PMID:15492029

BioASQ. (n.d.). *BioASQ project, Data*. Retrieved January 07, 2015, from http://www.bioasq.org/participate/data

Chinchor. (2001, January 12). *Named Entity Scores– English, in: Message Understanding Conference Proceedings.* Retrieved January 06,2015, from http://www.itl.nist.gov/iaui/894.02/related_projects/muc/proceedings/ne_english_score_report.html

Coorevits, P., Sundgren, M., Klein, G. O., Bahr, A., Claerhout, B., Daniel, C., ... De. (2013). Electronic health records: New opportunities for clinical research. *The Association for the Publication of the Journal of Internal Medicine*, *274*(6), 547–560. doi:10.1111/joim.12119 PMID:23952476

Dehghan, A., Kovacevic, A., Karystianis, G., Keane, J. A., & Nenadic, G. (2015). Combining knowledge- and data-driven methods for de-identification of clinical narratives. *Journal of Biomedical Informatics*, *58*, S53–S59. doi:10.1016/j.jbi.2015.06.029 PMID:26210359

Deleger, L., Molnar, K., Savova, G., Xia, F., Lingren, T., Li, Q., ... Solti, I. (2013). Large-scale evaluation of automated clinical note de-identification and its impact on information extraction. *Journal of the American Medical Informatics Association*, *20*(1), 84–94. doi:10.1136/amiajnl-2012-001012 PMID:22859645

Friedman, C., Shagina, L., Socratous, S. A., & Zeng, X. (1996). A WEB-based version of MedLEE: A medical language extraction and encoding system. In *Proceedings of the AMIA Annual Fall Symposium* (p. 938). American Medical Informatics Association.

Goeuriot, L., Jones, G.J., Kelly, L., Leveling, J., Hanbury, A., Müller, H., Salantera, S., Suominen, H. and Zuccon, G. (2013). ShARe/CLEF eHealth Evaluation Lab 2013, Task 3: Information retrieval to address patients' questions when reading clinical reports. *CLEF 2013 Online Working Notes, 8138*.

Hara, K. (2006, November). Applying a SVM based Chunker and a text classifier to the deid challenge. In *i2b2 Workshop on challenges in natural language processing for clinical data* (pp. 10-11).

He, B., Guan, Y., Cheng, J., Cen, K., & Hua, W. (2015). CRFs based de-identification of medical records. *Journal of Biomedical Informatics*, 58. PMID:26315662

Health Insurance Portability and Accountability Act. P.L. 104e191, 42 U.S.C. 1996.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hirschman, L., Yeh, A., Blaschke, C., & Valencia, A. (2005). Overview of BioCreAtIvE: Critical assessment of information extraction for biology. *BMC Bioinformatics*, *6*(Suppl 1), S1. doi:10.1186/1471-2105-6-S1-S1 PMID:15960821

Kim, J. D., Ohta, T., Pyysalo, S., Kano, Y., & Tsujii, J. I. (2009, June). Overview of BioNLP'09 shared task on event extraction. In *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing: Shared Task* (pp. 1-9). Association for Computational Linguistics. 10.3115/1572340.1572342

Lafferty, J., McCallum, A., & Pereira, F. C. (2001). *Conditional random fields: Probabilistic models for segmenting and labeling sequence data*. Academic Press.

Liu, H., Lussier, Y. A., & Friedman, C. (2001). Disambiguating ambiguous biomedical terms in biomedical narrative text: An unsupervised method. *Journal of Biomedical Informatics*, *34*(4), 249–261. doi:10.1006/jbin.2001.1023 PMID:11977807

Liu, Z., Chen, Y., Tang, B., Wang, X., Chen, Q., Li, H., ... Zhu, S. (2015). Automatic de-identification of electronic medical records using token-level and character-level conditional random fields. *Journal of Biomedical Informatics*, 58. PMID:26122526

McCallum, A., Freitag, D., & Pereira, F. C. (2000, June). *Maximum Entropy Markov Models for Information Extraction and Segmentation* (Vol. 17). ICML.

Meystre, S. M., Friedlin, F. J., South, B. R., Shen, S., & Samore, M. H. (2010). Automatic de-identification of textual documents in the electronic health record: A review of recent research. *BMC Med Res Methodol BMC Medical Research Methodology*, *10*(1), 70. doi:10.1186/1471-2288-10-70 PMID:20678228

Nédellec, C., Bossy, R., Kim, J. D., Kim, J. J., Ohta, T., Pyysalo, S., & Zweigenbaum, P. (2013, August). Overview of BioNLP shared task 2013. In *Proceedings of the BioNLP Shared Task 2013 Workshop* (pp. 1-7).

Pyysalo, S., Ohta, T., Rak, R., Sullivan, D., Mao, C., Wang, C., ... Ananiadou, S. (2012). Overview of the ID, EPI and REL tasks of BioNLP Shared Task 2011. *BMC Bioinformatics*, *13*(11), 1. PMID:22759456

Savova, G. K., Masanz, J. J., Ogren, P. V., Zheng, J., Sohn, S., Kipper-Schuler, K. C., & Chute, C. G. (2010). Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): Architecture, component evaluation and applications. *Journal of the American Medical Informatics Association*, *17*(5), 507–513. doi:10.1136/jamia.2009.001560 PMID:20819853

Stubbs, A., Kotfila, C., & Uzuner, Ö. (2015). Automated systems for the de-identification of longitudinal clinical narratives: Overview of 2014 i2b2/UTHealth shared task Track 1. *Journal of Biomedical Informatics*, *58*, S11–S19. doi:10.1016/j.jbi.2015.06.007 PMID:26225918

Sundheim, B. M. (1996, May). Overview of results of the MUC-6 evaluation. In *Proceedings of a workshop on held at Vienna*, *Virginia* (pp. 423-442). Association for Computational Linguistics. 10.3115/1119018.1119073

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Szarvas, G., Farkas, R., & Busa-Fekete, R. (2007). State-of-the-art anonymization of medical records using an iterative machine learning framework. *Journal of the American Medical Informatics Association*, *14*(5), 574–580. doi:10.1197/jamia.M2441 PMID:17823086

Torii, M., Fan, J. W., Yang, W. L., Lee, T., Wiley, M. T., Zisook, D., & Huang, Y. (2014, November). De-identification and risk factor detection in medical records. Seventh i2b2 Shared Task and Workshop: Challenges in Natural Language Processing for Clinical Data.

Uzuner, Ö., Solti, I., & Cadag, E. (2010). Extracting medication information from clinical text. *Journal of the American Medical Informatics Association*, *17*(5), 514–518. doi:10.1136/jamia.2010.003947 PMID:20819854

Uzuner, Ö., South, B. R., Shen, S., & DuVall, S. L. (2011). 2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text. *Journal of the American Medical Informatics Association*, *18*(5), 552–556. doi:10.1136/amiajnl-2011-000203 PMID:21685143

Wellner, B., Huyck, M., Mardis, S., Aberdeen, J., Morgan, A., Peshkin, L., ... Hirschman, L. (2007). Rapidly Retargetable Approaches to De-identification in Medical Records. *Journal of the American Medical Informatics Association*, *14*(5), 564–573. doi:10.1197/jamia.M2435 PMID:17600096

Wolf, M. S., & Bennett, C. L. (2006). Local perspective of the impact of the HIPAA privacy rule on research. *Cancer*, *106*(2), 474–479. doi:10.1002/cncr.21599 PMID:16342254

Xu, H., Stenner, S. P., Doan, S., Johnson, K. B., Waitman, L. R., & Denny, J. C. (2010). MedEx: A medication information extraction system for clinical narratives. *Journal of the American Medical Informatics Association*, *17*(1), 19–24. doi:10.1197/jamia.M3378 PMID:20064797

Xue, H., Chen, S., & Yang, Q. (2008). Structural support vector machine. *Advances in Neural Networks-ISNN*, *2008*, 501–511.

Yang, H., & Garibaldi, J. M. (2015). Automatic detection of protected health information from clinic narratives. *Journal of Biomedical Informatics*, 58. PMID:26231070

KEY TERMS AND DEFINITIONS

Classification: Classification in the machine learning is defined as the supervised learning technique where problem is to identify the class of the new observation with the already developed observations through the labeled data.

Conditional Random Field (CRF): An undirected probabilistic graphic model categorized under statistical modeling method used for structured prediction in machine learning and pattern recognition.

Health Insurance Portability and Accountability Act (HIPAA): An act formulated in 1996 for the secure transmission and usability of the private health information electronically.

Machine Learning: A field of computer science that exploits the development of the algorithms for making a prediction on the data on the basis of its learning.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Medical Natural Language Processing: A domain of natural language processing that focuses on text mining of the medical records like clinical texts.

Named Entity Recognition (NER): An entity extraction task which aims to identify and retrieve the text carrying relevant information into some predefined categories.

Natural Language Processing (NLP): A domain of computer science, computational linguistics and artificial intelligence that focuses on computers understand to human languages or natural language.

Support Vector Machine (SVM): A supervised learning model used in the classification that learns from the data by developing a model that maximizes the error margin.

This research was previously published in the Handbook of Research on Applied Cybernetics and Systems Science edited by Snehanshu Saha, Anand Narasimhamurthy, Abhyuday Mandal, Sarasvathi V, and Shivappa Sangam; pages 234-253; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 49
Neural Network Model for Semantic Analysis of Sanskrit Text

Smita Selot
SSTC, India

Neeta Tripathi
SSTC, India

A. S. Zadgaonkar
CV Raman University, India

ABSTRACT

Semantic analysis is the process of extracting meaning of the sentence, from a given language. From the perspective of computer processing, challenge lies in making computer understand the meaning of the given sentence. Understandability depends upon the grammar, syntactic and semantic representation of the language and methods employed for extracting these parameters. Semantics interpretation methods of natural language varies from language to language, as grammatical structure and morphological representation of one language may be different from another. One ancient Indian language, Sanskrit, has its own unique way of embedding syntactic information within words of relevance in a sentence. Sanskrit grammar is defined in 4000 rules by PaninI reveals the mechanism of adding suffixes to words according to its use in sentence. Through this article, a method of extracting meaningful information through suffixes and classifying the word into a defined semantic category is presented. The application of NN-based classification has improved the processing of text.

INTRODUCTION

Semantic, in basic form is the meaning of the sentence. In computer driven world of automation, it has become necessary for machine to understand the meaning of the given text for applications like automatic answer evaluation, summary generation, translation system etc. In linguistics, semantic analysis is the

DOI: 10.4018/978-1-7998-0951-7.ch049

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

process of relating syntactic structures, from words and phrases of a sentence to their language independent meaning. Given a sentence, one way to perform semantic analysis is to identify the relation of the words with action entity of the sentence. For example, *Rohit ate ice cream*, agent of action is Rohit, object on which action is performed is ice cream. This type of association creates predicate-arguments relation between the verb and its constituent. This association is achieved in Sanskrit language through *kArakA* analysis. Understanding of the language is guided by its semantic interpretation. Semantic analysis in Sanskrit language is guided by six basic semantic roles given by *pAninI* as *kAraka* values.

pAninI, an ancient Sanskrit grammarian, mentioned nearly 4000 rules called *sutra* in book called *asthadhyAyi;* meaning eight chapters. These rules describe transformational grammar, which transforms root word to number of dictionary words by adding proper suffix, prefix or both, to the root word. Suffix to be added depends on the category, gender, number of the word. Structured tables containing suffix are maintained for the purpose. These declension tables are designed in such a way that their position in the table are defined with respect to number, gender and *karka* value. Similar ending words follow the same declension, for example *rAma* is *a*-ending root word and words generated using *a*-ending declension table are r*AmH, rAmau rAmAH* by appending *H, au* and *AH* to *rAma*, respectively. Suffix based information of the word reveals not only syntactic but drives a way to find semantic based relation of words with verb using *kAraka* theory.

Next section describes Sanskrit language and *kAraka* theory, section three states the problem definition, followed by NN model for semantic analysis. Features extracted from corpus of pre-annotated text are supplied as input to system with objective of making system learn six *kAraka* defined by *pAninI.* This paper presents the concept of Neural Network, work done in the field of NN and Natural Language Processing, algorithm, annotated corpus and results obtained.

SANSKRIT AND *kAraka* THEORY

Sanskrit language, with well-defined grammatical and morphological structure, not only presents relation of suffix-affix with the word, but also provides syntactic and semantic information the of words in a sentence. Due to its rich inflectional morphological structure; it is predicted to be suitable for computer processing. Work at NASA on Sanskrit language reported that triplets (role of the word, word, action) generated from this language are equivalent to semantic net representation (Briggs 1995).

Sanskrit grammar is developed by three main individuals over the years – *pAninI*, *Katyanan* and *Patanjali*. *pAninI* was first one to identify nearly 4000 rules to define the language (Kak, 1987). Sanskrit is one of the 22 official languages of India and standardized dialect of old Indo-Aryan community. It is a rule-based language having flexibility and precision, both. It is an order free language – changing the position of the word, does not change the meaning of the sentence. Processing techniques for order free languages cannot be derived from rigid order language like English. *pAninI's* idea of semantic interpretation was based on identification of relations of words with action in a sentence. as agent, object etc. This was called as *kAraka* theory. In Sanskrit, this relationship is developed by adding specific, pre-set syllables, known as case-endings or *vibhakti* to the basic noun form.Natural language processing concept in *pAnini* way describes the processing of language

Sanskrit is highly inflectional language, where words are generated from the root word. Structure of the word is given as <word>=<prefix><root_word><suffix>. In pAninI system finite set of rules are enough to generate infinite sentences (Kak,1987). Extraction of prefix and suffix by morphologi-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

cal analyzer accurately, help in syntactic and semantic analysis of the sentence. Case endings are the driving force for semantic interpretation as it can change the meaning of the sentence as stated in the example below:

1. *RamH bAlakam grAm**At** yaati.* (Rama **brings** the boy **from** the village.)
2. *RamH bAlakam grAm**am** yaati*. (Rama **takes** the boy **to** the village.)

This paper focuses on *kAraka* based semantic analysis of text under six *kAraka* as described in Table 1. Identification of *kAraka* values of words forms the basis of fundamental semantic interpretation (Bharati et. al 2008; Dwivedi 2002). Ancient old *kAraka* theory rules for external sandhi are modeled with finite state machine (Hyman 2009) and potential of *pAninI* grammar exceeds power of finite state machines. (Vaidya et. al. 2005). *kAraka* roles are equivalent to case-based semantics in event-driven situations, where each word or phrase is related to action. This relation is labeled as cases like agent, object, location (Kak 1987; Hellwig 2007). Under *pAnini* sytem of *kAraka* they are identified as *karta, karma, karan, sampradan, apadana* and *adhikaran* which are interpreted as agent, object, instrument, cause, source of departure and location respectively as presented in Table 1 (Kiparsky 2002) various algorithms over the years have evolved for semantic classification of words, text and documents. Rule based system are also one of the conventional and good means of extracting information. Order of application of rules from NLP perspective is optimized and implemented on subantaprakaraṇa of Siddhāntakaumudī. This system works well as Sanskrit is rules-based language. There is no order of rules defined in theory; but for practical implementation an optimized order is defined by the authors. (Patel & Katuri, 2015) One of the major problems is designing and incorporating large volume of rules and to handle the overlapping or conflicting cases (Selot et al., 2010). Development of knowledge representation for Sanskrit based on stated theory is also experimented. (Selot et al., 2007) Dependency amongst the constituent of the words in a sentence is also a means of representing semantic relation. (Pedersen et al., 2004)

A stochastic Part of Speech (POS) tagger for unprocessed Sanskrit text using Hidden Markov model was designed and analyzed. Parameters for the system were obtained from annotated corpus of about 1,500,000 words. System analyzed unannotated Sanskrit text by repeated application of stochastic model (Hellwig, 2007). Mathematical models such as Hidden Markov, Maximum Entropy and Support Vector machines are extensively used in natural language processing (Timothy, 2005). With evolution of statistical methods, N-gram models became popular in POS tagging. Trigram's and Tags (TnT) for POS tagging has proved to be an efficient tool for Persian language. (Tasharofi, 2008). Microsoft Research India (MSRI) introduced hierarchical Indic language POS tagset (IL-POSTS) for tagging Indian lan-

Table 1. Six kAraka in pAnanian model

S.No	Case	kAraka	Vibhakti	Meaning
1	Nominative	*Karta*	*prathamA*	Agent
2	Accusavive	*Karma*	*dviteeyA*	Object
3	Instrumental	*karNa*	*Trieetiya*	Instrument
4	Dative	*Sampradan*	*Chaaurthi*	Recipient
5	Ablative	*apAdAn*	*panchami*	Departure
6	Locative	*adhikaraN*	*saptami*	Place

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

guages. An experiment on the use of the IL-POSTS for annotating Sanskrit corpus is conducted. (Jha et.al, 2010) Though statistical model are quite efficient and popular for language processing; neural network is equally efficient with reduced processing time in applications like text categorization (Babu, 2010).

With availability of natural language processing tools like nlptk; an open source platform; computations and analysis in the domain has taken a step forward in research areas. However, for Indian languages especially Sanskrit; availability of large volume of corpus and processing is still a challenge (Joon et al., 2017)

Use of Neural Network for complex task like NLP is rarely found as most of the work done using stochastic model reports an accuracy of more than 90%. Stochastic model requires large amount of pre-annotated corpus to calculate various probability values (Shan 2004). Lack of availability of such large size corpus for regional language hinders the development of NLP systems with high accuracy rate. Neural Network appears to be better choice for classifying the Sanskrit words in semantic classes as it has paradigm-based representation and had reported better result under sparse data.

PROBLEM DEFINITION

Processing of Sanskrit text is challenging one for two main reasons; firstly, it is order free and secondly words coagulate at junction to form a new word. Output oriented training of any language requires large tagged corpus for better results. Feature extraction from individual words is possible with such tagged data. Scarcity of tagged data in Sanskrit language restricts the number of algorithm that can be implemented for semantic extraction Hence, main objective is to identify the semantic relation of the word with action entity of the sentence under stated constraint environment. Given a word with its suffix, system identifies the *kAraka* value of the word from the set of predefined six *kAraka* roles. System is first trained for semantic classes using manually pre-annotated corpus. Features are extracted from corpus; set of input/output are stored and used for training. System's performance is then evaluated on new set tagged sentence. Manually annotated text was developed for experiment purpose. Challenge was mainly associated with conflicting cases. When a suffix maps to multiple semantic classes; means to resolve them was needed. Proposed system handled such conflict by maintaining context-based features as most of the conflicts are resolved by occurrence of a particular word or its constituents in context window. Neural Network is used to train the system as it works well in sparse data. Effort in development of this algorithm has an objective of accomplishing learning of semantic relation through a technical approach and analysing its inbuilt scientific way of its representation. Finally, performance of the system is discussed is reported.

FEATURE SELECTION

Annotated corpus describing syntactic and *kAraka* categories of words is manually generated. It includes the suffix-affix information, category, *vibhakti* in nouns and tense and mood in verbs. Set of features generally used to train the system are morphological feature, corpus-based feature, context-based feature and lexicon-based feature.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Morphological Features

It includes the suffix and prefixes of words in the context. Prefix or suffix is a sequence of first or last few characters of a word, which does not possess linguistically meaning. Use of prefix and suffix information works well for highly inflected languages.

Corpus Based Features

Corpus based features play a crucial role in tagging of words. Each word is to be tagged with one of the six *pAninI* class. Not all the words will fall under six tags; some words exhibit relation with subsequent words, they are categorized as relation class. For words which are neither in six tags nor in relation tag are assigned category other. To finalize a tag for a word, its historical information in form of probability is calculated from corpus. A probability vector is associated with each suffix which gives the chances of assigning a particular *kAraka* to a word with a defined suffix.

Context Based Features

In declension tables; in some cases, particular suffix occupies more than one place, resulting in ambiguity. To resolve the *kAraka* or *vibhakti* ambiguity associated with the word, values of proceeding and successive words are taken into consideration. It is quite arguable that for order free languages, like Sanskrit and Hindi, context-based feature should not be considered. But, it has been found that occurrence of a particular word before or after the word generates important clue in resolving the *kAraka* or *vibhakti* ambiguity. For example, in the word *vanam, am* suffix occupies two different *vibhakti* position- first and second. As observed in the grammar, occurrence of words like *dhik,hA, ubhyatH, pAritH* confirms the *vibhakti* value of *am* as *dviteeyA*. (Rao 2005). Context window of various sizes are used in problems like word sense disambiguation (WSD), POS tagging, semantic tagging etc. Window size of three gives better result maintaining the balance between accuracy and processing overhead. Hence context window of size three (w_{i+1}, w_i, w_{i-1}) is taken for resolving the *vibhakti* ambiguity associated with Sanskrit language.

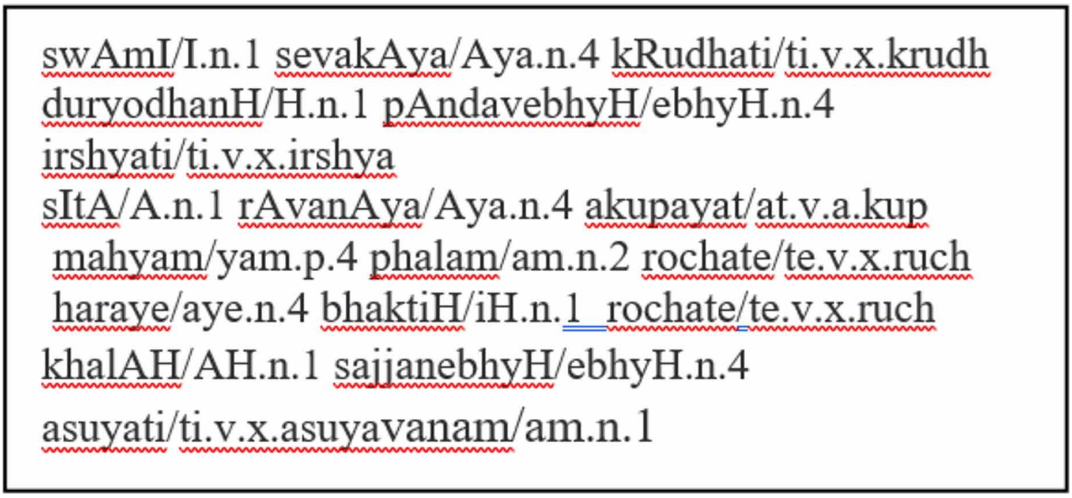
Lexicon Based Feature

Instead of maintaining a complete dictionary of words, list of verbal and nominal bases has been maintained. There is strong association of words with the verb occurring in the sentence. For example, in deciding *vibhakti* value for *am* suffix, occurrence of any one of sixteen-verbal base confirms the *vibhakti* value of *am* as second. These sixteen verbs are {*dhuH, yAch, pach, chi bru, Shas anD, rudh, pRich, ji, math, mush, ni, kriSh, hRi* }(Dwivedi 2002). Likewise, a complete database of nominal base is also maintained. Features are selected from the annotated corpus and are assigned numeric value after mapping them to numeric data. A sample tagged data used for training is shown in the Figure 1; where each word is tagged with its suffix, syntactic category and *kAraka* value. For example; first word in the Figure 1; *swAmI;* contains *I* suffix, part-of-speech is noun and *kAraka* value is 1 *prathamA*

Set of features selected are quantified and used for training network. Text coverage of suffix plays an important role in resolving the conflict.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Sample tagged data



NN MODEL FOR SEMANTIC CLASSIFICATION

Popular methods for NLP are rule based or stochastic models, where former emphasizes on rules of language and later uses large corpus for processing. Major task in NLP at syntactic and semantic level is classification of words or phrases. Neural network is popular tool for classification task which uses an extensive data driven approach for solving complex problem Capabilities of NN are exploited in classification task related to natural language, such as text categorization, POS tagging etc. It has been shown that NN based text categorization is better than N-gram approach giving 98% accuracy with speed five times faster than N-gram (Kumar 2010). A multi layered neural network with the back propagation is used for solving the problem. Input to the system is syntactic and morphological features of the previous and successive words in a sentence. Brief working of Multi layered perceptron (MLP) based NN is explained in the Figure 2.

MLP Network

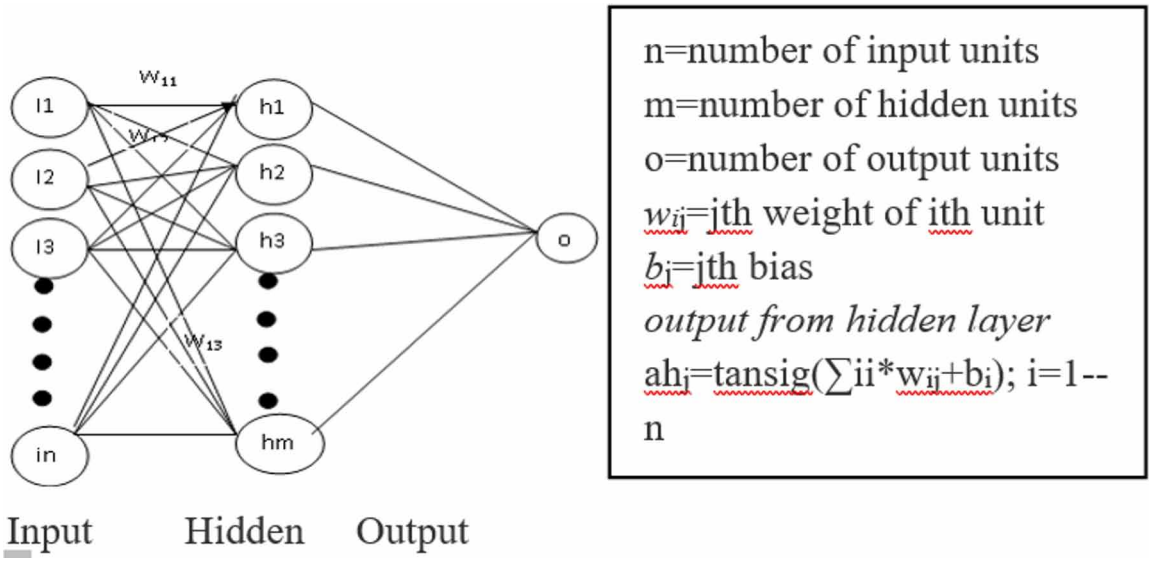
It consists of minimum of three layers input, hidden and output layer represented by symbol i, h and o as presented in Figure 2. Each layer consists of processing units termed as neurons. Fully connected network uses sigmoid activation function which limit the output from hidden unit between -1 and 1. MLP projects the data to be classified into different spaces, where neuron of output participates in the separation of classes. Final outcome is obtained by applying purelin function as output is in range 1 to 6 depending upon the *kAraka* value.

Back Propagation Learning Algorithm

Learning algorithms are commonly used to make computer system learn a concept for particular applications. Neural Network is one of the learning environments simulating human brain. It consists of three layers namely input, hidden and output layer. Network is trained with set of defined input-output

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Three-layer MLP network



data set. Performance of the system is tested by comparing actual output with desired one. Back propagation algorithm calculates error and sends it back to neurons for weight adjustment. After each pass error is reduced and system is tuned for better results. System works in two phases, in forward phase weights are multiplied to respective inputs and bias is added to it. Result is fed as input to hidden layer; after squashing the values. Data obtained from hidden is fed to output layer; to limit the range sigmoid function is applied. Output obtained is compared with actual output and error is measured as difference of the two. In the backward phase, error is transmitted back; and weights are modified as function of error. In the forward phase, weights of the directed links remain unchanged at each processing unit of the hidden layer. For n input values to the network; weighted sum is obtained and sigmod function is applied to this weighted sum as given in Equations (1) and (2):

$$u_j = \sum_{1}^{n} w_{ij}^{h} pj + b \tag{1}$$

$$h_j = 1 \Big/ 1 - e^{u_j} \tag{2}$$

where h_j is the output from the jth hidden unit. At each processing unit j of the output layer; weighted outputs from the all neuron of hidden layer are added using Equation 3 and resulting value is passed through a sigmoid function as given in Equation 4:

$$v_j = \sum_{1}^{n} w_{ij}^{y} h_j + b \tag{3}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$y_j = 1 \Big/ 1 - e^{v_j} \tag{4}$$

where y_j is the final output. If t_0 is the desired output and y_0 is the actual output, in the backward pass, for every output neuron, local gradient is calculated as per Equations 5:

$$g_o = E_o y_o (1 - y_o) \tag{5}$$

where $E_o = t_o - y_o$ (6)

likewise, for each hidden neuron, local gradient is calculated as:

$$g_h = E_h h_i (1 - h_i) \tag{7}$$

where $E_h = \sum_1^n g_o w_{ij}^y$ (8)

To speed up the training process, momentum term α and learning rate η, is added to the weight update formulae in hidden and output layers as given in Equations 9 and 10:

$$\left.\begin{aligned} \Delta w_{ij}^h (new) &= \eta g_h h_i + \alpha \Delta w_{ij}^h (old) \\ w_{ij}^h (new) &= w_{ij}^h (old) + \Delta w_{ij}^h (new) \end{aligned}\right\} \tag{9}$$

$$\left.\begin{aligned} \Delta w_{ij}^y (new) &= \eta g_O y_i + \alpha \Delta w_{ij}^y (old) \\ w_{ij}^y (new) &= w_{ij}^y (old) + \Delta w_{ij}^y (new) \end{aligned}\right\} \tag{10}$$

Outline of Algorithm

Step 1: Weights are initialized to small random variables
Step 2: For each input-output training pair
Input vector=<p_1, p_2… p_n>
Weight for first neuron = <w_{11}, w_{21}, w_{31}… w_{n1}>
Output=<y_0>
Step 3: Transmit input values to neurons of hidden layer.

$$u_j = \sum_1^n w_{ij}^h pj + b$$

Step 4: Apply activation function

$$h_j = 1 \Big/ 1 - e^{u_j}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Step 5: Each output unit receives sum of weighted input and applies an activation function to it

$$v_j = \sum_{1}^{n} w_{ij}^{y} h_j + b$$

$$y_j = 1 \Big/ 1 - e^{v_j}$$

Step 6: Each output unit, difference between target and desired value is calculated as an error.

$E_{rr} = t_o - y_o$

Step 7: At each hidden unit the error is calculated as error from the layer above it.
Step 8: Weights and biases are updated at input and hidden layer
Step 9: Stopping condition is tested
Step 10: Repeat steps 3 to 9 till stopping condition is achieved

Processing of text using NN requires a defined set of input data in numerical form. Numerically encoded data is generated by capturing various features of the input. For language processing applications, these features include – morphological, corpus-based, context based, lexicon-based features. Next section explains the generation of encoded data.

Encoding of the Input Data

Features vector, input to the NN, is generated by converting different features to their numeric form as described in the Table 2.

Algorithm for encoding input values reads the pre-annotated text and converts the data into feature vector.

System is trained using multilayer neural networks with *squashing* functions in the hidden layer. These functions compress an infinite input range of input values to a finite output range. In sigmoid functions, slope approaches zero as input gets larger. This causes problem in steepest descent, as gradient has small value, changes in weight and biases are very small, though they are far from their optimum values. Re-

Table 2. Sample set of encoded values

Class\number	1	2	3	4	5	6
category	Pronoun	Noun	Verb	Avaya	Adverb	
Prefix(verb)	Up	Anu	Adi	A	Abhini	
Verb(root)	Gachch	Pach	Kar	Bhu	Anay	as
Vibhakti	Agent	Object	instrument	cause	From	Relation
Suffix(verb)	ti	si	At	ami	Anti	AmH
Suffix(noun)	No suffix	Am	H	Aya	ebhyH	Asya

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

silient back propagation removes these problems of magnitude of partial derivative. It uses sign of the derivative to determine the direction of weight update. Amount of weight change is determined by the separate update value and is incremented by an amount delta_inc whenever the derivative change sign from previous iteration. Weight is unchanged if derivative is zero; change is decreased for oscillating weights and increased if weight continues to change in same direction.

In NN based semantic classification, resilient back propagation algorithm is implemented.

Input Data for the BPN

Feature vector, main input in network, is generated by extracting parameters from annotated corpus. These parameters are classified under different category as given in the Table 3. Suffixes and prefixes fall under morphological features; are crucial part of feature vector Apart from this, context based and syntactic class also form a part of the vector. Context based features are significant as they help in resolving the conflict. Occurrence of multiple matched suffixes in a particular *vibhakti* group is governed by occurrence of special words at preceding and successive positions.

Data set of pre-annotated sentences is created with tagged words revealing the syntactic feature, morphological and semantic features of the words such as suffix, category and *vibhakti* value. Annotated corpora contain sentences taken from various Sanskrit text books. An example of dataset and its generated encoded form is presented in Figure 3. Algorithm for generating the encoded data after reading each sentence of a pre-annotated data is presented below:

Step 1: Identify the word of each sentence
Step 2: Read the features tagged with word
Step 3: Identify the type of feature
Step 4: Convert them into the numeric value as described in Table 2.
Step 5: Identify the *vibhakti* value and store it as target value
Step 6: Store the information in the file

Output is *kAraka* value of the word stored in a fil. Each case or *kArka* is defined as follows: 1 for *kartA*,2 for *karma,*3 for *karan,*4 for *sampradan,*5 for *apAdAn* and 7 for *adhikaran.* Number 6 is used for *sambandh kAraka* which is not included in the *pAninI* class.

The encoded data is used for training network with back propagation using 200 sentences. Network is trained using resilient back propagation algorithm for data set and the result is shown in the Figure 4.

BPN data set for encoded input/output values were used and system performance is calculated. As shown in the figure system was trained in less than 400 epochs.

Table 3. Feature selected

Parameter	Feature type
Prefix and suffix of noun and verb	Morphological
Word category	Feature based on syntax
Verb root	Lexical feature
Successive and previous words	Context based feature
Probability occurrence of sufix	Feature based on corpus

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. (a) Anotated data (b) encoded values

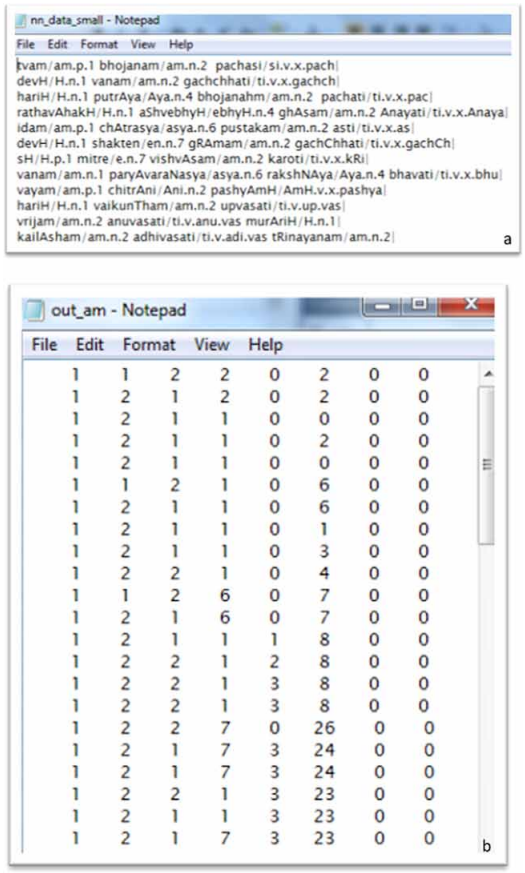
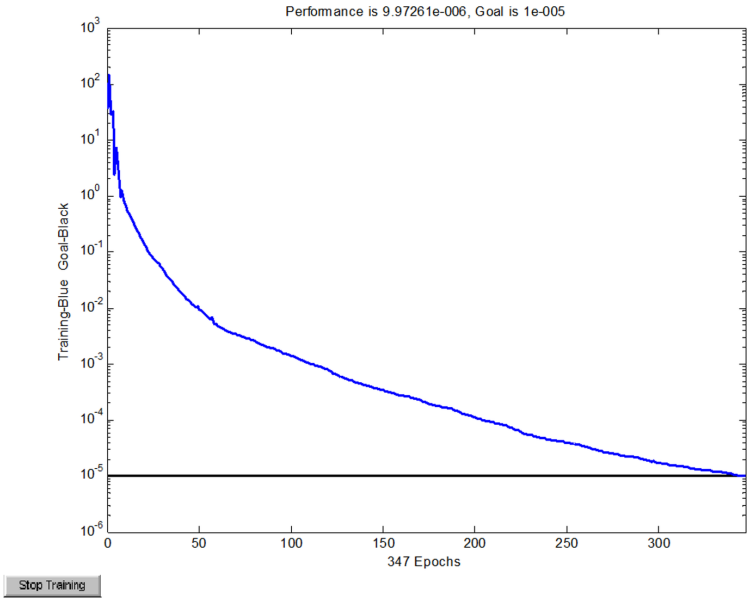


Figure 4. Training of the NN for semantic classification



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

Experiments for *kAraka* based semantic classification using neural networks were conducted as per algorithms discussed in previous section. Tagged data of 200 sentences were created as shown in the Figure 1 Out of 200 sentences, 80% of the data was used in training the network with back propagation and 20% of the data was used for testing. Back propagation algorithm with gradient descent and resilient learning is used to train the three-layered network where system was trained in 500-800 epochs. Performance of resilient BP algorithm has shown better results over gradient descent. Data was tested with different learning rate as shown in the Table 4 where n = number of epochs and lr = learning rate

Data at the serial number 2 shows the best performance with time, otherwise as the number of neuron increases, training time increases. Time is also dependent on the learning rate, higher the learning rate, faster the system and vice versa. Trained NN was tested on the 20 sentences. Out of which the identification of each class is shown in the Table 5. Using resilient BPN algorithm, best output is obtained with 200, 250 300 neurons and 0.01 learning rate. Using this data, the system was further tested, and the output was classified for the various semantic classes. Accuracy of the test was measured by F-measure given in Equation 11:

$$F = 2(p \times r)/(p + r) \tag{11}$$

Average performance of the system with respect to F-Score is 0.8243. Test data with 40 sentences were used for Case Frame generation out of which 22 were generated with 100% accuracy, 12 with accuracy between 50-100% and remaining with accuracy less than 50% thereby giving an overall performance of 72.5%. F_measure presents weighted score of the individual *kAraka* value and overall accuracy is measured degree of correctness achieved in each CF generation.

Statistical tools are commonly used in NLP applications like POS tagging, semantic classification, word sense disambiguation etc. They give an accuracy of 90% and above.

Table 4. Trained data with different learning rate

pAninI **class**	**f**	**n**	**c**	**p**	**r**	**F**
Karta	20	18	17	0.940	0.850	0.890
Karma	15	12	11	0.916	0.730	0.812
Karan	15	13	10	0.760	0.660	0.706
sampradAn	13	12	11	0.916	0.846	0.879
apAdAn	13	12	10	0.833	0.769	0.790
Adhikaran	12	11	10	0.909	0.833	0.869

Where p is precision; r is recall and F is F score.p and r are calculated as follows:
p = Number of accurate result /total number of returned result
r = Number of exact result / Number of results to be returned.
F score is calculated as weighted average of precision and recall using Equation (11). F-Score of each class under NN is given in Table 5.
f = frequency of occurrence of each class in corpus
n = total count of a particular class in data
c = number correctly identified by the system

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Classification of pAninI for test data

S. No	Epochs	n	Lr	Performance	Time (in sec)
1	460	200	0.01	0.009981	11.20
2	**322**	**200**	**0.03**	**0.009980**	**7.75**
3	507	200	0.05	0.009980	12.43
4	462	250	0.01	0.009980	18.62
5	454	250	0.03	0.009990	17.72
6	270	250	0.05	0.009980	10.59
7	243	300	0.05	0..009996	11.32
8	382	300	0.03	0.009995	17.66
9	247	300	0.01	0.009985	11.81

Due to inherent complexities and ambiguities in languages, Syntactic and semantic classifications become a great challenge. Collection of tagged corpus with wide coverage of all scenarios plays an important role in training NLP system. Statistical tool, supported by large set of tagged corpus, has given accuracy of 90% and above. All languages are not rich in resources. Tagged corpus rarely cover all scenarios or conflicting cases. Processing of Sanskrit, an ancient language, with high deep rooted morphological structure but with scarce resources, was a great challenge. At the same time applications of neural networks for classification purpose under sparse data had proved to be a better tool. Hence, experimental set up was built to analyze the semantic classification after training NN for these classes under *pAninI* framework. NN has proved to be an efficient tool for semantic classification of the Sanskrit words.

This work can further be enhanced for automatic feature extraction of each word. As structure of the basic word in Sanskrit contain prefix, root word and suffix; an algorithm for breaking word into its constituent and mapping it into feature set can be developed. Incorporation of deep learning in classification can be experimented for further improvement of the system.

REFERENCES

Babu, A., Suresh, K., & Pavan, P. N. V. S. (2010). Comparing neural network approach with n-gram approach for text categorization. *International Journal on Computer Science and Engineering*, *2*(1), 80–83.

Bharati, A., Kulkarni, A., & Sivaja, S. N. (2008). Use of Amarako'sa and Hindi WordNet in Building a Network of Sanskrit Words. In *6th International Conference on Natural Language Processing ICON '08*. India: Macmillan Publishers.

Bhattacharyya, P. (2010, May). Indowordnet. In *lexical resource engineering conference*, LREC Malta,

Briggs, R. (1985). Knowledge Representation in Sanskrit and Artificial Intelligence. *AI Magazine*, *6*(1), 32–39.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chaitanya, V., & Kulkarni, A. (2015). Sanskrit and Computational Linguistics. In *16th World Sanskrit Conference,* 28 June – 2 July , Silpakorn University, Bangkok.

Dwivedi, K. (2002). Prarambhik RachanAnuvAdakumaudi. VishavavidyAlaya PrakAshan. Varanasi (19th ed.).

Hellwig, O. (2011). Performance of a lexical and POS tagger for Sanskrit. In *Proceedings of fourth international Sanskrit computational linguistic symposium* (pp. 162-172).

Huet, G. (2003). Towards Computational Processing of Sanskrit. Recent advances in Natural Language Processing. In *Proceedings in International conference ICON*.

Hyman, D. M. (2009). *From pAninian sandhi to finite state calculus*. In Sanskrit Computational Linguistics. Springer-Verlag. doi:10.1007/978-3-642-00155-0_10

Jha, G. N., Gopal, M., & Mishra, D. (2009, November). Annotating Sanskrit corpus: adapting IL-POSTS. In Language and Technology Conference (pp. 371-379). Springer, Berlin, Heidelberg.

Joon, R., Gupta, V., Dua, C. & Kaur, G. (2017). Nltk based processing of sanskrit text. *International Journal of Management and Applied Science*, *3*(8).

Kak, S. C. (1987). The panian approach to natural language processing. *International Journal of Approximate Reasoning*, *1*(1), 117–130.

Kiparsky, P. (2002). On the architecture of Panini's grammar. In *International Conference on the Architecture of Grammar*.

Kumar, D., & Josan, G. (2010). Part of Speech Taggers for Morphologically Rich Indian Languages: A Survey. *International Journal of Computers and Applications*, *6*(5), 32–41. doi:10.5120/540-704

Mittal, V. (2010). Automatic Sanskrit Segmentizer Using Finite State Transducers. In *Proceedings of the ACL 2010* (pp. 85-90).

Bharati, A., Chaitanya, V., Sangal, R., & Ramakrishnamacharyulu, K. V. (1994). Natural language processing: A Paninian perspective. PHI New Delhi.

Patel, D. & Katuri, S. (2015). Prakriyāpradarśinī - an open source subanta generator Sanskrit and Computational Linguistics. In *16th World Sanskrit Conference,* 28 June – 2 July, Silpakorn University, Bangkok.

Pedersen, M., Eades, D., Amin, S. K. & Prakash, L. (2004). Relative Clauses in Hindi and Arabic: A Paninian Dependency Grammar Analysis.

Rao, B. N. (2005). Panini and Computer Science – into the future with knowledge from past.

Selot, S., & Singh, J. (2007). Knowledge representation and Information Retrieval in PAninI Grammar Framework. In *International Conference ICSCIS '07* (Vol. 2, pp. 45-51).

Selot, S., Tripathi, N., & Zadgaonkar, A. (2010). Semantic Extraction in PGF using POS tagged data for Sanskrit. i-manager's Journal on Software Engineering, 5(2).

Selot, S., Tripathi, N., & Zadgaonkar, A. S. (2009). Transition network for processing of Sanskrit text for identification of case endings. *ICFAI Journal of Computational Science*, *3*(4), 32–38.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Shan, H. & Gildea, D. (2004). *Semantic labelling by Maximum entropy model.* University of Rochester.

Tasharofi, S., Raja, F., Oroumchian, F., & Rahgozar, M. (2008). Evaluation of statistical part of speech tagging of Persian text. In *9th International symposium on signal processing and its applications*.

Timothy, J. D., Hauser, M., & Tecumseh, W. (2005). Using mathematical models of language experimentally. *Trends in Cognitive Sciences*, *9*(6), 284–289. doi:10.1016/j.tics.2005.04.011 PMID:15925807

Vaidya, A., Husain, S., Mannem, P., & Misra, D. S. (2009). A Karaka Based Annotation Scheme for English. In *International Conference on Intelligent Text Processing and Computational Linguistics, LNCS* (Vol. *5449,* pp. 41-52). Springer.

This research was previously published in the International Journal of Natural Computing Research (IJNCR), 7(1); edited by Nilanjan Dey and Vikrant Bhateja; pages 1-14; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 50
Probabilistic Nodes Combination (PNC): Formulas and Examples

Dariusz Jacek Jakóbczak
Technical University of Koszalin, Poland

ABSTRACT

The method of Probabilistic Nodes Combination (PNC) enables interpolation and modeling of two-dimensional curves using nodes combinations and different coefficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function, also inverse functions. This probabilistic view is novel approach a problem of modeling and interpolation. Computer vision and pattern recognition are interested in appropriate methods of shape representation and curve modeling. PNC method represents the possibilities of shape reconstruction and curve interpolation via the choice of nodes combination and probability distribution function for interpolated points. It seems to be quite new look at the problem of contour representation and curve modeling in artificial intelligence and computer vision. Function for γ calculations is chosen individually at each curve modeling and it is treated as probability distribution function: γ depends on initial requirements and curve specifications.

INTRODUCTION

Probabilistic modeling is still a developing branch of economic and computer sciences: operational research (for example probabilistic model-based prognosis) (Lorton, Fouladirad & Grall, 2013), decision making techniques and probabilistic modeling (Pergler & Freeman, 2008), artificial intelligence and machine learning. Different aspects of probabilistic methods are used: stochastic processes and stochastic model-based techniques, Markov processes (Cocozza-Thivent, Eymard, Mercier & Roussignol, 2006), Poisson processes, Gamma processes, Monte Carlo methods, Bayes rule, conditional probability and many probability distributions. In this chapter the goal of a probability distribution function is to describe the position of unknown

DOI: 10.4018/978-1-7998-0951-7.ch050

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

points between given interpolation nodes. Two-dimensional curve is used to represent the data points and extrapolation of the unknown values enables analyzing risk and then decision making.

The chapter clarifies the significance and novelty of the proposed method compared to existing methods (for example polynomial interpolations and Bézier curves). Previous published papers of the author were dealing with the method of Hurwitz-Radon Matrices (MHR method). Novelty of this chapter and proposed method consists in the fact that calculations are free from the family of Hurwitz-Radon Matrices. Problem statement of this chapter is: how to reconstruct (interpolate) missing points of 2D curve and how to anticipate (extrapolate) unknown values or data having the set of interpolation nodes (key points) and using the information about probabilistic distribution of unknown points. For example the simplest basic (uniform) distribution leads to the easiest interpolation – linear interpolation. Apart from probability distribution, additionally there is the second factor of proposed interpolation method: nodes combination. The simplest nodes combination is zero. Thus proposed curve modeling and extrapolation is based on two agents: probability distribution and nodes combination. First trial of probabilistic modeling in MHR version was described in (Jakóbczak, 2013). Significance of this chapter consists in generalization for MHR method: the computations are done without matrices in curve fitting and data anticipation, with clear point interpolation formula based on probability distribution function (continuous or discrete) and nodes combination. The chapter also consists of generalization for linear interpolation with different (non-uniform) probability distribution functions and nodes combinations. So this chapter answers the question: "Why and when should we use Probabilistic Nodes Combination (PNC) method in extrapolation and interpolation?".

Curve interpolation (Collins, 2003) represents one of the most important problems in mathematics and computer science: how to model the curve (Chapra, 2012) via discrete set of two-dimensional points (Ralston & Rabinowitz, 2001)? Also the matter of shape representation (as closed curve-contour) and curve parameterization is still opened (Zhang & Lu, 2004). Operational research in planning and scheduling, also decision making systems in risk analysis, solve the problems which are based on data modeling and extrapolation via the choice of key points. So interpolation and extrapolation is not only a pure mathematical problem but important task in economic and artificial intelligence. The chapter wants to approach a problem of curve modeling by characteristic points. Proposed method relies on nodes combination and functional modeling of curve points situated between the basic set of key points and outside of this set. The functions that are used in calculations represent whole family of elementary functions with inverse functions: polynomials, trigonometric, cyclometric, logarithmic, exponential and power function. These functions are treated as probability distribution functions in the range [0;1].

An important problem in operational research and computer sciences (Ballard, 1982) is that of appropriate shape representation and reconstruction. Classical discussion about shape representation is based on the problem: contour versus skeleton. This chapter is voting for contour which forms boundary of the object. Contour of the object, represented by contour points, consists of information which allows us to describe many important features of the object as shape coefficients (Tadeusiewicz & Flasiński, 1991). In the chapter contour is dealing with a set of curves. Curve modeling and generation is a basic subject in many branches of industry and computer science, for example in the CAD/CAM software.

The representation of shape has a great impact on the accuracy and effectiveness of object recognition (Saber, Yaowu & Murat, 2005). In the literature, shape has been represented by many options including curves (Sebastian & Klein, 2003), graph-based algorithms and medial axis (Liu & Geiger, 1999) to enable shape-based object recognition. Digital curve (open or closed) can be represented by chain code (Freeman's code). Chain code depends on selection of the started point and transformations of the

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

object. So Freeman's code is one of the methods how to describe and to find contour of the object. An analog (continuous) version of Freeman's code is the curve α-*s*. Another contour representation and reconstruction is based on Fourier coefficients calculated in Discrete Fourier Transformation (DFT). These coefficients are used to fix similarity of the contours with different sizes or directions. If we assume that contour is built from segments of a line and fragments of circles or ellipses, Hough transformation is applied to detect contour lines. Also geometrical moments of the object are used during the process of object shape representation (Choraś, 2005). Proposed method, called Probabilistic Nodes Combination (PNC), is the method of 2D curve interpolation and extrapolation using the set of key points (knots or nodes). Nodes can be treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PNC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. Two-dimensional curve is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function. Novelty of this book consists of two generalizations: generalization of previous MHR method with various nodes combinations and generalization of linear interpolation with different (no basic) probability distribution functions and nodes combinations.

Curve interpolation represents one of the most important problems in mathematics and computer science: how to model the curve via discrete set of two-dimensional points? Also the matter of shape representation (as closed curve-contour) and curve parameterization is still opened. For example pattern recognition, signature verification or handwriting identification problems are based on curve modeling via the choice of key points. So interpolation is not only a pure mathematical problem but important task in computer vision and artificial intelligence. The monograph wants to approach a problem of curve modeling by characteristic points. Proposed method relies on nodes combination and functional modeling of curve points situated between the basic set of key points. The functions that are used in calculations represent whole family of elementary functions with inverse functions: polynomials, trigonometric, cyclometric, logarithmic, exponential and power function. These functions are treated as probability distribution functions in the range [0;1]. The method of Probabilistic Nodes Combination (PNC) enables interpolation and modeling of two-dimensional curves using nodes combinations and different coefficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function, also inverse functions. This probabilistic view is novel approach a problem of modeling and interpolation. Computer vision and pattern recognition are interested in appropriate methods of shape representation and curve modeling. PNC method represents the possibilities of shape reconstruction and curve interpolation via the choice of nodes combination and probability distribution function for interpolated points. It seems to be quite new look at the problem of contour representation and curve modeling in artificial intelligence and computer vision.

BACKGROUND

All interpolation theory is based on polynomials. But why? Many kinds of polynomials are used for interpolation: classical polynomials, trigonometric polynomials, orthogonal polynomials (Tschebyscheff, Legendre, Laguerre), rational polynomials. But what about the exceptional situations with unexpected features of curve, data or nodes. Then polynomials are not the solution, for example when:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. The curve is not a graph of function (no matter – opened or closed curve);
2. The curve does not have to be smooth at interpolation nodes: for example curve representing symbols, signature, handwriting or other specific data;
3. Nodes are fixed and there is no possibility to choose "better" nodes as for orthogonal polynomials;
4. The curve differs considerably from any interpolation polynomial;
5. The curve fails to be differentiable at some points;
6. Between each pair of nodes we are not interested in linear interpolation (uniform probability distribution and zero nodes combination) but there ought to be some generalization (even for two nodes only) with other probability distributions and nodes combinations;
7. Interpolated points depend on some chosen nodes (two nearest nodes or more) via nodes combination $h(p_1,p_2,\ldots,p_m)$ in (1);
8. We are not interested in the formula of interpolation function (for lower computational costs) but only calculated points of modeled curve are ready to be used in numerical computations;
9. The formula of curve or function is known but from some reason (for example high computational costs or hard polynomial interpolation) the curve has to be modeled or fitted in some way for numerical calculations – the examples for PNC interpolation (in MHR version) of functions $f(x) = 2/x$ and $f(x) = 1/(1+5x^2)$ with quantified measures and experimental comparison with classical polynomial interpolation in (Jakóbczak, 2010);
10. Extrapolation problem is also a big numerical challenge and PNC interpolation enables the extension into extrapolation (Jakóbczak, 2011) with α outside of [0;1] and $\gamma = F(\alpha)$ still strictly monotonic, $F(0) = 0$, $F(1) = 1$. So for example $\gamma = \alpha^2$ is impossible for extrapolation if $\alpha < 0$ (Jakóbczak, 2011). Polynomial or other interpolations are sometimes useless for extrapolation;
11. Having only nodes the user may have "negative" information (from specific character of data): no polynomial interpolation;
12. All calculations are numerical (discrete) – even $\gamma = F(\alpha)$ is to be given in tabular (discrete) form. There is no need to build continuous function: polynomial or others;
13. Parametric version of modeled curve is to be found.

Above thirteen important and heavy individual and characteristic features of some curves and their interpolations show that there may exist the situations with unexpected assumptions for interpolation.

Why not classical interpolation? Classical methods are useless to interpolate the function that fails to be differentiable at one point, for example the absolute value function $f(x) = |x|$ at $x = 0$. If point (0;0) is one of the interpolation nodes, then precise polynomial interpolation of the absolute value function is impossible. Also when the graph of interpolated function differs from the shape of polynomial considerably, for example $f(x) = 1/x$, interpolation is very hard because of existing local extrema and the roots of polynomial. We cannot forget about the Runge's phenomenon: when nodes are equidistance then high-order polynomial oscillates toward the end of the interval, for example close to -1 and 1 with function $f(x) = 1/(1+25x^2)$ (Ralston & Rabinowitz, 2001). These classical negative cases do not appear in proposed PNC method. Experimental comparison for PNC with polynomial interpolation is to be found in (Jakóbczak, 2009).

Nowadays methods apply mainly polynomial functions in different versions (trigonometric, orthogonal, rational) and for example Bernstein polynomials in Bezier curves, splines (Schumaker, 2007) and NURBS (Rogers, 2001). But Bezier curves don't represent the interpolation method (rather interpolation-approximation method) and cannot be used for example in handwriting modeling with key points (interpolation

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

nodes). In comparison PNC method with Bézier curves, Hermite curves and B-curves (*B-splines*) or NURBS one unpleasant feature of these curves has to be mentioned: small change of one characteristic point can result in unwanted change of whole reconstructed curve. Such a feature does not appear in proposed PNC method which is more stable than Bézier curves. Only first and last characteristic point is situated on the Bézier curve (interpolation), the rest characteristic points lay outside the Bézier curve (approximation). Numerical methods for data interpolation are based on polynomial or trigonometric functions, for example Lagrange, Newton, Aitken and Hermite methods. These methods have many weak sides (Dahlquist & Bjoerck, 1974) and are not sufficient for curve interpolation in the situations when the curve cannot be build by polynomials or trigonometric functions. Also there exists several well established methods of curve modeling, for example shape-preserving techniques (Dejdumrong, 2007), subdivision algorithms (Dyn, Levin & Gregory, 1987) and others (Kozera, 2004) to overcome difficulties of polynomial interpolation, but probabilistic interpolation with nodes combination seems to be quite novel in the area of shape modeling. Proposed 2D curve interpolation is the functional modeling via any elementary functions and it helps us to fit the curve during the computations.

This chapter presents novel Probabilistic Nodes Combination (PNC) method of curve interpolation. This chapter takes up new PNC method of two-dimensional curve modeling via the examples using the family of Hurwitz-Radon matrices (MHR method) (Jakóbczak, 2007), but not only this method (other nodes combinations). The method of PNC requires minimal assumptions: the only information about a curve is the set of at least two nodes. Proposed PNC method is applied in curve modeling via different coefficients: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power. Function for PNC calculations is chosen individually at each interpolation and it represents probability distribution function of parameter $\alpha \in [0;1]$ for every point situated between two interpolation knots. PNC method uses two-dimensional vectors (x,y) for curve modeling-knots $p_i = (x_i,y_i) \in \boldsymbol{R^2}$ in PNC method, $i = 1,2,\ldots n$:

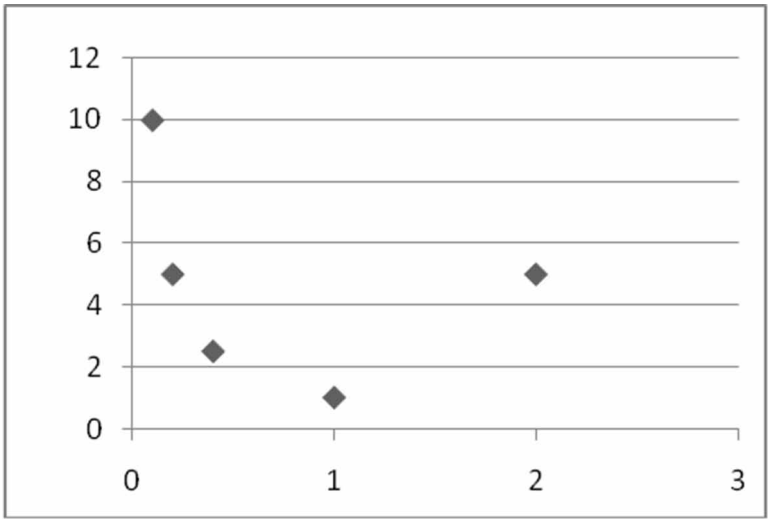
1. PNC needs 2 knots or more ($n \geq 2$);
2. If first node and last node are the same ($p_1 = p_n$), then curve is closed (contour);
3. For more precise modeling knots ought to be settled at key points of the curve, for example local minimum or maximum and at least one node between two successive local extrema.

Condition 3 means for example the highest point of the curve in a particular orientation, convexity changing or curvature extrema. So this chapter wants to answer the question: how to interpolate the curve by a set of knots (Jakóbczak, 2010)?

Nodes on Figure 1 represent characteristic points of the handwritten letter or symbol: if $n = 5$ then curve is opened and if $n = 6$ then curve is closed (contour). The examples of PNC curve modeling for these nodes are described later in this chapter. Coefficients for PNC curve modeling are computed using nodes combinations and probability distribution functions: polynomials, power functions, sine, cosine, tangent, cotangent, logarithm, exponent or arc sin, arc cos, arc tan or arc cot.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Five knots of the curve before modeling
Source: Author, 2016.



NOVELTY OF PROBABILISTIC INTERPOLATION AND EXTRAPOLATION

Issues

The method of PNC enables to compute points between two successive nodes of the curve: calculated points are interpolated and parameterized for real number $\alpha \in [0;1]$ in the range of two successive nodes. PNC method uses the combinations of nodes $p_1=(x_1,y_1)$, $p_2=(x_2,y_2),\ldots$, $p_n=(x_n,y_n)$ as $h(p_1,p_2,\ldots,p_m)$ and $m = 1,2,\ldots n$. Nodes combination h is defined individually for each curve to interpolate points (x,y) with second coordinate $y = y(c)$ for any first coordinate $x = c$ situated between nodes (x_i,y_i) and (x_{i+1},y_{i+1}):

$c = \alpha \cdot x_i + (1-\alpha) \cdot x_{i+1}$, $i = 1,2,\ldots n-1$,

$$y(c) = \gamma \cdot y_i + (1-\gamma) y_{i+1} + \gamma(1-\gamma) \cdot h(p_1, p_2, \ldots, p_m), \quad (1)$$

$\alpha \in [0;1]$, $\gamma = F(\alpha) \in [0;1]$, $F:[0;1] \rightarrow [0;1]$, F(0)=0, F(1)=1 and F is strictly monotonic.

PNC extrapolation requires α outside of $[0;1]$: $\alpha < 0$ (anticipating points right of last node for $c > x_n$) or $\alpha > 1$ (extrapolating values left of first node for $c < x_1$), $\gamma = F(\alpha)$, $F:P \rightarrow \mathbf{R}$, $P \supset [0;1]$, F(0)=0, F(1)=1 and F is still strictly monotonic for the arguments from P. So c and α represent the same – coordinate x of any point (x,y) between two successive nodes (x_i,y_i) and (x_{i+1},y_{i+1}): having c we can calculate α and vice versa. PNC curve modeling relies on two factors: function $\gamma = F(\alpha)$ and nodes combination $h(p_1,p_2,\ldots,p_m)$. Function F is a probabilistic distribution function for random variable $\alpha \in [0;1]$ and parameter γ leads PNC interpolation into probabilistic modeling. Second factor, the combination of nodes h, is responsible for making dependent a reconstructed point on the coordinates of several nodes. The simplest case is for $h = 0$. Here are the examples of h computed for MHR method (Jakóbczak, 2009):

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$h(p_1, p_2) = \frac{y_1}{x_1} x_2 + \frac{y_2}{x_2} x_1$$

(only two neighboring nodes are taken for PNC calculations) or

$$h(p_1, p_2, p_3, p_4) = \frac{1}{x_1^2 + x_3^2}(x_1 x_2 y_1 + x_2 x_3 y_3 + x_3 x_4 y_1 - x_1 x_4 y_3) + \frac{1}{x_2^2 + x_4^2}(x_1 x_2 y_2 + x_1 x_4 y_4 + x_3 x_4 y_2 - x_2 x_3 y_4)$$

(more than two neighboring nodes are used in PNC interpolation).

The examples of other nodes combinations are presented below. Formula (1) represents curve parameterization ($x(\alpha)$,$y(\alpha)$) between two successive nodes (x_i,y_i) and (x_{i+1},y_{i+1}) as $\alpha \in [0;1]$:

$$x(\alpha) = \alpha \cdot x_i + (1\text{-}\alpha) \cdot x_{i+1}$$

and

$$y(\alpha) = F(\alpha) \cdot y_i + (1 - F(\alpha)) y_{i+1} + F(\alpha)(1 - F(\alpha)) \cdot h(p_1, p_2, ..., p_m),$$

$$y(\alpha) = F(\alpha) \cdot (y_i - y_{i+1} + (1 - F(\alpha)) \cdot h(p_1, p_2, ..., p_m)) + y_{i+1}.$$

Proposed parameterization gives us the infinite number of possibilities for curve calculations (determined by choice of F and h) as there is the infinite number of human handwritten letters and symbols. Nodes combination is the individual feature of each modeled curve (for example a handwritten character). Coefficient $\gamma = F(\alpha)$ and nodes combination h are key factors in PNC curve interpolation and data extrapolation.

Distribution Functions in PNC Interpolation and Extrapolation

Points settled between the nodes are computed using PNC method. Each real number $c \in [a;b]$ is calculated by a convex combination $c = \alpha \cdot a + (1\text{-}\alpha) \cdot b$ for

$$\alpha = \frac{b - c}{b - a} \in [0;1].$$

Key question is dealing with coefficient γ in (1). The simplest way of PNC calculation means $h = 0$ and $\gamma = \alpha$ (basic probability distribution – uniform distribution). Then PNC represents a linear interpolation. MHR method (Jakóbczak, 2010) is not a linear interpolation. MHR is the example of PNC modeling.

Each interpolation requires specific distribution of parameter α and γ (1) depends on parameter $\alpha \in [0;1]$:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$\gamma = F(\alpha)$, F:[0;1]→[0;1], $F(0) = 0$, $F(1) = 1$ and F is strictly monotonic.

Coefficient γ is calculated using different functions (polynomials, power functions, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan or arc cot, also inverse functions) and choice of function is connected with initial requirements and curve specifications. Different values of coefficient γ are connected with applied functions $F(\alpha)$. The functions (2)-(34) represent the examples of probability distribution functions for random variable $\alpha \in [0;1]$ and real number $s > 0$:

1. **Power Function:**

$$\gamma = \alpha^s \text{ with } s > 0. \tag{2}$$

For $s = 1$: basic version of PNC and MHR (Jakóbczak, 2010) methods when $\gamma = \alpha$.

2. **Sine:**

$$\gamma = sin(\alpha^s \cdot \pi/2),\ s > 0 \tag{3}$$

or

$$\gamma = sin^s(\alpha \cdot \pi/2),\ s > 0. \tag{4}$$

For $s = 1$:

$$\gamma = sin(\alpha \cdot \pi/2). \tag{5}$$

3. **Cosine:**

$$\gamma = 1\text{-}cos(\alpha^s \cdot \pi/2),\ s > 0 \tag{6}$$

or

$$\gamma = 1\text{-}cos^s(\alpha \cdot \pi/2),\ s > 0. \tag{7}$$

For $s = 1$:

$$\gamma = 1\text{-}cos(\alpha \cdot \pi/2). \tag{8}$$

4. **Tangent:**

$$\gamma = tan(\alpha^s \cdot \pi/4),\ s > 0 \tag{9}$$

or

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$\gamma = tan^{s}(\alpha \cdot \pi/4),\ s > 0. \tag{10}$$

For $s = 1$:

$$\gamma = tan(\alpha \cdot \pi/4). \tag{11}$$

5. **Logarithm:**

$$\gamma = log_2(\alpha^{s} + 1),\ s > 0 \tag{12}$$

or

$$\gamma = log_2^{s}(\alpha + 1),\ s > 0. \tag{13}$$

For $s = 1$:

$$\gamma = log_2(\alpha + 1). \tag{14}$$

6. **Exponent:**

$$\gamma = \left(\frac{a^{\alpha} - 1}{a - 1}\right)^{s},\ s > 0 \text{ and } a > 0 \text{ and } a \neq 1. \tag{15}$$

For $s = 1$ and $a = 2$:

$$\gamma = 2^{\alpha} - 1. \tag{16}$$

7. **Arc Sine:**

$$\gamma = (2/\pi) \cdot arcsin(\alpha^{s}),\ s > 0 \tag{17}$$

or

$$\gamma = (2/\pi \cdot arcsin\ \alpha)^{s},\ s > 0. \tag{18}$$

For $s = 1$:

$$\gamma = (2/\pi) \cdot arcsin(\alpha). \tag{19}$$

8. **Arc Cosine:**

$$\gamma = 1\text{-}(2/\pi) \cdot arccos(\alpha^{s}),\ s > 0 \tag{20}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

or

$$\gamma = 1\text{-}(2/\pi \cdot arccos\ \alpha)^s,\ s > 0. \tag{21}$$

For $s = 1$:

$$\gamma = 1\text{-}(2/\pi) \cdot arccos(\alpha). \tag{22}$$

9. **Arc Tangent:**

$$\gamma = (4/\pi) \cdot arctan(\alpha^s),\ s > 0 \tag{23}$$

or

$$\gamma = (4/\pi \cdot arctan\ \alpha)^s,\ s > 0. \tag{24}$$

For $s = 1$:

$$\gamma = (4/\pi) \cdot arctan(\alpha). \tag{25}$$

10. **Cotangent:**

$$\gamma = ctg(\pi/2 - \alpha^s \cdot \pi/4),\ s > 0 \tag{26}$$

or

$$\gamma = ctg^s\ (\pi/2\text{-}\alpha \cdot \pi/4),\ s > 0. \tag{27}$$

For $s = 1$:

$$\gamma = ctg(\pi/2\text{-}\alpha \cdot \pi/4). \tag{28}$$

11. **Arc Cotangent:**

$$\gamma = 2 - (4/\pi) \cdot arcctg(\alpha^s),\ s > 0 \tag{29}$$

or

$$\gamma = (2\text{-}4/\pi \cdot arcctg\ \alpha)^s,\ s > 0. \tag{30}$$

For $s = 1$:

$$\gamma = 2 - (4/\pi) \cdot arcctg(\alpha). \tag{31}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Functions used in γ calculations (2)-(31) are strictly monotonic for random variable $\alpha \in [0;1]$ as $\gamma = F(\alpha)$ is probability distribution function. Of course any monotonic combination of functions (2)-(31) between points (0;0) and (1;1) is also good, for example

$$\gamma = \alpha^5 \cdot (4/\pi) \cdot arctan(\alpha)$$

or

$$\gamma = \tfrac{1}{2}(\alpha^{7.6} + sin^3(\alpha \cdot \pi/2))$$

or

$$\gamma = \tfrac{1}{4}(\alpha^{1.6} + \alpha^{0.6} + \alpha^6 + sin^{3.9}(\alpha \cdot \pi/2))$$

or

$$\gamma = \tfrac{1}{2}(\alpha^{7.98} + \alpha^3 \cdot (4/\pi) \cdot arctan(\alpha))$$

or other modeling functions.

Also inverse function $F^{-1}(\alpha)$ is appropriate for γ calculations. Choice of function and value s depends on curve specifications and individual requirements. Proposed (2)-(31) probability distributions are continuous, but of course parameter γ can represent discrete probability distributions, for example: $F(0.1)=0.23$, $F(0.2)=0.3$, $F(0.3)=0.42$, $F(0.4)=0.52$, $F(0.5)=0.63$, $F(0.6)=0.69$, $F(0.7)=0.83$, $F(0.8)=0.942$, $F(0.9)=0.991$. What is very important in PNC method: two curves (for example a handwritten letter) may have the same set of nodes but different h or γ results in different interpolations (Figure 2-10).

Algorithm of PNC interpolation and modeling (1) looks as follows:

Step 1: Choice of knots p_i at key points.
Step 2: Choice of nodes combination $h(p_1,p_2,\ldots,p_m)$.
Step 3: Choice of distribution $\gamma = F(\alpha)$: (2)-(31) or others (continuous or discrete).
Step 4: Determining values of α: $\alpha = 0.1, 0.2 \ldots 0.9$ (nine points) or $0.01, 0.02 \ldots 0.99$ (99 points) or others.
Step 5: The computations (1).

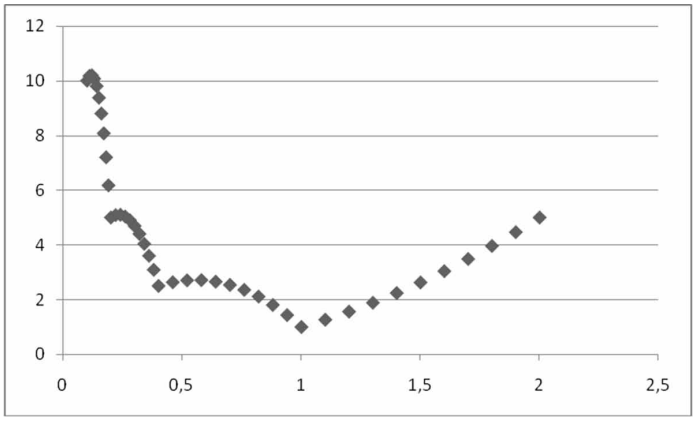
These five steps can be treated as the algorithm of PNC method of curve modeling and interpolation (1). Without knowledge about the formula of curve or function, PNC interpolation has to implement the coefficients γ (2)-(31), but PNC is not limited only to these coefficients. Each strictly monotonic function F between points (0;0) and (1;1) can be used in PNC modeling.

Data Modeling and Curve Fitting

Curve knots $p_1 = (0.1;10)$, $p_2 = (0.2;5)$, $p_3 = (0.4;2.5)$, $p_4 = (1;1)$ and $p_5 = (2;5)$ from Figure 1 are used in some examples of PNC method in handwritten character modeling. Figures 2-9 represent PNC as MHR interpolation (Jakóbczak, 2011) with different γ. Points of the curve are calculated with no matrices ($N = 1$) and $\gamma = \alpha$ in example 1 and with matrices of dimension $N = 2$ in examples 2-8 for $\alpha = 0.1, 0.2, \ldots, 0.9$.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. PNC character modeling for nine reconstructed points between nodes
Source: Author, 2016.



Example 1: (refer to Figure 2)
PNC curve interpolation (1) for $\gamma = \alpha$ and

$$h(p_1, p_2) = \frac{y_1}{x_1} x_2 + \frac{y_2}{x_2} x_1 :$$

For $N = 2$ (examples 2 – 8) MHR version (Jakóbczak, 2011) as PNC method gives us

$$h(p_1, p_2, p_3, p_4) = \frac{1}{x_1^{\,2} + x_3^{\,2}} (x_1 x_2 y_1 + x_2 x_3 y_3 + x_3 x_4 y_1 - x_1 x_4 y_3) + \frac{1}{x_2^{\,2} + x_4^{\,2}} (x_1 x_2 y_2 + x_1 x_4 y_4 + x_3 x_4 y_2 - x_2 x_3 y_4)$$

Example 2: (refer to Figure 3)
PNC sinusoidal interpolation with $\gamma = sin(\alpha \cdot \pi/2)$.
Example 3: (refer to Figure 4)
PNC tangent interpolation for $\gamma = tan(\alpha \cdot \pi/4)$.
Example 4: (refer to Figure 5)
PNC tangent interpolation with $\gamma = tan(\alpha^s \cdot \pi/4)$ and $s = 1.5$.
Example 5: (refer to Figure 6)
PNC tangent curve interpolation for $\gamma = tan(\alpha^s \cdot \pi/4)$ and $s = 1.797$.
Example 6: (refer to Figure 7)
PNC sinusoidal interpolation with $\gamma = sin(\alpha^s \cdot \pi/2)$ and $s = 2.759$.
Example 7: (refer to Figure 8)
PNC power function modeling for $\gamma = \alpha^s$ and $s = 2.1205$.
Example 8: (refer to Figure 9)
PNC logarithmic curve modeling with $\gamma = log_2(\alpha^s + 1)$ and $s = 2.533$.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Sinusoidal modeling with nine reconstructed curve points between nodes
Source: Author, 2016.

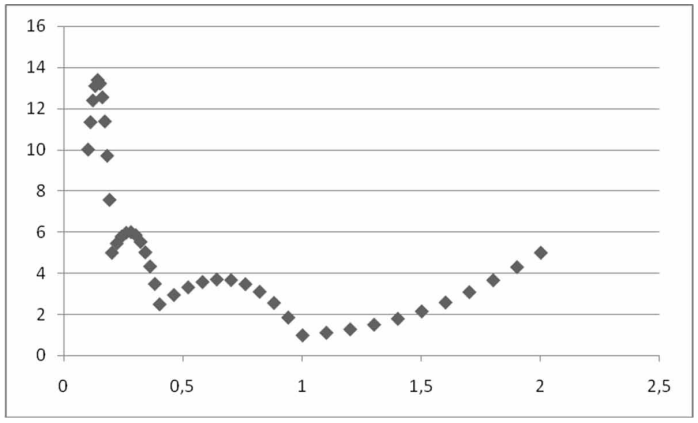


Figure 4. Tangent character modeling with nine interpolated points between nodes
Source: Author, 2016

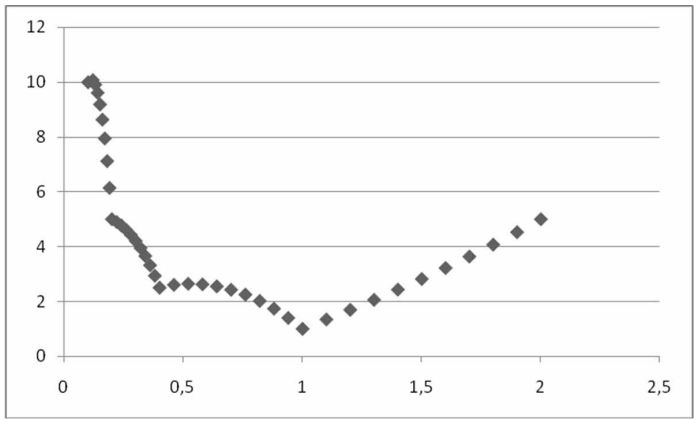
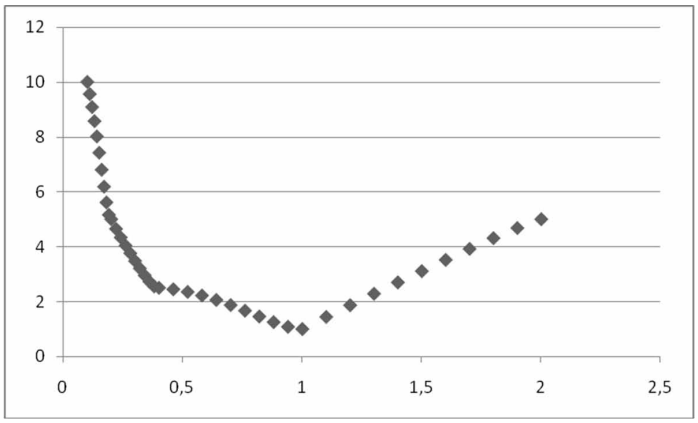


Figure 5. Tangent curve modeling with nine recovered points between nodes
Source: Author, 2016.



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Tangent symbol modeling with nine reconstructed points between nodes
Source: Author, 2016.

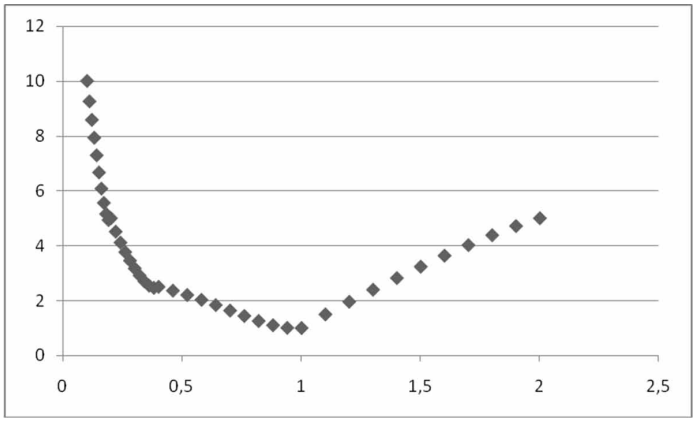


Figure 7. Sinusoidal modeling with nine interpolated curve points between nodes
Source: Author, 2016.

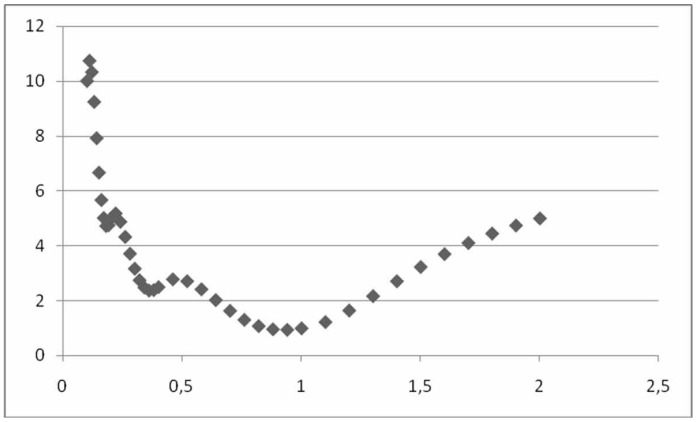
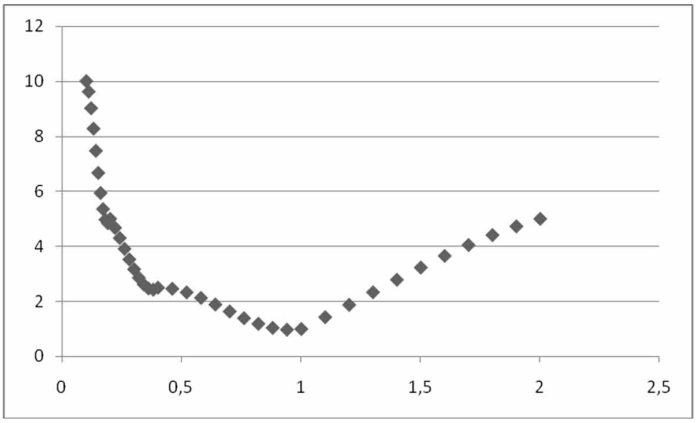
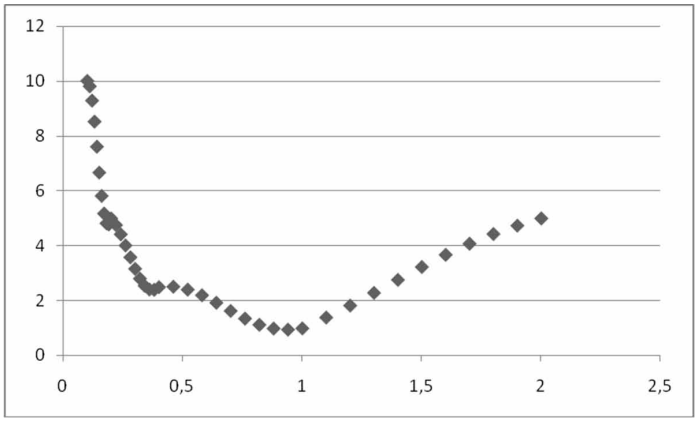


Figure 8. Power function curve modeling with nine recovered points between nodes
Source: Author, 2016.



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Logarithmic character modeling with nine reconstructed points between nodes
Source: Author, 2016.



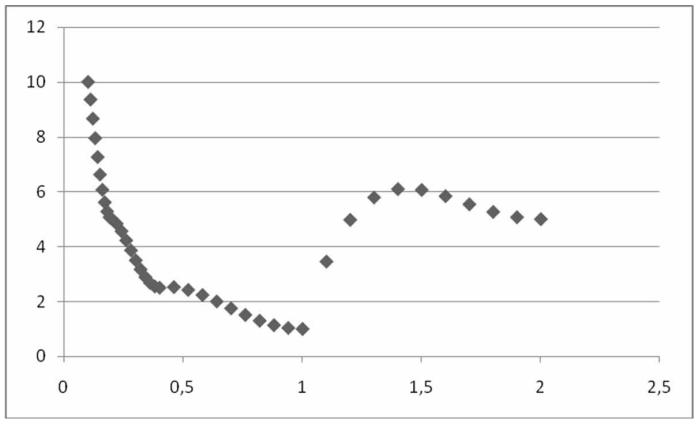
These eight examples demonstrate possibilities of PNC curve interpolation and handwritten character modeling for key nodes in MHR version. And here are other examples of PNC modeling (but not MHR):

Example 9: (refer to Figure 10)
PNC for $\gamma = \alpha^2$ and $h(p_1,p_2) = x_1y_1 + x_2y_2$.
Example 10: (refer to Figure 11)
PNC for $\gamma = \alpha^3$ and $h(p_1,p_2) = x_1y_1 + x_2y_2$.

If there is considered Figure 1 as closed curve (contour) with the node $p_6 = p_1 = (0.1;10)$ then examples 9 and10 give the shapes:

Example 11: (refer to Figure 12)
PNC for $\gamma = \alpha^2$ and $h(p_1,p_2) = x_1y_1 + x_2y_2$.

Figure 10. Quadratic symbol modeling with nine reconstructed points between nodes
Source: Author, 2016.



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Cubic character modeling with nine reconstructed points between nodes
Source: Author, 2016.

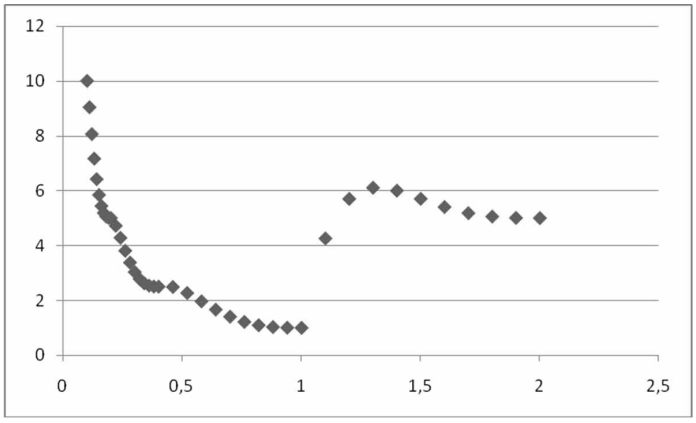
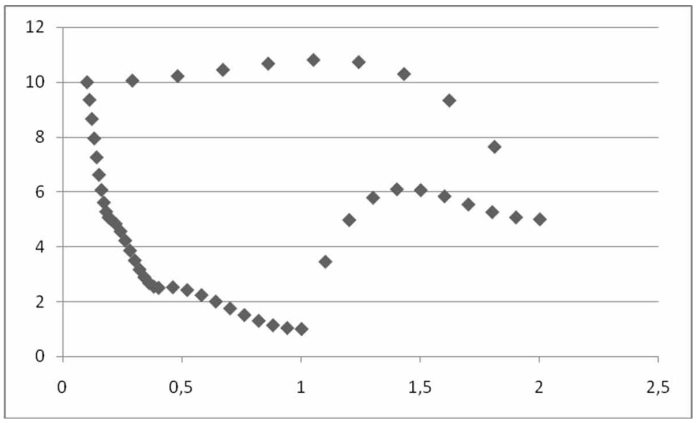


Figure 12. Quadratic contour modeling with nine reconstructed points between nodes
Source: Author, 2016.



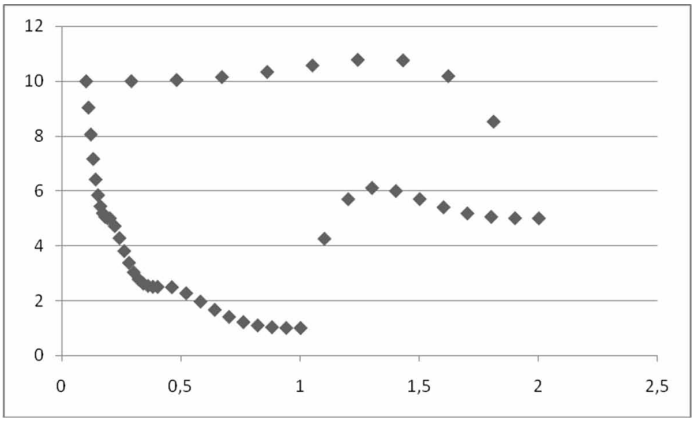
Example 12: (refer to Figure 13)
PNC for $\gamma = \alpha^3$ and $h(p_1,p_2) = x_1y_1 + x_2y_2$.

Every man has individual style of handwriting. Recognition of handwritten letter or symbol need modeling and the model of each individual symbol or character can be built by choice of γ and h in (1). PNC modeling via nodes combinations h and parameter γ as probability distribution function enables curve interpolation for each specific letter or symbol.

Number of reconstructed points depends on a user by value α. If for example $\alpha = 0.01, 0.02,\ldots,0.99$ then 99 points are interpolated for each pair of nodes. Reconstructed values and interpolated points, calculated by PNC method, are applied in the process of curve modeling. Every curve can be interpolated by some distribution function as parameter γ and nodes combination h. Parameter γ is treated as probability distribution function for each curve.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 13. Cubic shape modeling with nine reconstructed points between nodes
Source: Author, 2016.



Beta Distribution

Considering nowadays used probability distribution functions for random variable $\alpha \in [0;1]$-one distribution is dealing with the range [0;1]: beta distribution. Probability density function f for random variable $\alpha \in [0;1]$ is:

$$f(\alpha) = c \bullet \alpha^s \bullet (1-\alpha)^r,\ s \geq 0,\ r \geq 0. \tag{32}$$

When $r = 0$ probability density function (32) represents $f(\alpha) = c \bullet \alpha^s$ and then probability distribution function F is like (2), for example $f(\alpha)=3\alpha^2$ and $\gamma = \alpha^3$. If s and r are positive integer numbers then γ is the polynomial, for example $f(\alpha)=6\alpha(1-\alpha)$ and $\gamma = 3\alpha^2-2\alpha^3$. So beta distribution gives us coefficient γ in (1) as polynomial because of interdependence between probability density f and distribution F functions:

$$f(\alpha) = F'(\alpha),\ \ F(\alpha) = \int_0^{\alpha} f(t)dt\,. \tag{33}$$

For example (33):

$$f(\alpha) = \alpha \bullet e^{\alpha} \text{ and } \gamma = F(\alpha) = 1 - (1-\alpha)e^{\alpha}. \tag{34}$$

Basic (uniform) distribution ($\gamma = \alpha$) with nodes combination $h = 0$ turns PNC interpolation (1) to linear interpolation. What about PNC in the case of yet another distribution on the range [0;1]: beta distribution (32)? Power functions as γ used in examples 1, 7 and 9-12 are also connected with beta distribution. Here are the examples of PNC modeling for beta distribution with nodes combination $h = 0$.

Example 13: (refer to Figure 14)
PNC for $\gamma = 3\alpha^2-2\alpha^3$ and $h(p_1,p_2)=0$.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Example 14: (refer to Figure 15)
PNC for $\gamma = 4\alpha^3\text{-}3\alpha^4$ and $h(p_1,p_2)=0$.
Example 15: (refer to Figure 16)
PNC for $\gamma = 2\alpha\text{-}\alpha^2$ and $h(p_1,p_2)=0$.

Examples 9-12 represent beta distribution with $h(p_1,p_2) = x_1y_1 + x_2y_2$.

Exponential Distribution

Exponential distribution is dealing with random variable ≥ 0, but in PNC interpolation random variable $\alpha\in[0;1]$. Then exponential distribution is represented by distribution function (34):

$$\gamma = F(\alpha) = 1 - (1 - \alpha)e^{\alpha}.$$

Figure 14. Beta distribution in handwritten character modeling
Source: Author, 2016.

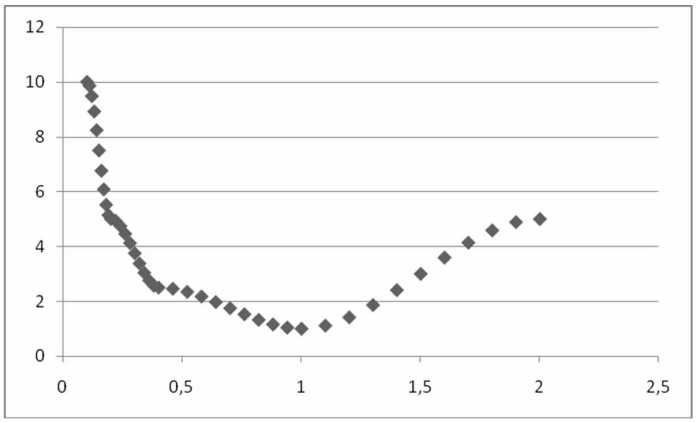
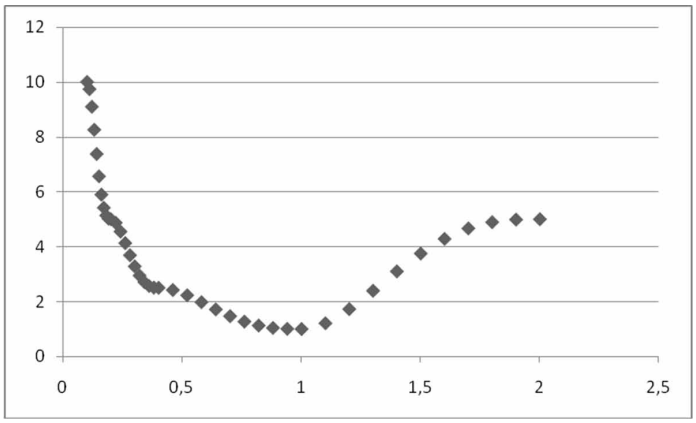
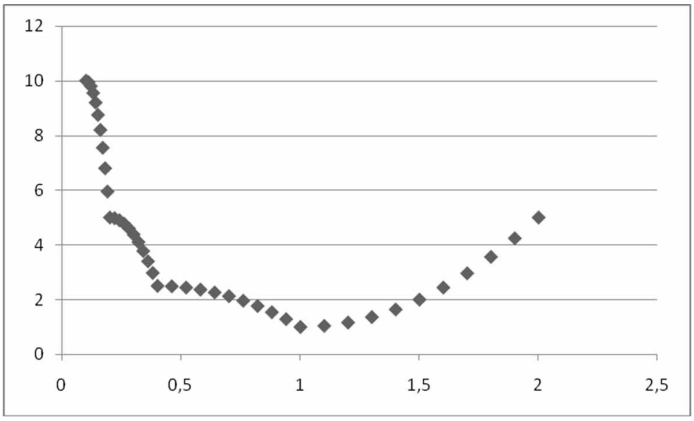


Figure 15. Beta distribution in handwritten symbol modeling
Source: Author, 2016.



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 16. Beta distribution in handwritten letter modeling
Source: Author, 2016.



Example 16: (refer to Figure 17)

PNC for $\gamma = 1\text{-}(1\text{-}\alpha)e^{\alpha}$ and $h(p_1,p_2)=0$.

Example 17: (refer to Figure 18)

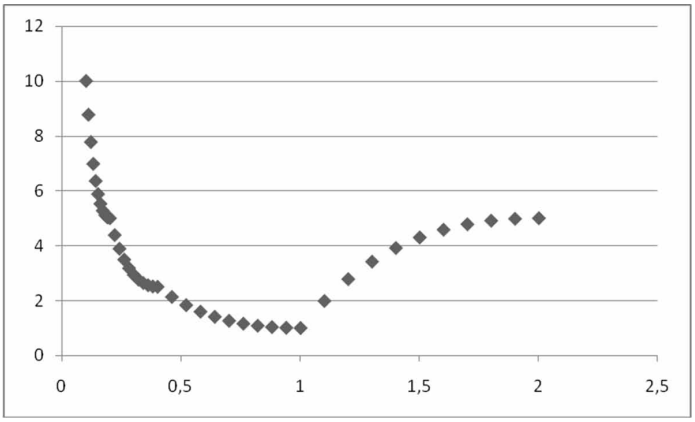
PNC for $\gamma = 1\text{-}(1\text{-}\alpha)e^{\alpha}$ and $h(p_1, p_2) = \frac{y_2}{y_1} + \frac{x_2}{x_1}$:

These examples show the variety of possibilities in curve modeling via the choice of nodes combination and probability distribution function for interpolated points.

PNC Extrapolation as the Support in Planning

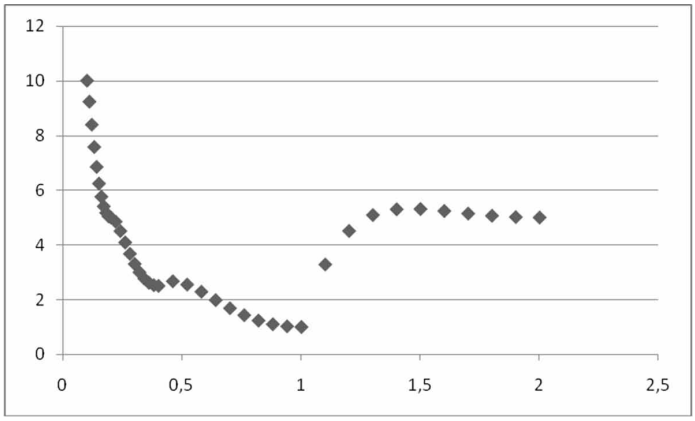
Unknown data, important for planning or decision making, are modeled (interpolated or extrapolated) by the choice of nodes, determining specific nodes combination and characteristic probabilistic distribution function. Less complicated models take $h(p_1,p_2,\ldots,p_m) = 0$ and then the formula of interpolation (2) looks as follows:

Figure 17. Exponential distribution in handwritten character modeling
Source: Author, 2016



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 18. Exponential distribution in handwritten symbol modeling
Source: Author, 2016.



$y(c) = \gamma \bullet y_i + (1-\gamma) y_{i+1}$.

It is linear interpolation for basic (uniform) probability distribution ($\gamma = \alpha$).

Example 1: Nodes (1;3), (3;1), (5;3), (7;3) and $h = 0$, $\gamma = F(\alpha) = \alpha^2$. This function F requires $\alpha > 1$ and extrapolation is computed with (4)-(5) (see Figure 19).
Extrapolated points: (7.2;3), (7.4;3), (7.6;3), (7.8;3), (8;3), (8.2;3), (8.4;3), (8.6;3), (8.8;3).
Example 2: Nodes (1;3), (3;1), (5;3), (7;2) and $h = 0$, $\gamma = F(\alpha) = \alpha^2$. This function F requires $\alpha > 1$ and extrapolation is computed with (4)-(5) too (see Figure 20)
Extrapolated points: (7.2;1.79), (7.4;1.56), (7.6;1.31), (7.8;1.04), (8;0.75), (8.2;0.44), (8.4;0.11), (8.6;-0.24), (8.8;-0.61).
Example 3: Nodes (1;3), (3;1), (5;3), (7;4) and $h = 0$, $\gamma = F(\alpha) = \alpha^3$ (see Figure 21).
Extrapolated points: (7.2;4.331), (7.4;4.728), (7.6;5.197), (7.8;5.744), (8;6.375), (8.2;7.096), (8.4;7.913), (8.6;8.832), (8.8;9.859).

These three examples 1-3 (Figures 1-3) with nodes combination $h = 0$ differ at fourth node and probability distribution functions $\gamma = F(\alpha)$. Much more possibilities of modeling are connected with a choice of nodes combination $h(p_1,p_2,\ldots,p_m)$. MHR method uses the combination (3) with good features because of orthogonal rows and columns at Hurwitz-Radon family of matrices:

$$h(p_i, p_{i+1}) = \frac{y_i}{x_i} x_{i+1} + \frac{y_{i+1}}{x_{i+1}} x_i$$

and then (2):

$$y(c) = \gamma \cdot y_i + (1-\gamma) y_{i+1} + \gamma(1-\gamma) \cdot h(p_i, p_{i+1}).$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 19. PNC modeling for nine interpolated points between successive nodes and nine extrapolated points right of the last node
Source: Author, 2016.

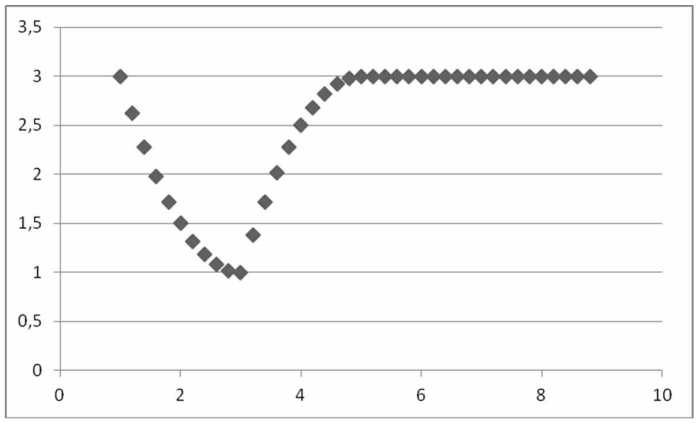
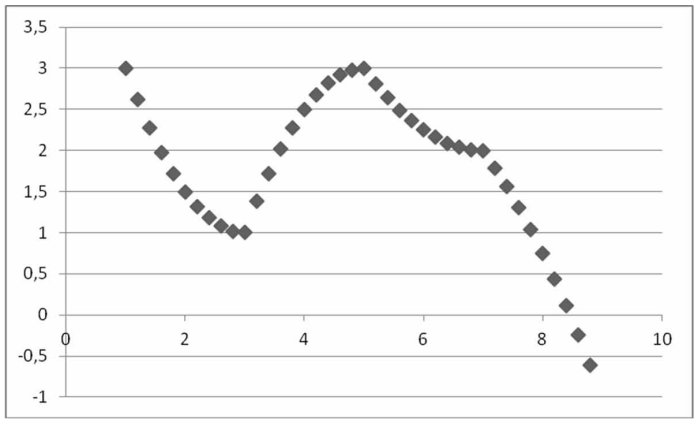


Figure 20. PNC modeling with nine interpolated points between successive nodes and nine extrapolated points right of the last node
Source: Author, 2016.



Here are two examples 4 and 5 of PNC method with MHR combination (3).

Example 4: Nodes (1;3), (3;1), (5;3) and $\gamma = F(\alpha) = \alpha^2$. This function F requires $\alpha > 1$ and extrapolation is computed with (4)-(5) (see Figure 22).

Extrapolated points: (5.2;2.539), (5.4;1.684), (5.6;0.338), (5.8;-1.603), (6;-4.25), (6.2;-7.724), (6.4;-12.155), (6.6;-17.68), (6.8;-24.443).

Example 5: Nodes (1;3), (3;1), (5;3) and $\gamma = F(\alpha) = \alpha^{1.5}$. This function F requires $\alpha > 1$ and extrapolation is computed with (4)-(5) (see Figure 23).

Extrapolated points: (5.2;2.693), (5.4;2.196), (5.6;1.487), (5.8;0.543), (6;-0.657), (6.2;-2.136), (6.4;-3.915), (6.6;-6.016), (6.8;-8.461).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 21. PNC modeling for nine interpolated points between successive nodes and nine extrapolated points right of the last node
Source: Author, 2016.

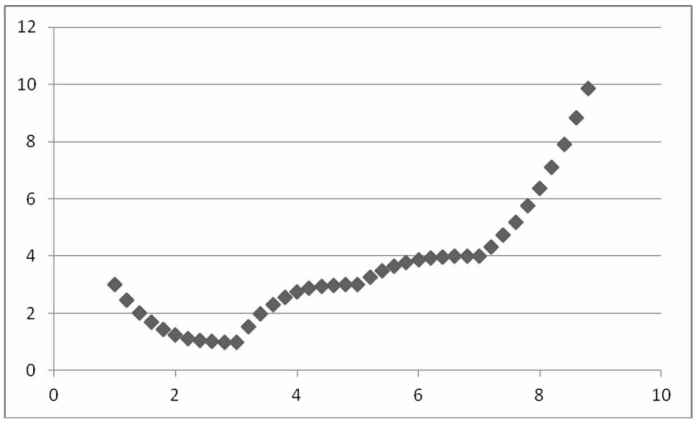
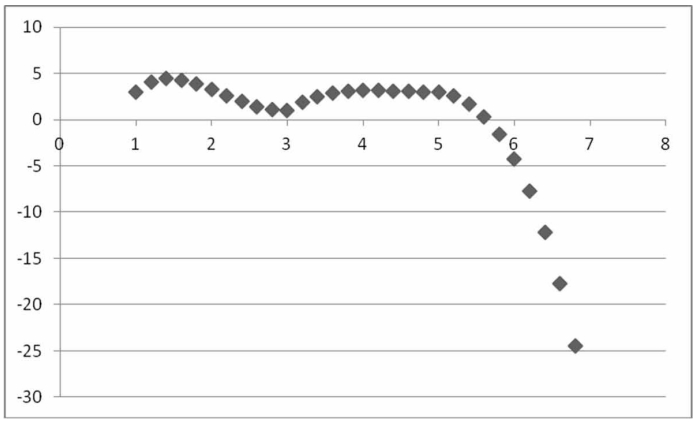


Figure 22. PNC modeling with nine interpolated points between successive nodes and nine extrapolated points right of the last node
Source: Author, 2016.



Now let us consider PNC method with other functions F than power functions, $\alpha < 0$ for extrapolation (1)-(2) and nodes combination $h = 0$.

Example 6: Nodes (2;2), (3;1), (4;2), (5;1), (6;2) and $\gamma = F(\alpha) = sin(\alpha \cdot \pi/2)$, $h = 0$ (see Figure 24).
Extrapolated points: (6.1;2.156), (6.2;2.309), (6.3;2.454), (6.4;2.588), (6.5;2.707), (6.6;2.809), (6.7;2.891), (6.8;2.951), (6.9;2.988).
Example 7: Nodes (2;2), (3;1), (4;2), (5;1), (6;2) and $\gamma = F(\alpha) = sin^3(\alpha \cdot \pi/2)$, $h = 0$ (see Figure 25).
Extrapolated points: (6.1;2.004), (6.2;2.03), (6.3;2.094), (6.4;2.203), (6.5;2.354), (6.6;2.53), (6.7;2.707), (6.8;2.86), (6.9;2.964).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 23. PNC modeling with nine interpolated points between successive nodes and nine extrapolated points right of the last node
Source: Author, 2016.

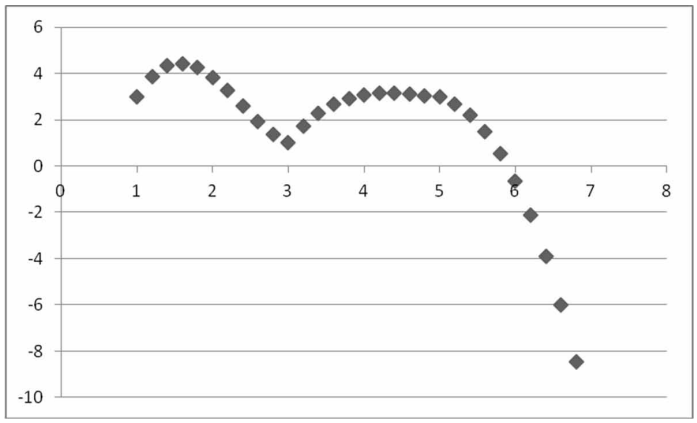
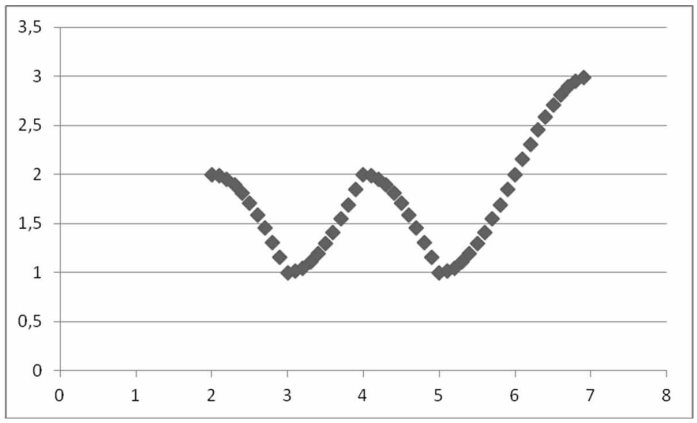


Figure 24. PNC modeling with nine interpolated points between successive nodes and nine extrapolated points right of the last node
Source: Author, 2016.



These two examples 6 and 7 (Figures 6-7) with nodes combination $h = 0$ and the same set of nodes differ only at probability distribution functions $\gamma = F(\alpha)$. Figure 8 is the example of nodes combination h as (3) in MHR method.

Example 8: Nodes (2;2), (3;1), (4;1), (5;1), (6;2) and $\gamma = F(\alpha) = 2^{\alpha}$-1 (see Figure 26).
Extrapolated points: (6.1;2.067), (6.2;2.129), (6.3;2.188), (6.4;2.242), (6.5;2.293), (6.6;2.34), (6.7;2.384), (6.8;2.426), (6.9;2.464).

Examples that are calculated above have one function $\gamma = F(\alpha)$ and one combination h for all ranges between nodes. But it is possible to create a model with functions $\gamma_i = F_i(\alpha)$ and combinations h_i individually for a range of nodes $(p_i;p_{i+1})$. Then it enables very precise modeling of data between each suc-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 25. PNC modeling with nine interpolated points between successive nodes and nine extrapolated points right of the last node
Source: Author, 2016.

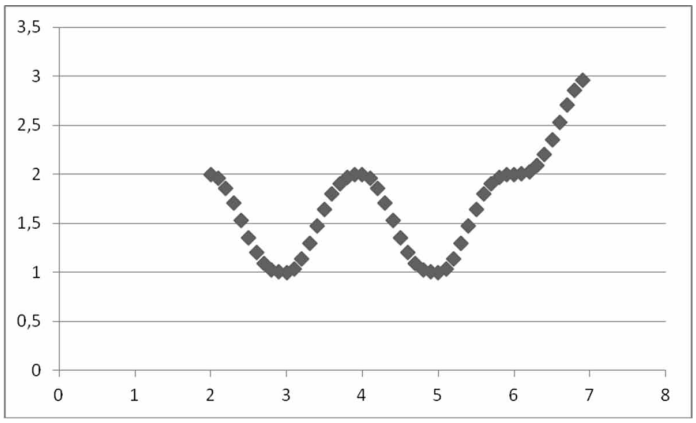
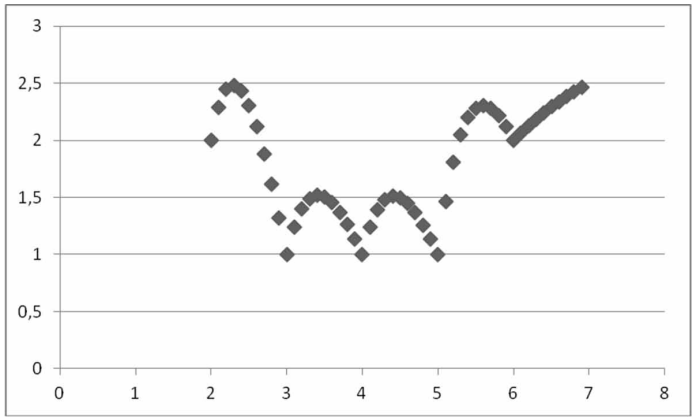


Figure 26. PNC modeling with nine interpolated points between successive nodes and nine extrapolated points right of the last node
Source: Author, 2016.



cessive pair of nodes. Each data point is interpolated or extrapolated by PNC via three factors: the set of nodes, probability distribution function $\gamma = F(\alpha)$ and nodes combination h. These three factors are chosen individually for each data, therefore this information about modeled points seems to be enough for specific PNC curve interpolation and extrapolation. Function γ is selected via the analysis of known points before extrapolation, we may assume $h = 0$ at the beginning and after some time exchange h by more adequate.

These eight examples illustrate the extrapolation of some important values in planning process, for example anticipation of some costs or expenses and foreseeing the prices or other significant data in the process of planning.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

SOLUTIONS AND RECOMMENDATIONS

Proposed method, called Probabilistic Nodes Combination (PNC), is the method of 2D curve interpolation and extrapolation using the set of key points (knots or nodes). Nodes can be treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PNC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. Two-dimensional curve is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function. Novelty of the chapter consists of two generalizations: generalization of previous MHR method with various nodes combinations and generalization of linear interpolation with different (no uniform) probability distribution functions and nodes combinations.

FUTURE RESEARCH DIRECTIONS

Future trends will go to various directions: how to fix the best probability distribution function for the nodes, how to calculate the most appropriate nodes combination and what extrapolation is the most valuable in decision making and risk analysis.

CONCLUSION

Planning process requires the anticipation of unknown values or data and foreseeing some important factors. The method of Probabilistic Nodes Combination (PNC) enables interpolation and extrapolation of two-dimensional curves using nodes combinations and different coefficients γ: polynomial, sinusoidal, cosinusoidal, tangent, cotangent, logarithmic, exponential, arc sin, arc cos, arc tan, arc cot or power function, also inverse functions. Function for γ calculations is chosen individually at each planning and it is treated as probability distribution function: γ depends on initial requirements and data specifications. PNC method leads to point extrapolation and interpolation via discrete set of fixed knots. Main features of PNC method are:

1. The smaller distance between knots the better;
2. PNC method develops a linear interpolation and extrapolation into other functions as probability distribution functions;
3. PNC is a generalization of MHR method via different nodes combinations;
4. Interpolation and extrapolation of L points is connected with the computational cost of rank $O(L)$ as in MHR method;
5. Nodes combination and coefficient γ are crucial in the process of data probabilistic anticipation and foreseeing.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Proposed method, called Probabilistic Nodes Combination (PNC), is the method of 2D curve interpolation and extrapolation using the set of key points (knots or nodes). Nodes can be treated as characteristic points of data for modeling and analyzing. The model of data can be built by choice of probability distribution function and nodes combination. PNC modeling via nodes combination and parameter γ as probability distribution function enables value anticipation in risk analysis and decision making. Two-dimensional curve is extrapolated and interpolated via nodes combination and different functions as discrete or continuous probability distribution functions: polynomial, sine, cosine, tangent, cotangent, logarithm, exponent, arc sin, arc cos, arc tan, arc cot or power function. Novelty of the chapter consists of two generalizations: generalization of previous MHR method with various nodes combinations and generalization of linear interpolation with different (no basic) probability distribution functions and nodes combinations. Computer vision needs suitable methods of shape representation and contour reconstruction. One of them, invented by the author and called method of Hurwitz-Radon Matrices (MHR), can be used in representation and reconstruction of shapes of the objects in the plane. Proposed method is based on a family of Hurwitz-Radon (HR) matrices. The matrices are skew-symmetric and possess columns composed of orthogonal vectors. 2D shape is represented by the set of successive nodes. It is shown how to create the orthogonal and discrete OHR operator and how to use it in a process of shape representation and reconstruction. Then MHR method is generalized to Probabilistic Nodes Combination (PNC) method. This work clarifies the significance and novelty of the proposed method compared to existing methods (for example polynomial interpolations and Bézier curves). Previous published papers of the author were dealing with the method of Hurwitz-Radon Matrices (MHR method). Novelty of this monograph and proposed method consists in the fact that calculations are free from the family of Hurwitz-Radon Matrices. Problem statement of this chapter is: how to reconstruct (interpolate) missing points of 2D curve having the set of interpolation nodes (key points) and using the information about probabilistic distribution of unknown points. For example the simplest basic distribution leads to the easiest interpolation – linear interpolation. Apart from probability distribution, additionally there is the second factor of proposed interpolation method: nodes combination. The simplest nodes combination is zero. Thus proposed curve modeling is based on two agents: probability distribution and nodes combination. Significance of this book consists in generalization for MHR method: the computations are done without matrices in curve fitting and shape modeling, with clear point interpolation formula based on probability distribution function (continuous or discrete) and nodes combination. This book also consists of generalization for linear interpolation with different (no basic) probability distribution functions and nodes combinations. So this chapter answers the question: "Why and when should we use PNC method?".

Curve interpolation represents one of the most important problems in mathematics and computer science: how to model the curve via discrete set of two-dimensional points? Also the matter of shape representation (as closed curve-contour) and curve parameterization is still opened. For example pattern recognition, signature verification or handwriting identification problems are based on curve modeling via the choice of key points. So interpolation is not only a pure mathematical problem but important task in computer vision and artificial intelligence. The book wants to approach a problem of curve modeling by characteristic points. Proposed method relies on nodes combination and functional modeling of curve points situated between the basic set of key points. The functions that are used in calculations represent whole family of elementary functions with inverse functions: polynomials, trigonometric, cyclometric, logarithmic, exponential and power function. These functions are treated as probability distribution functions in the range [0;1]. Significant problem in machine vision and computer vision is that of appropriate 2D shape representation and reconstruction. Classical discussion about shape representation is based on

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

the problem: contour versus skeleton. This monograph is voting for contour which forms boundary of the object. Contour of the object, represented by successive contour points, consists of information which allows us to describe many important features of the object as shape coefficients. 2D curve modeling and generation is a basic subject in many branches of industry and computer science, for example in the cad/cam software. The representation of shape can have a great impact on the accuracy and effectiveness of object recognition. In the literature, shape has been represented by many options including curves, graph-based algorithms and medial axis to enable shape-based object recognition. Digital 2D curve (open or closed) can be represented by chain code (Freeman's code). Chain code depends on selection of the started point and transformations of the object. So Freeman's code is one of the method how to describe and to find contour of the object. Analog (continuous) version of Freeman's code is the curve α-s. Another contour representation and reconstruction is based on Fourier coefficients calculated in discrete Fourier transformation (DFT). These coefficients are used to fix similarity of the contours with different sizes or directions. If we assume that contour is built from segments of a line and fragments of circles or ellipses, hough transformation is applied to detect contour lines. Also geometrical moments of the object are used during the process of object shape representation. Contour is also applied in shape decomposition. Many branches of medicine, industry and manufacturing are looking for methods connected with geometry of the contour.

Future works are going to: application of PNC method in decision making, choice and features of nodes combinations and coefficient γ, implementation of PNC in computer vision and artificial intelligence: shape geometry, contour modelling, object recognition and curve parameterization.

REFERENCES

Ballard, D. H. (1982). *Computer Vision*. New York: Prentice Hall.

Chapra, S. C. (2012). *Applied Numerical Methods*. McGraw-Hill.

Choraś, R. S. (2005). *Computer Vision*. Warsaw, Poland: Exit.

Cocozza-Thivent, C., Eymard, R., Mercier, S., & Roussignol, M. (2006). Characterization of the Marginal Distributions of Markov Processes Used in Dynamic Reliability. *Journal of Applied Mathematics and Stochastic Analysis*, 1–18.

Collins, G. W. II. (2003). *Fundamental Numerical Methods and Data Analysis*. Case Western Reserve University.

Dahlquist, G., & Bjoerck, A. (1974). *Numerical Methods*. New York: Prentice Hall.

Dejdumrong, N. (2007). A Shape Preserving Verification Techniques for Parametric Curves. *Computer Graphics, Imaging and Visualization*, *2007*, 163–168.

Dyn, N., Levin, D., & Gregory, J. A. (1987). A 4-Point Interpolatory Subdivision Scheme for Curve Design. *Computer Aided Geometric Design*, *4*(4), 257–268. doi:10.1016/0167-8396(87)90001-X

Jakóbczak, D. (2007). 2D and 3D Image Modeling Using Hurwitz-Radon Matrices. *Polish Journal of Environmental Studies*, *4A*(16), 104–107.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Jakóbczak, D. (2009). Curve Interpolation Using Hurwitz-Radon Matrices. *Polish Journal of Environmental Studies*, *3B*(18), 126–130.

Jakóbczak, D. (2010a). Shape Representation and Shape Coefficients via Method of Hurwitz-Radon Matrices. *Lecture Notes in Computer Science*, *6374*, 411–419. doi:10.1007/978-3-642-15910-7_47

Jakóbczak, D. (2010b). Object Modeling Using Method of Hurwitz-Radon Matrices of Rank k. In W. Wolski & M. Borawski (Eds.), *Computer Graphics: Selected Issues* (pp. 79–90). Szczecin, Poland: University of Szczecin Press.

Jakóbczak, D. (2011a). Curve Parameterization and Curvature via Method of Hurwitz-Radon Matrices. *Image Processing & Communications- International Journal (Toronto, Ont.)*, *1-2*(16), 49–56.

Jakóbczak, D. (2011b). Data Extrapolation and Decision Making via Method of Hurwitz-Radon Matrices. *Lecture Notes in Computer Science*, *6922*, 173–182. doi:10.1007/978-3-642-23935-9_17

Jakóbczak, D. (2011c). Curve Extrapolation and Data Analysis using the Method of Hurwitz-Radon Matrices. *Folia Oeconomica Stetinensia, 9*(17), 121-138.

Jakóbczak, D. (2013). Probabilistic Modeling of Signature using the Method of Hurwitz-Radon Matrices. *Global Perspectives on Artificial Intelligence*, *1*(1), 1–7.

Kozera, R. (2004). *Curve Modeling via Interpolation Based on Multidimensional Reduced Data*. Gliwice, Poland: Silesian University of Technology Press.

Liu, T., & Geiger, D. (1999). Approximate tree matching and shape similarity. *Int. Conf. Computer Vision,* Corfu, Greece.

Lorton, A., Fouladirad, M., & Grall, A. (2013). A Methodology for Probabilistic Model-based Prognosis. *European Journal of Operational Research*, *225*(3), 443–454. doi:10.1016/j.ejor.2012.10.025

Pergler, M., & Freeman, A. (2008). Probabilistic Modeling as an Exploratory Decision-Making Tool. *McKinsey Working Papers on Risk*, *6*, 1-18.

Ralston, A., & Rabinowitz, P. (2001). *A First Course in Numerical Analysis* (2nd ed.). New York: Dover Publications.

Rogers, D. F. (2001). *An Introduction to NURBS with Historical Perspective*. Morgan Kaufmann Publishers.

Saber, E., Xu, Y., & Murat Tekalp, A. (2005). Partial shape recognition by sub-matrix matching for partial matching guided image labeling. *Pattern Recognition*, *38*(10), 1560–1573. doi:10.1016/j.patcog.2005.03.027

Schumaker, L. L. (2007). *Spline Functions: Basic Theory*. Cambridge Mathematical Library. doi:10.1017/CBO9780511618994

Sebastian, T. B., Klein, P. N., & Kimia, B. B. (2003). On aligning curves. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*(1), 116–124. doi:10.1109/TPAMI.2003.1159951

Tadeusiewicz, R., & Flasiński, M. (1991). *Image Recognition*. Warsaw, Poland: PWN.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Zhang, D., & Lu, G. (2004). Review of Shape Representation and Description Techniques. *Pattern Recognition*, *1*(37), 1–19. doi:10.1016/j.patcog.2003.07.008

ADDITIONAL READING

Basu, S., & Bresler, Y. (2000). $O(N^2 \log_2 N)$ filtered backprojection reconstruction algorithm for tomography. *IEEE Transactions on Image Processing*, *9*(10), 1760–1773. doi:10.1109/83.869187 PMID:18262914

Brankov, J. G., Yang, Y., & Wernick, M. N. (2004). Tomographic image reconstruction based on a Content – Adaptive Mesh Model. *IEEE Transactions on Medical Imaging*, *2*(23), 202–212. doi:10.1109/TMI.2003.822822 PMID:14964565

Brasse, D., & Defrise, M. (2004). Fast fully 3-D image reconstruction in PET using planograms. *IEEE Transactions on Medical Imaging*, *4*(23), 413–425. doi:10.1109/TMI.2004.824231 PMID:15084067

Bulacu, M., & Schomaker, L. (2007). Text-independent writer identification and verification using textural and allographic features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(4), 701–717. doi:10.1109/TPAMI.2007.1009 PMID:17299226

Bulacu, M., Schomaker, L., & Brink, A. (2007). *Text-independent writer identification and verification on off-line Arabic handwriting*. In: *International Conference on Document Analysis and Recognition*. 769–773.

Cetin, M., Karl, W. C., & Willsky, A. S. (2002, September). *Edge – preserving image reconstruction for coherent imaging application.* Paper presented at the IEEE International Conference on Image Processsing, Rochester, NY, USA. 10.1109/ICIP.2002.1039992

Chen, J., Cheng, W., & Lopresti, D. (2011). Using perturbed handwriting to support writer identification in the presence of severe data constraints. In: Document Recognition and Retrieval. 1–10. doi:10.1117/12.876497

Chen, J., Lopresti, D., & Kavallieratou, E. (2010). *The impact of ruling lines on writer identification*. In: *International Conference on Frontiers in Handwriting Recognition*. 439–444. 10.1109/ICFHR.2010.75

Chlebus, E., & Cholewa, M. (1999). Rapid prototyping – rapid tooling. *CADCAM Forum, 11*, 23-28.

Cormen, T. H., Leiserson, C. E., & Rivest, R. L. (1996). *Introduction to algorithms*. Massachusetts, USA: the Massachusetts Institute of Technology Press and McGraw-Hill.

Defrise, M. (2001). A short readers guide to 3D tomographic reconstruction. *Computerized Medical Imaging and Graphics*, *25*(2), 113–116. doi:10.1016/S0895-6111(00)00061-6 PMID:11137787

Djeddi, C., & Souici-Meslati, L. (2010). *A texture based approach for Arabic writer identification and verification*. In: *International Conference on Machine and Web Intelligence*. 115–120.

Djeddi, C., & Souici-Meslati, L. (2011). *Artificial immune recognition system for Arabic writer identification*. In: *International Symposium on Innovation in Information and Communication Technology*. 159–165. 10.1109/ISIICT.2011.6149612

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dryja, M., Jankowska, J., & Jankowski, M. (1982). *Survey of numerical methods and algorithms. Part II*. Warsaw, Poland: WNT.

Eldar, Y. C. (2001). *Quantum Signal Processing*. (Unpublished doctoral dissertation). Massachusetts Institute of Technology, USA.

Eldar, Y. C., & Oppenheim, A. V. (2002). Quantum Signal Processing. *IEEE Signal Processing Magazine*, *6*(19), 12–32. doi:10.1109/MSP.2002.1043298

Fortuna, Z., Macukow, B., & Wąsowski, J. (1982). *Numerical methods*. Warsaw, Poland: WNT.

Galloway, M. M. (1975). Texture analysis using gray level run lengths. *Computer Graphics and Image Processing*, *4*(2), 172–179. doi:10.1016/S0146-664X(75)80008-6

Garain, U., & Paquet, T. (2009). *Off-line multi-script writer identification using AR coefficients*. In: *International Conference on Document Analysis and Recognition*. 991–995. 10.1109/ICDAR.2009.222

Ghiasi, G., & Safabakhsh, R. (2013). Offline text-independent writer identification using codebook and efficient code extraction methods. *Image and Vision Computing*, *31*(5), 379–391. doi:10.1016/j.imavis.2013.03.002

Jakóbczak, D. (2005). Hurwitz-Radon matrices and their children. *Computer Science*, *5*(8), 29–38.

Jakóbczak, D. J. (2014). *2D Curve Modeling via the Method of Probabilistic Nodes Combination-Shape Representation, Object Modeling and Curve Interpolation-Extrapolation with the Applications*. Saarbrucken: LAP Lambert Academic Publishing.

Jankowska, J., & Jankowski, M. (1981). *Survey of numerical methods and algorithms. Part I*. Warsaw, Poland: WNT.

Kontaxakis, G., & Strauss, L. G. (1998). Maximum likelihood algorithms for image reconstruction in Positron Emission Tomography. *Radionuclides for Oncology – Current Status and Future Aspects, 1998,* 73-106.

Kowalczuk, Z., & Wiszniewski, B. (Eds.). (2007). *Intelligent data mining in diagnostic purposes: Automatics and informatics*. Gdansk, Poland: PWNT.

Kundur, D., & Hatzinakos, D. (1998). A novel blind deconvolution scheme for image restoration using recursive filtering. *IEEE Transactions on Signal Processing*, *2*(46), 375–390. doi:10.1109/78.655423

Laine, A., & Zong, X. (1996). *Border identification of echocardiograms via multiscale edge detection and shape modeling*. Paper presented at the IEEE International Conference on Image Processsing, Lausanne, Switzerland. 10.1109/ICIP.1996.560486

Lang, S. (1970). *Algebra*. Reading, Massachusetts, USA: Addison-Wesley Publishing Company.

Le Buhan Jordan, C., Bossen, F., & Ebrahimi, T. (1997). *Scalable shape representation for content based visual data compression*. Paper presented at the International Conference on Image Processing, Santa Barbara, CA, USA. 10.1109/ICIP.1997.647962

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Marker, J., Braude, I., Museth, K., & Breen, D. (2006). Contour-based surface reconstruction using implicit curve fitting, and distance field filtering and interpolation. *Volume Graphics*, *2006*, 1–9.

Marti, U.-V., & Bunke, H. (2002). The IAM-database: An English sentence database for offline handwriting recognition. *Int. J. Doc. Anal. Recognit.*, *5*(1), 39–46. doi:10.1007100320200071

Meyer, Y. (1993). *Wavelets: algorithms & applications*. Philadelphia, USA: Society for Industrial and Applied Mathematics.

Nosary, A., Heutte, L., & Paquet, T. (2004). Unsupervised writer adaption applied to handwritten text recognition. *Pattern Recognition Letters*, *37*(2), 385–388. doi:10.1016/S0031-3203(03)00185-7

Ozaki, M., Adachi, Y., & Ishii, N. (2006). *Examination of effects of character size on accuracy of writer recognition by new local arc method*. In: *International Conference on Knowledge-Based Intelligent Information and Engineering Systems*. 1170–1175. 10.1007/11893004_148

Poggio, T., & Smale, S. (2003). The mathematics of learning: Dealing with data. *Notices of the American Mathematical Society*, *5*(50), 537–544.

Przelaskowski, A. (2005). *Data compression*. Warsaw, Poland: BTC.

Rutkowski, L., Siekmann, J., Tadeusiewicz, R., & Zadeh, A. (Eds.). (2004). *Lecture notes on artificial intelligence: Artificial intelligence and soft computing*. Berlin-Heidelberg, Germany: Springer-Verlag.

Schlapbach, A., & Bunke, H. (2004). *Using HMM based recognizers for writer identification and verification. 9th Int. Workshop on Frontiers in Handwriting Recognition*. 167–172. 10.1109/IWFHR.2004.107

Schlapbach, A., & Bunke, H. (2006). *Off-line writer identification using Gaussian mixture models*. In: *International Conference on Pattern Recognition*. 992–995.

Schlapbach, A., & Bunke, H. (2007). A writer identification and verification system using HMM based recognizers. *Pattern Analysis & Applications*, *10*(1), 33–43. doi:10.100710044-006-0047-5

Schomaker, L., Franke, K., & Bulacu, M. (2007). Using codebooks of fragmented connected-component contours in forensic and historic writer identification. *Pattern Recognition Letters*, *28*(6), 719–727. doi:10.1016/j.patrec.2006.08.005

Shahabinejad, F., & Rahmati, M. (2007). A new method for writer identification and verification based on Farsi/Arabic handwritten texts, *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*. 829–833.

Siddiqi, I., Cloppet, F., & Vincent, N. (2009). Contour based features for the classification of ancient manuscripts. In: Conference of the International Graphonomics Society. 226–229.

Siddiqi, I., & Vincent, N. (2010). Text independent writer recognition using redundant writing patterns with contour-based orientation and curvature features. *Pattern Recognition Letters*, *43*(11), 3853–3865. doi:10.1016/j.patcog.2010.05.019

Vakhania, N. (1993). Orthogonal random vectors and the Hurwitz – Radon-Eckmann theorem. *Proc. of the Georgian Academy of Sciences-Mathematics, 1(1)*, 109-125.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Van, E. M., Vuurpijl, L., Franke, K., & Schomaker, L. (2005). The WANDA measurement tool for forensic document examination. *J. Forensic Doc. Exam.*, *16*, 103–118.

Willis, M. (2000). *Algebraic reconstruction algorithms for remote sensing image enhancement*. Unpublished doctoral dissertation, Department of Electrical and Computer Engineering, Brigham Young University.

Xu, Fang, & Mueller, K. (2005). Accelerating popular tomographic reconstruction algorithms on commodity PC graphics hardware. *IEEE Transactions on Nuclear Science*, *3*(52), 654–661.

Zaletelj, J., & Tasic, J. F. (2003). *Optimization and tracking of polygon vertices for shape coding*. Berlin-Heidelberg, Germany: Springer-Verlag. doi:10.1007/978-3-540-45179-2_52

Zhang, J. K., Davidson, T., & Wong, K. M. (2004). Efficient design of orthonormal wavelet bases for signal representation. *IEEE Transactions on Signal Processing*, *7*(52), 1983–1996. doi:10.1109/TSP.2004.828923

KEY TERMS AND DEFINITIONS

Artificial Intelligence: Intelligence of machines and computers, as a connection of algorithms and hardware, which makes that a man – human being can be simulated by the machines in analyzing risk, decision making, reasoning, knowledge, planning, learning, communication, perception and the ability to move and manipulate objects.

Curve Interpolation: Computing new and unknown points of a curve and creating a graph of a curve using existing data points – interpolation nodes.

Data Extrapolation: Calculation of unknown values for the points situated outside the ranges of nodes.

Hurwitz – Radon Matrices: A family of skew – symmetric and orthogonal matrices with columns and rows that create, together with identical matrix, the base in vector spaces of dimensions $N = 2, 4$ or 8.

MHR Method: The method of curve interpolation and extrapolation using linear (convex) combinations of OHR operators.

OHR Operator: Matrix operator of Hurwitz – Radon built from coordinates of interpolation nodes.

Value Anticipation: Foreseeing next value when last value is known.

This research was previously published in Probabilistic Nodes Combination (PNC) for Object Modeling and Contour Reconstruction edited by Dariusz Jacek Jakóbczak; pages 44-86; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 51
Can Computers Create Comics and Animations?

Miki Ueno
Toyohashi University of Technology, Japan

Kiyohito Fukuda
Osaka Prefecture University, Japan

Naoki Mori
Osaka Prefecture University, Japan

ABSTRACT

The objective of this chapter is to develop a method for generating and analyzing creative work by using computers. In this chapter, novels, comics, and animations are considered representative creative work because the main method for representing creative work is the use of natural language and pictures. These works focus on interesting issues, which can be explained by computational approaches to narratology. In this chapter, the authors discuss the following three topics. First, a method of semi-automatic story generation by agent-based simulation is presented. Second, methods of generating and analyzing comics on the basis of the features of pictures and stories used are described. Third, the authors introduce a method of estimating the difficulties of translating a storyboard into original pictures in the process of creating animations.

INTRODUCTION

How do humans produce creative work? What do individuals imagine, and what do they think of during the creative process? The ultimate purpose of this study is to find an algorithm for the generation of creative, intellectual works, such as novels, comics, and animations, with regard to artificial intelligence. The creative works listed above involve stories and are constructed using sequential components. Further, the popular representation of a narrative utilizes natural language and pictures. For example, novels are constructed using natural language, and this type of language is a very useful means of logically depicting

DOI: 10.4018/978-1-7998-0951-7.ch051

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

a given situation. However, the number of sentences comprising individual scenes varies considerably, and it is therefore difficult for computers to determine scene boundaries. In addition, comics are constructed using both pictures and natural language, and animations are comprised of pictures and sound based on natural language. Note that pictures constitute an effective means of creating stories, because they can convey postures and positions more directly than verbal or written language. This approach allows stories to become comprehensible to a wider audience, regardless of age or nationality.

In Japan, comics and animations are common in popular culture. A key characteristic of these representations are the deformation or exaggeration of characters and other objects in the story. These visual representations contain lyrical and descriptive aspects to depict each situation clearly. Further, these works are often very intellectual, and some popular authors have produced exceptional creative pieces.

Story creation is something of which the vast majority of humans are capable; however, computers can neither create nor understand stories. The authors are very interested in the process of story construction, particularly as regards the development of appropriate models to facilitate computer story creation. Therefore, our goal is to define complete models of story creation for use by computers. However, it is difficult to define models for the entire process and every objective, therefore the authors suggest models for a specific process of creation. Hence, the authors have conducted several studies regarding computational creative stories, focusing on the topics listed below.

- **Story Generation:** Semi-automatic story generation using log data. Continuous transitions are automatically created and are given to writers in order to write unexpected stories that cannot be created by humans only. The aim of this section is to propose a computational method of coherent and unexpected story creation.
- **Comic Generation:** Automatic comic generation focusing on the relationship between transitions of stories and expression of pictures and comic analysis based on two of their features. Continuous transitions expressed by pictures are modeled and the relations between characteristic expressions and the patterns of stories are analyzed. The aim of this section is to suggest computational models for comics focusing on the expressions in pictures.
- **Animation Generation:** Feature vector definition to provide storyboard information and to determine the difficulties involved in scene translation from storyboards to equivalent key animations. The difficulties in transitioning storyboards into original pictures are estimated. The aim of this section is to appropriately assign each scene to each creator of original pictures according to the index of difficulty by a computer functioning like a facilitator.

In this chapter, these studies are introduced using examples, in order to demonstrate the fundamental techniques of computational narratology and the challenges involved in the quantification of story creation.

BACKGROUND

Recently, automatic story generation and comprehension has attracted considerable interest as a challenging problem in the fields of artificial intelligence. The types of stories are broadly classified into two categories. One is utilizing natural language, the other is utilizing pictures. Novels are written in natural language, while comics and animations are described mainly by pictures. Story information is

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

common, but expression information is specific. Firstly, the background of story generation with natural language is described. With a view towards realizing effective automatic story generation, a number of research papers on automatic story generation based on Case-Based Reasoning (CBR) (Gervás, Diaz-Agudo, Peinado, & Hervás, 2005; Pérez y Pérez & Sharples, 2001; Tuner, 1994) and the problem-solving process (Meehan, 1980), the project on automatic composition of works like Shinichi Hoshi, who was a famous Japanese short story writer (Kouyama & Sumi, 2014; Matsubara et al., 2013; Ogata, Sato, & Komatani, 2014) and some studies on automatic story generation system based on well-known Propp theory (Imabuchi & Ogata, 2012) have been conducted. However, a problem has arisen in that these previous studies have only succeeded in generating lower quality novels rather than original, higher quality ones. The present authors believe that this problem can be attributed to automatic story generation research conducted based on the approach that "a human generates novels via computer"; this means that the computer must reuse original sentences and stories directly.

Secondly, the background of picture stories. Compared with the text stories, there exist only a few studies focused on picture stories such as comics and animation. Pictures are composed of lines and shapes, respectively. Although all of these components are simply collections of dots from a local perspective (Eitz, Hays, & Alexa, 2012), humans perceive them as characters, figures, and symbols based on semantics. However, the semantic comprehension method utilized by humans is unclear. In the field of image processing, the methods of automatically generating in-betweening pictures (Gohara, Sugimoto, & Morishima, 2010) have been proposed. However, limited research has been conducted on computational representations of the process of animation creation.

STORY GENERATION

In this section, the authors aim to propose a new automatic novel generation technique that does not utilize previously created original novels. This new approach essentially means that "the computer generates novels," that is, the computer can generate coherent and unexpected stories for humans. As a first step towards realizing this goal, the authors have proposed a novel semi-automatic story generation method that utilizes log data of agent-based simulation (ABS) (Fukuda, Mori, & Matsumoto, 2015). Here, ABS is a simulation to investigate the actions and interactions between autonomous agents (Macal & North, 2010). The proposed method is a "middle" approach, where "a human generates novels in cooperation with the computer," i.e., novel generation is conducted via interaction between a human and a computer. In this section, the proposed method is introduced, and its effectiveness and the related difficulties are discussed.

Proposed Story Generation Method

In this section, the abovementioned semi-automatic story generation method using log data of the ABS is described. This is based on the idea that a human sometimes writes stories and novels based on memories about interactions between humans. To provide a computer with that memory, the authors believe that log data can function as human memory. The proposed method can utilize any log data, such as the results of a numerical simulation or application usage history. However, the log data of the ABS is used here, because of its greater affinity. The proposed method proceeds as follows.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Step 1: Setting of the ABS. Some parameters of the ABS are determined by the user.
Step 2: Log Data Generation. Two types of log data are generated via the ABS: Status and action logs.
Step 3: Event Generation. Events are generated from the action logs. Events are important action logs to generate the story.
Step 4: Story Vector Generation. Story vectors are generated from the events and status logs. Story vectors are basic information represented in vector form that is required to generate the story.
Step 5: Story Generation. Stories are generated using story vectors through interaction between the user and computer.

Steps 2−4 of the proposed method can be performed automatically. However, steps 1 and 5 cannot be performed by the computer alone. Thus, the former are referred to as the "computer part" and the latter are referred to as the "user part."

Setting of the ABS (Step 1)

The ABS algorithms are implemented for the proposed method. In this ABS, agents have two internal values (friendship and emotion). In moving toward the sub-goal in the field, agents interact with each other if they are located close to one another. Interactions may change their internal values. The ABS is halted if the friendship value of the agent exceeds a threshold and the agent feels love or affection towards the other agent. The details of the algorithms are as follows. Here, a random number in a normal distribution is represented by $N(\mu,\sigma^2)$, where μ is the mean and σ^2 is the variance.

1. The field size is $x_{\max} \times y_{\max}$, the maximum number of turns is $T_{\max}$, the number of player agents is N_p, the number of roles is N_r, and the action number is N_a. Here, turn is a parameter that defines order of each interaction.
2. The environment agent is created, which has scene name s_{xy} at position (x, y) and environment values b_{xy} and $\beta_{xy,t}$. The environment values represent the influence of the environment on the changes of state at each position. Here, b_{xy} is 0 or 1 binary, $\beta_{xy,t}$ is a real number of [-15, 15], and the default of $\beta_{xy,0}$ is defined as [-10, 10].
3. N_p player agents are created. Each player agent has a name, (x, y) point, sex m, role in the story r, friendship values (f-values) with each player agent $F_t = \left\{f_{1,t}, f_{2,t}, ..., f_{N_p,t}\right\}$, love flags with each player agent $L = \left\{\ell_1, \ell_2, ..., \ell_{N_p}\right\}$, and an emotion value (e-value) $E_t = \left\{e_{1,t}, e_{2,t}\right\}$. Here, $f \in F$ is a real number of [0, 100], $e \in E$ is a real number of [-100, 100], m and $\ell \in L$ are 0 or 1 binary, and r is a natural number of $(0, N_r]$. In addition, the default of f_0 is defined as [25, 75] and the default of e_0 is defined as [-25, 25]. Finally, a set of player agents is defined as P.
4. Then, the turn t is set to 0. The set of player agents who do not take action is defined as P_t and the set of player agents who have already taken some action is defined as P_t'.
5. P_t and P_t' are then updated as $P_t = P, P_t' = \varnothing$.
6. A player agent i is randomly selected from P_t. P_t and P_t' are updated as $P_t = P_t / i$, $P_t' = P_t' \cup i$, respectively.
7. i moves to certain Moore neighborhoods as it traverses the distance between the current and target positions. The next target position of i is randomly set when i reaches its existing target.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

8. A player agent of non- *i*, labeled *j*, is randomly selected. In addition, the friendship value of *i* for *j*, represented by $f_{j,t}^{i}$ and the emotion value of *i* e_{t}^{i} are updated as follows. This update represents psychological influences, such that these values converge on a constant value to indicate that *i* loses interest in *j* once a given time has elapsed. Then, T_{fri} and T_{emo} are used to represent the time-elapsed rates of the friendship and emotion values, respectively.

$$f_{j,t}^{i} = f_{j,t}^{i} - \frac{\left|N\left(0,1\right)\right| \times b_{xy}^{i}\left(f_{j,t}^{i} - 50\right)}{T_{fri}} \tag{1}$$

s.t. $\{j \mid j \neq i, j \in O\}$

$$e_{t}^{i} = e_{t}^{i} - \frac{\left|N\left(0,1\right)\right| b_{xy}^{i} e_{t}^{i}}{T_{emo}} \tag{2}$$

9. If the other player agents exist in Neumann neighborhoods of *i*, a player agent *k* is randomly selected from amongst these agents. If no player agents exist, the process moves on to 12.
10. In order to represent the interaction between two player agents, the friendship value of *k* towards *i*, here labeled $f_{i,t}^{k}$, and the emotion values of *k*, $e_{1,t}^{k}$, $e_{2,t}^{k}$ as follows. The action log for the interaction is saved and N_a is updated to $N_a = N_a + 1$. Now, I_{fri} and I_{emo} represent the friendship and emotion value interaction rates, respectively.

$$f_{i,t+1}^{k} = f_{i,t}^{k} + \Delta f_{i,t}^{k} \tag{3}$$

$$e_{1,t+1}^{k} = e_{1,t}^{k} + \Delta e_{1,t}^{k} \tag{4}$$

$$e_{2,t+1}^{k} = e_{2,t}^{k} + \Delta e_{2,t}^{k} \tag{5}$$

$$\Delta f_{i,t}^{k} = \left|N\left(0,1\right)\right| \times b_{xy}^{k} d_{i,t}^{k} + \beta_{xy,t}^{k} \tag{6}$$

$$d_{i,t}^{k} = \frac{2f_{i,t}^{k} + f_{k,t}^{i}}{I_{fri}} + \frac{e_{1,t}^{k}\left(e_{2,t}^{k} + 100\right)}{I_{emo}} \tag{7}$$

$$\Delta e_{1,t}^{k} = \left|N\left(0,1.5^{2}\right)\right| \times \Delta f_{i,t}^{k} \tag{8}$$

$$\Delta e_{2,t}^{k} = N\left(0,1.5^{2}\right) \times \left|\Delta f_{i,t}^{k}\right| \tag{9}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

11. If $\Delta f_{i,t}^{k}$ or $\Delta f_{k,t}^{i}$ exceeds a threshold value *h*, the love flag of *k* or *i* for *i* or *k*, which is ℓ_{i}^{k} or ℓ_{k}^{i}, is changed to another binary value.
12. If $P_t \neq \varnothing$, the process returns to 6.
13. A player agent *p* is randomly selected from P_t'. If the other player agents exist in Neumann neighborhoods of *p*, a player agent p' is randomly selected from among them. If $\ell_{p'}^{p}$ is 1, *p* confesses his or her love for p' with probability *R*. This confession succeeds with a rate of R_{suc}. If *p* confesses, the confession is saved in the action log, regardless of its success or failure. If *p* does not confess, P_t' is updated to $P_t' = P_t' / p$. *R* and R_{suc} are defined as follows.

$$R = \frac{\left| f_{p',t}^{p} - 50 \right|}{100} \tag{10}$$

$$R_{suc} = \begin{cases} \dfrac{0.8 f_{p',t}^{p} + 0.2\alpha}{200} & \left(m^{p} = m^{p'} \right) \\ \dfrac{0.8 f_{p',t}^{p} + 0.2\alpha}{100} & \left(m^{p} \neq m^{p'} \right) \end{cases} \tag{11}$$

$$\alpha = \frac{e_{1,t}^{p'} \times e_{2,t}^{p'}}{100} \tag{12}$$

14. If $P_t' \neq \varnothing$, the process returns to 13.
15. The value of $\beta_{xy,t}$ always varies over time. Hence, $\beta_{xy,t}$ is updated as follows. Here, a_{env} represents the environment variation cycle.

$$\beta_{xy,t+1} = \beta_{xy,0} + 5 \sin\left(\frac{2\left(t+1\right)\pi}{a_{env}} \right) \tag{13}$$

16. Finally, *t* is updated to $t=t+1$ and the status log is saved. If $t \geq t_{\max}$, the ABS is halted. If $t < t_{\max}$, the process returns to 5.

Log Data Generation (Step 2)

The log data is generated using the ABS, which is implemented for the proposed method, along with the two types of log data mentioned above: Status and action logs. The details of these log types are given below.

- **Action Log:** Action log information is saved if agents perform a specific action in the ABS. The turn, scene name, main agent name, target agent name, action type, action number, the degree of change in the friendship value caused by the action, and the degree of change in the emotion value caused by the action in the action log are saved. Table 1 shows the example of the action logs.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. The example of the action logs

Turn	Main Agent	Target Agent	Action Type (Action Number)	Position	The Change of F-Value	The Change of E-Value 1	The Change of E-Value 2
178	0	2	INTERACTION (0)	7,64,lectureB	-0.6626	-1.6809	-0.0026
179	0	2	INTERACTION (1)	8,64,lectureB	-1.2161	-0.4531	0.6638
323	2	1	INTERACTION (2)	50,111,lectureA	1.0534	1.3433	-0.5078
422	1	0	INTERACTION (3)	49,101,clubroom	3.0015	11.004	-5.5396
422	0	1	INTERACTION (4)	48,101,clubroom	10.663	13.993	-23.709
431	2	0	INTERACTION (5)	48,109,lectureA	10.497	15.894	-17.420
432	2	0	INTERACTION (6)	49,110,lectureA	-6.6434	-5.2733	-10.923
432	0	2	INTERACTION (7)	48,110,lectureA	3.3707	3.5269	-6.1631
467	1	0	INTERACTION (8)	46,113,lectureA	13.929	20.447	16.303
484	0	1	INTERACTION (9)	57,112,lectureA	16.168	28.519	-19.999
484	1	0	INTERACTION (10)	58,112,lectureA	-26.625	-47.057	-55.744
484	1	0	SUCCESS (11)	58,112,lectureA	100	100	100

Table 2. The example of the status logs

Turn	Name	Id	Position	Sex	Role	F-Value to 1	F-Value to 2	F-Value to 3	E-Value 1	E-Value 2
0	Hiro	1	90,39,home	1	1	72	45	33	-5	-4
0	Kana	2	42,44,home	2	2	47	62	50	13	-8
0	Yuu	3	67,72,home	1	3	37	64	60	-13	8
1	Hiro	1	89,40,way	1	1	71.883	45.002	33.065	-4.9897	-3.9986
1	Yuu	3	66,73,way	1	3	36.955	64.053	60.009	-13.053	8.0127
1	Kana	2	43,44,way	2	2	46.985	62.027	50	13.025	-8.0139
2	Yuu	3	65,73,way	1	3	36.986	64.050	59.981	-13.009	7.9878
2	Hiro	1	88,41,way	1	1	71.845	45.018	33.145	-4.9727	-3.9928
2	Kana	2	43,45,way	2	2	46.980	62.028	50	13.043	-8.0200

- **Status Log:** The status logs are saved at the end of each turn of the ABS. These comprise the turn information and all the parameters of all the agents. Table 2 shows the example of the status logs.

Event Generation (Step 3)

In this step, the direct use of action logs to generate story vectors tends to yield an unnatural story. Therefore, events are generated through the compression of action logs. The event generation algorithms are listed below. Table 3 shows the example of the events.

1. The set of action logs obtained via the ABS are defined as *G*.
2. Action logs within *G* that have the same scene names are defined as *G*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

3. G is divided into sets of action logs with identical main agent names and action types. Furthermore, a new action log is generated from the set of action logs obtained from G. The values of t, s_{xy}, and N_a are defined as those corresponding to the action log with the minimum t, and the degree of change in the F_t and E_t values are defined as the sum of all action logs in the set. G is updated as $G = G \,/\, G'$.
4. If $G \neq \varnothing$, the method returns to 2.
5. The new action logs generated in 3. are defined as the events.

Story Vector Generation (Step 4)

Story generation using events directly is difficult. Hence, stories are generated using the story vectors s, which are in turn generated by the events. The story vector generation algorithms are the following. Table 4 shows the example of the story vectors.

1. The main agent and the target agent are defined as i and j, respectively. The values of the turn t, the action number N_a, the position of j ($x_{j,t}$, $y_{j,t}$), the degree of change in the friendship value Δf_t and degree of change in the emotion value Δe_t are obtained from the event. Furthermore, m^i, m^j, r^i and r^j are obtained, along with the friendship values of i and j for j and i, which are $f_{j,t}^i$ and $f_{i,t}^j$, respectively, and the emotion values of i and j, e_t^i and e_t^j, respectively, from the status logs.
2. The s is defined as follows.

$$s = \left(t, N_a, m^i, m^j, r^i, r^j, x_{j,t}, y_{j,t}, \Delta f_t, f_{j,t}^i, f_{i,t}^j, e_{1,t}^i, e_{2,t}^i, e_{1,t}^j, e_{2,t}^j, \Delta e_{1,t}, \Delta e_{2,t}\right) \tag{14}$$

Table 3. The example of the events

Turn	Main Agent	Target Agent	Action Type (Action Number)	Position	The Change of F-Value	The Change of E-Value 1	The Change of E-Value 2
178	0	2	INTERACTION (0, 1)	lectureB	-0.6626	-1.6809	-0.0026
323	2	1	INTERACTION (2)	lectureA	1.0534	1.3433	-0.5078
422	1	0	INTERACTION (3)	clubroom	3.0015	11.004	-5.5396
422	0	1	INTERACTION (4)	clubroom	10.663	13.993	-23.709
431	2	0	INTERACTION (5, 6)	lectureA	10.497	15.894	-17.420
432	0	2	INTERACTION (7)	lectureA	3.3707	3.5269	-6.1631
467	1	0	INTERACTION (8)	lectureA	13.929	20.447	16.303
484	0	1	INTERACTION (9)	lectureA	16.168	28.519	-19.999
484	1	0	INTERACTION (10)	lectureA	-26.625	-47.057	-55.744
484	1	0	SUCCESS (11)	lectureA	100	100	100

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. The example of the story vectors

Turn	Main Agent (*i*)	Target Agent (*j*)	Scene Name	The Change of F-Value	F-Value of *i*	F-Value of *j*	The Change of Emotion of *i*	The Change of Emotion of *j*
178	Hiro, Man, Hero	Yuu, Man, Rival	lectureB	becomes slightly hostile (-1.88)	normal (35.32)	normal (38.74)	sadness → sadness	anger → anger
323	Yuu, Man, Rival	Kana, Woman, Heroine	lectureA	becomes slightly friendly (1.05)	normal (59.51)	normal (50.0)	anger → anger	fun → fun
422	Kana, Woman, Heroine	Hiro, Man, Hero	clubroom	becomes slightly friendly (3.0)	normal (48.2)	normal (46.59)	fun → fun	sadness → fun
422	Hiro, Man, Hero	Kana, Woman, Heroine	clubroom	becomes slightly friendly (10.66)	normal (46.59)	normal (48.2)	sadness → sadness	fun → fun
431	Yuu, Man, Rival	Hiro, Man, Hero	lectureA	becomes slightly friendly (3.85)	normal (40.17)	normal (38.92)	anger → anger	fun → fun
432	Hiro, Man, Hero	Yuu, Man, Rival	lectureA	becomes slightly friendly (3.37)	normal (49.43)	normal (40.22)	fun → fun	anger → anger
467	Kana, Woman, Heroine	Hiro, Man, Hero	lectureA	becomes slightly friendly (13.93)	normal (58.58)	normal (49.61)	fun → fun	fun → fun
484	Hiro, Man, Hero	Kana, Woman, Heroine	lectureA	becomes slightly friendly (16.17)	normal (63.2)	normal (58.4)	fun → fun	fun → fun
484	Kana, Woman, Heroine	Hiro, Man, Hero	lectureA	becomes more friendly (-26.62)	normal (58.4)	normal (63.2)	fun → fun	fun → sadness
484	Kana, Woman, Heroine	Hiro, Man, Hero	lectureA	accepts asking out	normal (58.4)	normal (63.2)	fun → fun	fun → joy

Story Generation (Step 5)

In this step, story generation requires six elements: "when," "where," "who (main)," "who (target)," "action," and "result". Five elements are generated using *s*. The "action" element is excluded as it is generated via the interaction between the user and the computer. The story generation algorithms are as follows.

1. "When" is defined as *t* in *s*.
2. "Who (main)" is defined as the name of main player agent *i*, having m^i and r^i in *s*. "Who (target)" is defined as the name of target player agent *j*, having m^j and r^j in *s*.
3. "Where" is defined as the scene name $s_{x_{j,t}y_{j,t}}$, based on the position ($x_{j,t}$, $y_{j,t}$) in *s*.
4. 4. The change in the friendship value caused by the action is defined as w_f. Note that w_f is defined as follows.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$w_f = \begin{cases} "becomes\ more\ friendly" \left(\Delta f_t > 20\right) \\ "becomes\ slightly\ friendly" \left(0 < \Delta f_t \leq 20\right) \\ "becomes\ slightly\ hostile" \left(-20 \leq \Delta f_t < 0\right) \\ "becomes\ more\ hostile" \left(\Delta f_t < -20\right) \end{cases} \tag{15}$$

5. The friendship values at the action $w^i_{f,t}$ are obtained, along with and the change in the emotion values caused by the action " $w^i_{e,t}$, " $w^j_{e,t}$. Here, $w^i_{f,t}$, $w^j_{f,t}$, " $w^i_{e,t}$, and " $w^j_{e,t}$ are defined as follows. Now, if the emotion values of *i* and *j* are equal to 0, $w'_{e^k,t}$ is defined as "---" in order to indicate that the player agents have no emotion.

$$w_{f,t} = \begin{cases} "high" \left(f_t \geq 70\right) \\ "normal" \left(30 < f_t < 70\right) \\ "low" \left(f_t \leq 30\right) \end{cases} \tag{16}$$

$$\Delta w^i_{e,t} = w_{e^i,t} \rightarrow w_{e^i,t} \tag{17}$$

$$\Delta w^j_{e,t} = w_{e^j,t} \rightarrow w'_{e^j,t} \tag{18}$$

$$w_{e^k,t} = \begin{cases} "joy" \left(e^k_{1,t} > 0, e^k_{2,t} > 0\right) \\ "anger" \left(e^k_{1,t} < 0, e^k_{2,t} > 0\right) \\ "sadness" \left(e^k_{1,t} < 0, e^k_{2,t} < 0\right) \\ "fun" \left(e^k_{1,t} > 0, e^k_{2,t} < 0\right) \\ "- - -" \left(otherwise\right) \end{cases} \tag{19}$$

$$w'_{e^k,t} = \begin{cases} "joy" \left(e^k_{1,t} + \Delta e_{1,t} > 0, e^k_{2,t} + \Delta e_{2,t} > 0\right) \\ "anger" \left(e^k_{1,t} + \Delta e_{1,t} < 0, e^k_{2,t} + \Delta e_{2,t} > 0\right) \\ "sadness" \left(e^k_{1,t} + \Delta e_{1,t} < 0, e^k_{2,t} + \Delta e_{2,t} < 0\right) \\ "fun" \left(e^k_{1,t} + \Delta e_{1,t} > 0, e^k_{2,t} + \Delta e_{2,t} < 0\right) \\ "- - -" \left(otherwise\right) \end{cases} \tag{20}$$

6. The proposed method provides the user with the five elements defined according to 1. - 4., the friendship values at the action, and the change of emotion values caused by the action. The user then generates an "action" based on this information. The user selects the action from options that have been prepared in advance.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Experimental conditions of generation of story

(x_{max}, y_{max})	**(120, 120)**
N_p	3
N_r	3
Agent Name	Hiro, Kana, Yuu
Sex	Man, Woman, Man
Role	Hero, Heroine, Rival
Parameters	t_{max}=6000, h=30, T_{fri}=200, T_{emo}=500, I_{fri}=15, I_{emo}=500, a_{env}=200

Story Generation Experiments

To examine the effectiveness of the proposed method and any associated difficulties, user experiments were conducted. Table 5 shows the experimental conditions. In these experiments, five users (Users A−E), all of whom had written novels, made use of the proposed method. These users were aged between 20 and 29. The user-experiment procedure was as follows:

1. The proposed method was used to generate the story vector in the computer part.
2. The users generated stories utilizing the story vector generated in 1. After the story generation stage, the users evaluated the story generation difficulty on a scale of 1 (easy)−5 (difficult). In addition, the users' views on the proposed method were obtained via interview.

Story Generation Experiment Results

Figure 1 shows the example of the stories generated by the user experiment and Table 6 provides details of the evaluation about the difficulty of story generation evaluation. From Figure 1, Kana asks Hiro out after cutting class taking with him in a part of turn 485. It is difficult for humans to create the story like this. In addition, from Table 6, it is apparent that the mean of the story generation difficulty was 3.6. These results show that the proposed method allows users to create stories that cannot be created by only humans and suggest that the proposed method can create coherent and unexpected stories by automating the user part. In addition, the users stated that the range of action options was limited. Therefore, a range of actions must be prepared in order to overcome the lack of action versatility and create various stories automatically. Furthermore, the users stated that story planning was easy, because the character settings were unknown. This evaluation shows that it is necessary to provide detailed character settings automatically to realize applications in which "the computer generates novels."

Story Generation Conclusion

In this section, a proposed semi-automatic story generation method that utilizes log data of the ABS was discussed as a first step towards automatic unexpected story generation for a new automatic novel generation. The algorithms of the proposed method were explained and its efficacy and associated difficulties were discussed. The following future developments are necessary to further refine this method.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Experimental results of generation of story

Turn179: Hiro ignores Yuu in lecture room B. As a result, Yuu becomes slightly hostile towards Hiro.
Turn323: Yuu shows his textbook to Kana in lecture room A. As a result, Kana becomes slightly friendly towards Yuu.
Turn423: Kana talks with Hiro in Clubroom. As a result, Hiro becomes slightly friendly towards Kana.
Turn423: Hiro helps Kana with her task in Clubroom. As a result, Kana becomes slightly friendly towards Hiro.
Turn432: Yuu gives his loose-leaf paper to Hiro in lecture room A. As a result, Hiro becomes slightly friendly towards Yuu.
Turn433: Hiro falls asleep next to Yuu in lecture room A. As a result, Yuu becomes slightly friendly towards Hiro.
Turn467: Kana studies with Hiro in lecture room A. As a result, Hiro becomes slightly friendly towards Kana.
Turn484: Hiro studies with Kana in lecture room A. As a result, Kana becomes slightly friendly towards Hiro.
Turn485: Kana cuts class taking with Hiro in lecture room A. As a result, Hiro becomes more hostile towards Kana.
Turn485: Kana asks Hiro out in lecture room A. Hiro accepts.

Table 6. Evaluation about the difficulty of story generation

User Name	Difficulty of Story Generation
A	4
B	4
C	4
D	2
E	4
Mean	3.6

- User-selection model construction, which would model users' action choices.
- Development of a control method that would manage the logical coherence and unpredictability of the story.
- Automatic plot generation needed to generate attractive novels. Plots will be generated by rearranging the story.

AUTOMATIC COMIC GENERATION

In the previous section, the method of automatically generating stories by natural language was discussed. Here, the authors have focused on creative expressions using pictures, which can be understood by humans regardless of their age and nationality. Therefore, the study discussed in this section reports on computer creation and comprehension of comics whose stories are represented by pictures. There have been studies reporting comic creation by combining prepared datasets (Sumi, Sakamoto, Nakao, Mase, & Kunifuji, 2002). Stories represented by pictures involve large numbers of possible transition and expressions of pictures so that it is difficult to define models of pictures for every story. To solve this problem, the authors focused on four-scene comics. It is possible for creators to present any story in four-scene comics. On the other hand, the scene lengths are limited and the stories must be deformed in the interests of simplification. Thus, four-scene comics have clear stories and structures and are therefore suitable for computer engineering-related studies. In this section, research on picture model for automatic comic generation is first introduced (Ueno, Mori, & Matsumoto, 2014). Then, a two-scene comic creation system is constructed. After that, the analysis for comic creation is discussed.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Comic Generation Picture Model

In order to represent the semantics of pictures to facilitate computer comprehension, each of the elements in a figure, including the background images and balloons, are defined as certain objects. It is difficult to create a reasonable picture model for all pictures. Therefore, two-scene comics are considered here, because two sequential pictures provide sufficient information compared to a single picture, without excessively increasing the problem complexity. In this section, the authors will introduce the method of generating comics and the two-scene generating system. In this method, it is necessary to prepare a dataset comprising images of objects and related object information, area types, drawing operators, and transitions. The picture contains the above information along with a certain history.

The outline of the comic creation method is given below. First, an initial picture p_0 is generated by the system. p_0 is a blank or a background picture and has several areas determined by the system. The user can utilize p_0 as user input. The user operation set in time t is represented by $\boldsymbol{u}_t$, and $\boldsymbol{u}_t$ contains all potential user operations for picture p_t, p_{t+1} is created by applying $\boldsymbol{u}_t$ to p_t. This is expressed as

$$p_{t+1} = \prod_{u_{i,t} \in \boldsymbol{u}_t} u_{i,t} p_t \,, \tag{21}$$

where u_i represents each user's operations over t. The relations between p_t, o_t, $\boldsymbol{u}_t$, $\{O_t\}$, and U_t are

$$\{O_t\} = \{o_t \in p_t\}, \tag{22}$$

$$\{U_t\} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_t\}, \tag{23}$$

$$p_{t+1} = \prod_{u_{i,t} \in \boldsymbol{u}_t} u_{i,t} p_t \ldots = \prod_{u_{i,t} \in \boldsymbol{u}_t} u_{i,t} \prod_{u_{i,t-1} \in \boldsymbol{u}_{t-1}} u_{i,t-1} p_{t-1} = \prod_{j=0}^{t} \prod_{u_{i,j} \in \boldsymbol{u}_j} \boldsymbol{u}_{i,j} p_0 \tag{24}$$

The definite operation of $u_{i,t} \in \boldsymbol{u}_t$ is represented by the drawing operator ϕ and the operand object in p_t. Each ϕ also has specific augments, such as its color, angle, or other objects. Object o_t can represent the object transition. Finally, object o_n in the n-th scene adheres to the Markov model, such that

$$o_n = \prod_{i=0}^{N_n^{\max}} \phi_i o_0 \tag{25}$$

where $N^{\max}$ is the total number of ϕ to applied to o, and $N_n^{\max} \geq n-1$, because each u_i contains at least one ϕ.

ϕ operates to transform pictures. Further, the product of several ϕ is also ϕ, and the set of base operators is defined as $\left\{\hat{\phi}\right\}$. If the number of base operators is increased, the system has greater representative capacity, because it can create a large number of new images utilizing the existing objects. The system described in this section was written using Java and the operator was implemented using a command design pattern (Gamm, Helm, Johnson, & Vlissides, 1995). The base ϕ corresponds to the com-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Cup (apply ϕ_{crack})



mand and the product of ϕ corresponds to a macro-command. In Java implementation, an operator product is represented using a queue. For example, ϕ_{crack}, which represents a cracked transition, is composed of the base drawing operators $\hat{\phi}_{rotate}\hat{\phi}_{divide}$. Thus, two base drawing operators are applied to o in the case of

$$\phi_{crack} o = \hat{\phi}_{rotate}\hat{\phi}_{divide} o \tag{26}$$

Figure 2 shows the image variation caused by applying the base drawing operators of ϕ_{crack}.

Comic Generation System Flow

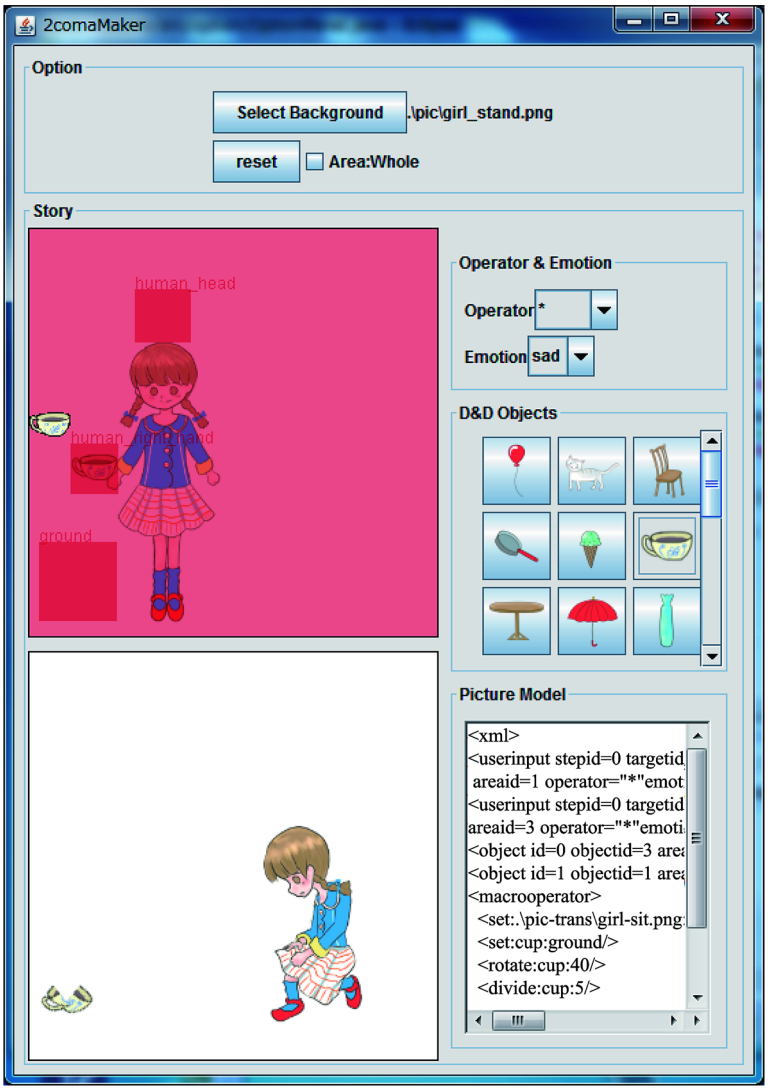
Based on the picture model, the authors propose the comic generation system. The basic flow of the proposed system is as follows.

1. The base image is set as the background for the initial states. This picture is defined as p_0.
2. The system waits for the user input. The user can decide to apply several actions to p_0. In this section, the only available user action involves the setting of an object in the object set displayed on the system graphical user interface (GUI). The user can set their chosen object on a specific area of p_0 using the ϕ and emotion information. Any number of objects can be added to p_0, under the condition the only one object is placed in each area. When the user places an object in p_0', new areas are added to the picture in accordance with the set object.
3. After all of the user inputs u are applied to p_0, p_0' is obtained, which becomes the 1st scene.
4. The system attempts to change objects in p_0' based on the user-defined drawing operator and the emotion information of the target object. The new image comprising the adjusted objects is obtained via transition DB.
5. After all of the changes have been applied to p_0', p_1 is obtained. p_1 represents the 2nd scene.
   a. **Two-Scene Comic:** $\boldsymbol{p}_1$ is the final output.
   b. ***n*-Scene Comic:** p_1' becomes $\boldsymbol{p}_2$. The system again waits for user input and repeats the above flow until $\boldsymbol{p}_{n-1}'$ becomes $\boldsymbol{p}_n$.

Using this method, two-scene comics can be created via the system. Figure 3 shows the outline of this system procedure.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Outline of two-scene comic creating system



Comic Analysis Method

In the previous paragraph, general models of transition of pictures are defined and are applied to the system. Although, several operators are implemented in the system, to depict various scenes, lots of drawing operators should be defined. However, it is difficult to classify and define drawing operator for expressing story transition. Therefore, the relation between object features and story patterns should be analyzed by considering the terms of stories. In comics, there are various kinds of objects and state transitions. The character object usually plays an important role for story transition because characters can act autonomously. The state of character objects are orientations, postures, expressions and so on. The authors assume that character orientations in an image are strongly related to stories and propose

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

a data format of quantifying these orientations. In order to expand the drawing operator, the relationships between orientation transition patterns, impressions of stories, and balloon shape features are also analyzed utilizing existing comics.

The authors propose a data format in which stories are patterned based on the character orientations in accordance with the following two concepts.

1. Suitable features, for example, automatic modeling based on improved image recognition.
2. Stories have several transitions. Characters are typically the main objects that can transition, so their facial orientation may be strongly related to the content and atmosphere of the stories.

The speech balloon shapes are also used to model the data, in accordance with the following two points.

1. The essential elements of comics are pictures, but natural language such as a character's words are also important. Although analysis of the semantics of words is difficult, the shapes of speech balloons can indicate kinds of character's words such as a whisper or loud voice.
2. Automatic detection of balloon areas and shapes has been reported (Tanaka, Toyama, Miyamichi, & Shoji, 2010).

The proposed data format and each value are as shown in Table 7. The table shows the example of four-scene comic data.

***n*:** Story number.
***t*:** Story title.
$\boldsymbol{o_j}$**:** Object j name.
$\boldsymbol{d_i^{oj}}$**:** Facial orientation of object o_j in scene i.
$\boldsymbol{c_j^k}$**:** 1 or 0; indicating whether or not type-k balloons appear.

To distinguish between objects of the same type, the object name is added to its number. A lower number is assigned to an object that has appeared previously. In a scene, an object for which the right end is located near the right side of the scene is considered to appear first. This is because scenes in Japanese comics are usually read from top to bottom and from right to left.

Figure 4 shows the example of facial orientations.

The facial orientation of a character is approximately assigned to a single direction based on the following indicators.

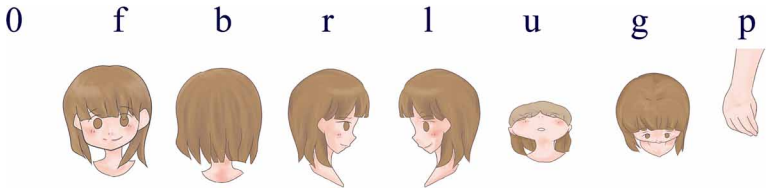
0: The human object does not appear.
f: Front.
b: Back.
r: Right.

Table 7. The example of data of four-scene comics

$\#n$; t ; o_1:d_1 o_1d_2 o_1d_3 o_1d_4 o_1,o_2:$d_1o_2d_2$ o_2d_3 o_2d_4 o_2; c_1^1 c_1^2 c_1^3 $c_1^4c_1^5$, c_2^1 c_2^2 c_2^3 $c_2^4c_2^5$, c_3^1 c_3^2 c_3^3 $c_3^4c_3^5$, c_4^1 c_4^2 c_4^3 $c_4^4c_4^5$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. The example of facial orientation



l: Left.
u: Top.
g: Bottom.
p: Part of the face does not appear.

Figure 5 shows these balloon types.

Further, the balloons are classified into the following five types based on shape, and are assigned corresponding numbers.

1. Oval frame,
2. Zigzag frame,
3. Square frame,
4. Other frame shapes,
5. Printed without frame.

Handwritten balloons are not included in this study. Finally, some cases for which it is difficult to create data are described below.

- **Character Definition:** In some creative works, animals and inanimate objects are personified, so any object that can autonomously act is defined as a character.
  - **The Face Does Not Appear:** Only part of a body is drawn in many comic scenes. If the face does not appear, dioj is p.
  - **Different Facial and Physical Orientations:** This method focuses on the facial orientation only.
  - **It is Difficult to Assign a Single Orientation:** The nearest orientation is assigned.

Figure 6 shows four-scene comics and the example data. In this study, the data were created by hand.

Experiment on Detailed Analysis of One Book

In this experiment, the frequency of object occurrence was investigated and the relationships between expressions and stories are discussed.

- 80 stories from the beginning of “Sketchbook 1” (Kobako, 2003) were used.
- The data were annotated by three graduate students.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. The example of balloon types

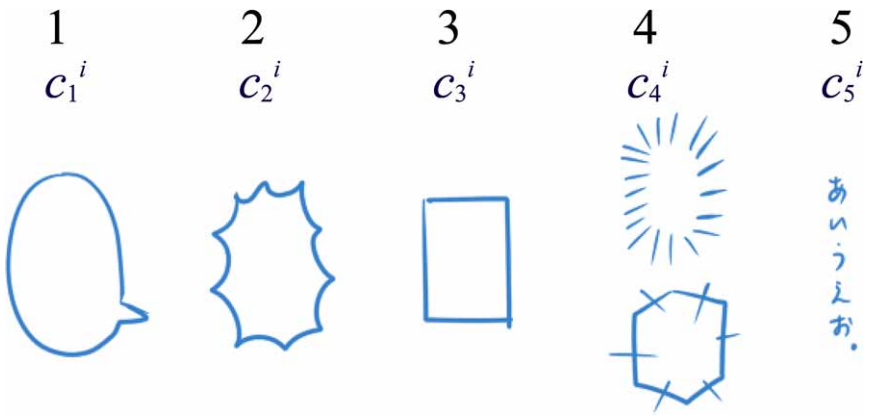
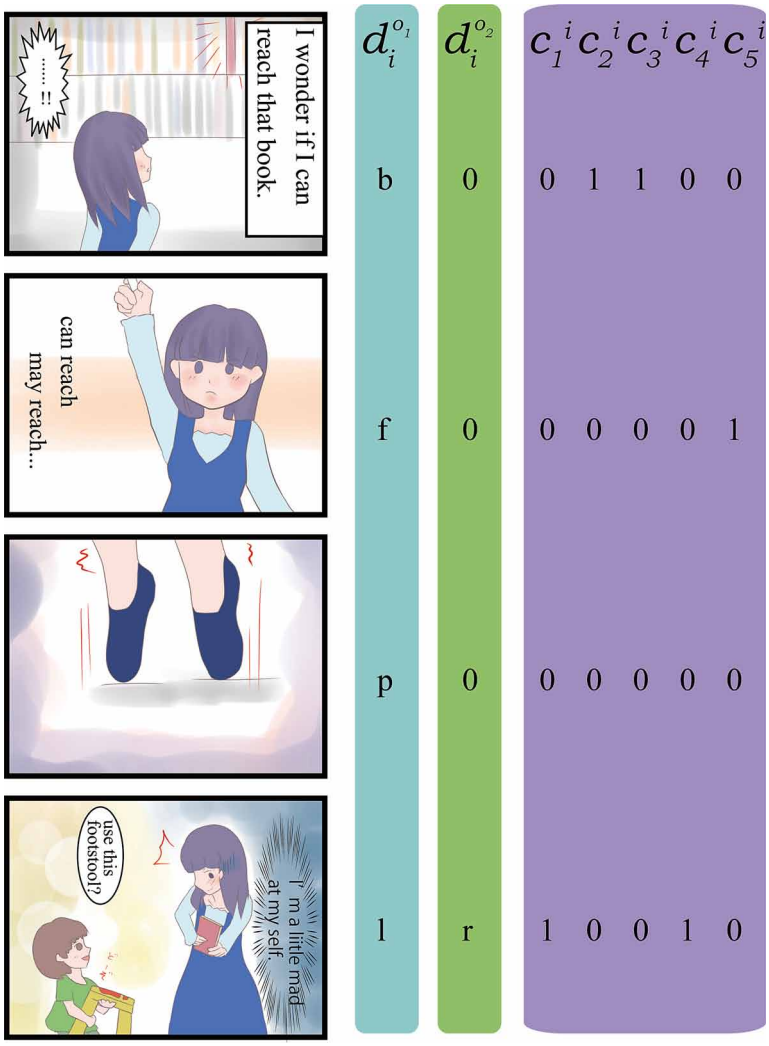


Figure 6. The example of four-scene comics



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Stories that had two characters or less were used, while stories with more than three characters were omitted. Here, o_1 and o_2 represent characters 1 and 2 in each story respectively. Based on these conditions, 66 stories were chosen from the original 80.

The results produced by one of the three annotators are given in this paragraph. Table 8 shows the number of stories per character, Table 9 shows the frequency and order of the o_1 and o_2 orientations, and Table 10 shows the frequency of the o_1 and o_2 orientation patterns for the 66 stories. Finally, Table 11 and Table 12 show the balloon pattern frequency. The results of this experiment can be summarized as follows.

- The orientation patterns $d_1^{o_j}$, $d_2^{o_j}$, $d_3^{o_j}$ and $d_4^{o_j}$ can adopt $8^4 = 4096$ potential patterns. A total of 67 orientation patterns were obtained for the 112 characters considered in this study. This constitutes 1.64% of all possible cases. This data number is not sufficient, so the large-scale dataset must be reviewed. Regardless of whether the data is sparse or an increased dataset is used, consideration of these orientations is a highly effective means of estimating story patterns in comics.
- The balloon patterns c_k^1, c_k^2, c_k^3, and c_k^4 can be any of $2^{20} = 1048576$ patterns. A total of 61 balloon patterns were obtained for the 80 stories considered in this study. Thus, the number of patterns whose frequency was more than 2 was only 7. This corresponds to 6.68×10^{-4}% of the potential dataset, so this can be regarded as sparse data.

Experiment Comparing Three Books

To compare data, the relationships between the stories and balloon patterns for three books were analyzed. The beginnings of each of the 80 stories were used; these stories were contained within two of three books have the same author, and it was therefore assumed that several common patterns would exist between these stories. Table 13 shows the number of variations of two orientation patterns, $d_1^{o_1}$, $d_2^{o_1}$, $d_3^{o_1}$, $d_4^{o_1}$ and $d_1^{o_2}$, $d_2^{o_2}$, $d_3^{o_2}$, $d_4^{o_2}$, and the balloon patterns c_k^1, c_k^2, c_k^3, and c_k^4 in the three books. The number of common pattern types are also shown.

Table 8. Comparison between the number of stories and the number of characters in a story

The Num. of Character	The Num. of Stories
0	0
1	20
2	46
3	10
4	2
5	0
6	2
Total	80

Table 9. The frequency of types of direction

Direction	Freq. of d^{o_1} (Order)	Freq. of d^{o_2} (Order)
0	26 (4)	46 (2)
f	138 (1)	63 (1)
b	20 (5)	10 (5)
r	38 (2)	36 (3)
l	38 (2)	26 (4)
u	0 (8)	0 (8)
g	2 (6)	2 (6)
p	2 (6)	1 (7)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 10. The frequency of the orientation pattern in 112 characters of 66 stories

o_1		o_2		$o_1 \cap o_2$	
Freq.	**Pattern**	**Freq.**	**Pattern**	**Freq.**	**Pattern**
10	ffff	4	0fff	11	ffff
3	frrr	3	rrrf	5	ff0f
3	ff0f	2	rfff	5	0fff
3	f0ff	2	fffr	4	f0ff
2	llll	2	ff0f	3	rrrf
2	lllf	2	00rr	3	rfff
2	frfl	2	000f	3	lllf
2	frff	2	000b	3	frrr
2	fflr			2	llll
2	fffl			2	lffl
2	fffb			2	frfl
2	fbbb			2	frff
2	f000			2	flll
				2	ffrr
				2	fflr
				2	fffr
				2	fffl
				2	fffb
				2	fbbb
				2	f000
				2	0lff
				2	00rr
				2	000f
				2	000b
(The pattern whose frequency lower than 1 is omitted)					
	42 types		35 types		67 types

Result of Three-Book Comparison

The same balloon patterns were frequently found in the two stories by the same author, while seven balloon patterns were common in all three stories. Table 14 shows the seven types of balloon patterns and their frequencies, while Figure 7 shows four higher-ranking balloon patterns. These seven balloon patterns were found in 81 of 240 stories, and correspond to approximately 30% of all patterns. Therefore, it can be concluded that typical balloon patterns exist. For each of the seven patterns, typical object types and character emotions were drawn in certain scenes. Therefore, balloon patterns can possibly be regarded as a semantic class of stories. In future research, scene and story meanings will be obtained based on training data using automatically extracted balloon patterns.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 11. The frequency of the balloon pattern in 80 stories

c_k^1		c_k^2		c_k^3		c_k^4	
Freq.	**Pattern**	**Freq.**	**Pattern**	**Freq.**	**Pattern**	**Freq.**	**Pattern**
27	00000	28	00000	28	00000	23	10000
27	10000	26	10000	28	10000	17	00000
11	00001	9	00001	6	00001	9	00001
6	00010	6	00010	4	00100	8	10010
4	01000	5	10010	3	01000	7	00010
2	10010	2	00100	3	10010	4	01000
1	00100	1	01000	2	00010	3	00100
1	10001	1	01001	1	00101	2	00101
1	11000	1	10001	1	00110	2	01001
		1	11000	1	01001	2	01010
				1	01100	2	11000
				1	10001	1	10100
				1	11000		
9 types		10 types		13 types		12 types	

Table 12. The frequency of the balloon pattern two in 80 stories

$c_k^1 \sim c_k^4$		$c_k^1 c_k^2 c_k^3 c_k^4$	
Freq.	**Pattern**	**Freq.**	**Pattern**
104	10000	7	10000,10000,10000,10000
100	00000	5	00001,00001,00001,00001
35	00001	5	00000,00000,00000,00000
21	00010	3	00000,00000,10000,10000
18	10010	2	10000,10000,10000,00000
12	01000	2	10000,00000,00000,00000
10	00100	2	00000,00000,00000,00100
5	11000		
4	01001		The pattern whose the frequency
3	00101		lower than 1 is omitted
3	10001		
2	01010		
15 types		61 types	

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

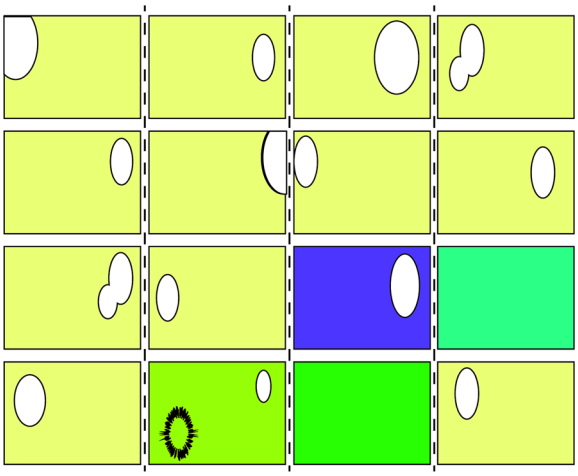
Table 13. Comparing the features of three books

	Direction o_1	Direction o_2	Balloon
Book 1	42	35	61
Book 2	25	29	42
Book 3	20	21	46
A ∩ B	3	6	8
A ∩ C	4	5	7
B ∩ C	4	6	10
A ∩ B ∩ C	1	3	7

Table 14. Common seven balloon patterns among three books

Frequency	Balloon Pattern
48	10000,10000,10000,10000
10	10000,10000,10000,10010
6	10000,10000,10000,00000
6	10000,10000,00000,10000
5	10000,10000,10000,01000
3	10000,10000,11000,11000
3	10000,10000,10010,10010

Figure 7. Examples of higher-ranking balloon patterns



Comic Analysis Conclusion

In this section, with a view towards realizing story representation for computer operations, a data format through which stories are modeled based on character orientations and balloon types was shown. Using the proposed data format, example data from existing comics can be created and the possibility that their characteristic features are well related to the terms of stories is shown. The following further developments are now required.

- Construction of large-scale data.
- Representation of orientations using angles and vector definitions, in order to realize a class estimation method for application to detailed comic stories based on machine learning.
- Surveying of comic authors to identify the essential elements necessary to create comics.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ANALYSIS OF ANIMATION STORYBOARD

In the previous section, the authors focused on stories told by pictures. Pictures are the key components not only for comics, but also for animation. The creation of animation has common processes compared to that of comics because one animation is usually a collaborative creation by several animators. In the animation creation process, the storyboard is a very important component as regards story design. Animators draw original pictures in accordance with the storyboard, and these original pictures depict certain events within a very limited timeframe. The directors then connect large numbers of original pictures to create stories. Namely, one story board requires several creators of original pictures. Facilitators should assign the drawing of each scene to the appropriate creator depending on the difficulty in drawing. However, the field of animation is facing labor shortage. Thus, humans strongly hope that computers can play an important role in this process. Here, the process of translating original pictures from storyboards is considered as using computers in one of the important processes of animation. Furthermore, the cost of creating original pictures from storyboards is calculated based on the number of original pictures, but the difficulty is not considered. In order to eliminate such discrepancies, a method for calculating the difficulty of the process should be proposed. In this section, the final goal is to define a difficulty index regarding the creation of an original picture from a storyboard created by an animator.

Target Works and Features

A method of classifying the difficulty index using a multi-class support vector machine (SVM) (Cortes & Vapnik, 1995) is described. The SVM receives feature values of each cut in a storyboard as input and classifies the output index.

Here, part A of the 10th episode of "Inou Battle Wa Nichijou Kei No Naka De" (Nozomi, 2014) is taken as the dataset. Note that a storyboard contains several elements: pictures, a number of cuts, a number of frames, lines, and cut contents. Thus, a storyboard is a blueprint of an original picture, because the creators create an original picture based on those components. In this section, the following storyboard components are extracted and defined as an SVM input vector.

- **Number of Original Pictures:** The number of original pictures, including the cut.
- **Number of Frames:** The number of frames, including the cut.
- **Cut Content:** The explanations of the cut given in the storyboard.
- **Lines:** The lines (dialogue) of the animated person described in the storyboard.
- **Sound Effects:** Any sound effects or music described in the storyboard.
- **Difficulty Index:** An index representing the difficulty of creating an original picture from the storyboard.

Note that the difficulty index is evaluated from the storyboard and the original picture by the annotators, who produce animation, based on their experience. The lines and cut content are translated into a numerical vector using a bag of words, because they are described in natural language. In addition, the number of original pictures is also used as a vector element.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 15. Number of pictures per class

Class	The Number of Pictures
1	28
3	95
5	7
Total	130

Storyboard Analysis Experiment

In order to evaluate the difficulty of translating original pictures into cuts, the index of each cut is evaluated on a five-level scale, where: 1 = easy, 2 = relatively easy, 3 = normal, 4 = slightly difficult, and 5 = difficult. In order to simplify the problem structure and to avoid data bias, indices 1 and 2 are classified as class 1, index 3 is classified as class 3, and indices 4 and 5 are classified as class 5. To evaluate this method, a multi-class SVM was trained using the data set, utilizing a linear kernel and setting the parameter cost to 32. These settings were determined in preliminary experiments. The number of pictures in Part A of the 10th episode of "Inou Battle Wa Nichijou Kei No Naka De" is 130, and Table 15 shows the number of pictures in each class. A five-fold cross-validation identification result of 79.03% was obtained through 100 iterations. As this value is above the 73.08% current baseline, it indicates that the SVM can learn and identify the difficulty index of translating a given storyboard into original pictures.

Conclusions from Animation Storyboard Analysis

In this section, the authors focus on one of the important processes of animation and a method through which the animation process can be evaluated was shown. The following future developments are required in order to further refine this technique.

- The results of other kernels must be considered.
- Data correlation between multiple evaluators must be considered.
- Cooperation with storyboard creators must be incorporated into the research.

FUTURE RESEARCH DIRECTIONS

Thus far, the generation of creative work has been mainly discussed in the fields of literature and art; however, recently, it has become a topic of engineering research. Although the authors know very well that a large number of difficult tasks are yet to be achieved for the analysis of creative work by computers, the computational models discussed in this chapter will give readers novel ideas and generate interest on the topic. Some directions for future work are as follows:

- There exists inter-textuality among existing stories. This information is useful for finding common flows in human story comprehension. Some creative work is available on the web under the Creative Commons license. On the other hand, a considerable number of creative works are pro-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

tected by copyrights. Copyright is a big problem that prevents the advance of research in this field. In this chapter, the authors prepared the creative work ourselves and will provide these creative products with the related metadata for use in future research work.

- The personality of the characters is the key information of stories because at times, the story patterns are strongly related to the character personalities. Character personality is regarded as a unique distribution of the transition that the character has. Thus, the development of a model of character personality should be focused upon.

CONCLUSION

In this chapter, the authors introduced several methods and computational models to automatically generate creative work. Novels, comics, and animations were focused upon because they are depicted by pictures and natural languages. Such creative work gives rise to a considerable number of interesting and challenging issues related to computational narratology.

The results of this chapter can be summarized as follows:

1. The authors proposed a method of semi-automatic generation of stories by agent-based simulation (ABS). Compared to the previous method based on existing novels, the proposed method can create flexible transitions for stories. A story is a set of transitions from the current states to the next states; therefore, the definition of appropriate units of transitions and the maintenance of logical coherence and unpredictability are important issues to be tackled in the future.
2. Computational models for generating and analyzing comics are constructed. The authors developed the comic generating system utilizing the prepared object images and the operators that represent image transition. Further, the authors found individual relations between the states of objects in the pictures and the stories.
3. The authors also proposed a method using an SVM to estimate the difficulties of each scene in the storyboard, which is the key process of creating animations. The method encourages human beings to co-create animations with computers by referring to the difficulties of each part.

Computational models for creative work are being developed, and their structure varies according to the aim of the research. If complete models are defined in the future, the authors believe that the computer will be able to create comics and animations along with human beings.

REFERENCES

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. doi:10.1007/BF00994018

Eitz, M., Hays, J., & Alexa, M. (2012). How do humans sketch objects? *ACM Trans. Graph., 31*(4), 44:1-44:10.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Fukuda, K., Mori, N., & Matsumoto, K. (2015). Analysis of automatic novel generation without existing novels information. In *Proceedings of the 29th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI'15)*. Hakodate, Japan: The Japanese Society for Artificial Intelligence.

Gamma, E., Helm, R., Johnson, R., & Vlissides, J. (1995). *Design patterns: elements of reusable object-oriented software*. Addison-Wesley.

Gervás, P., Diaz-Agudo, B., Peinado, F., & Hervás, R. (2005). Story plot generation based on CBR. *Knowledge-Based Systems Journal*, *18*(4-5), 235–242. doi:10.1016/j.knosys.2004.10.011

Gohara, H., Sugimoto, S., & Morishima, S. (2010). Data driven in-betweening for hand drawn rotating face. In *Proceedings of the ACM SIGGRAPH 2010 Posters Article* No. 7. Los Angeles, CA: ACM. 10.1145/1836845.1836853

Imabuchi, S., & Ogata, T. (2012). A story generation system based on Propp theory: As a mechanism in an integrated narrative generation system. In H. Isahara & K. Kanazaki (Eds.), *Advances in Natural Language Processing: Proceedings of the 8th International Conference on Natural Language Processing* (Vol. 7614, pp. 312-321). Berlin, Germany: Springer. 10.1007/978-3-642-33983-7_31

Kobako, T. (2003). *Sketchbook 1*. Tokyo, Japan: Mag Garden.

Kouyama, S., & Sumi, K. (2014). A discussion of humorous story generation system by using stories of Shinichi Hoshi. In *Proceedings of the 28th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI'14)*. Ehime, Japan: The Japanese Society for Artificial Intelligence.

Macal, C. M., & North, M. J. (2010). Tutorial on agent-based modeling and simulation. *Journal of Simulation*, *4*(3), 151–162. doi:10.1057/jos.2010.3

Matsubara, H., Sato, S., Akaishi, M., Sumi, K., Mukaiyama, K., Nakashima, H., & Otsuka, Y. (2013). An attempt at automatic composition of Shin'ichi Hoshi-like short short stories. In *Proceedings of the 27th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI'13)*. Toyama, Japan: The Japanese Society for Artificial Intelligence.

Meehan, J. R. (1980). *The metanovel: writing stories by computer*. New York, NY: Garland Publishing.

Nozomi, K. (2014). *Inou battle wa nichijou kei no naka de* [Battles using special skills in usual scenes]. Tokyo, Japan: SB Creative.

Ogata, K., Sato, S., & Komatani, K. (2014). Automatic generation of very short stories by imitation and substitution. In *Proceedings of the 28th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI'14)*. Ehime, Japan: The Japanese Society for Artificial Intelligence.

Pérez y Pérez, R., & Sharples, M. (2001). MEXICA: A computer model of cognitive account of creative writing. *Journal of Experimental & Theoretical Artificial Intelligence*, *13*(2), 119–139. doi:10.1080/09528130010029820

Sumi, Y., Sakamoto, R., Nakao, K., Mase, K., & Kunifuji, S. (2002). Change of individual experiences and interests in comics style. *IPSJ Journal*, *43*(12), 3582–3595.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Tanaka, T., Toyama, F., Miyamichi, J., & Shoji, K. (2010). Detection and classification of speech balloons in comic images. *The Journal of The Institute of Image Information and Television Engineers*, *64*(12), 1933–1939. doi:10.3169/itej.64.1933

Turner, S. R. (1994). *The creative process: A computer model of storytelling and creativity*. Oxford, UK: Psychology Press.

Ueno, M., Mori, N., & Matsumoto, K. (2014). *2-scene comic creating system based on the distribution of picture state transition. Advances in Intelligent Systems and Computing* (Vol. 290, pp. 459–467). New York, NY: Springer, International Publishing.

KEY TERMS AND DEFINITIONS

Agent-Based Simulation: Simulation to investigate actions and interactions between autonomous agents.

Creative Works: Intellectual products created by humans, such as pictures, novels, and animation.

Drawing Operator: The representation of the transition of stories to transform pictures.

Original Picture: Key animation which is the start, key, or end scene of motion.

Plot: A set of state transitions that are arranged in reading order in narratives.

Story: A set of state transitions that are arranged in chronological order in narratives.

Story Vector: Information pertaining to the elements of a story that is represented in vector form.

Storyboard: Story design that has information required for original picture generation.

This research was previously published in Computational and Cognitive Approaches to Narratology edited by Takashi Ogata and Taisuke Akimoto; pages 164-190; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 52
A Novel Approach to Managing the Dynamic Nature of Semantic Relatedness

Youngseok Choi
Brunel University London, UK

Jungsuk Oh
Seoul National University, South Korea

Jinsoo Park
Seoul National University, South Korea

ABSTRACT

This research proposes a novel method of measuring the dynamics of semantic relatedness. Research on semantic relatedness has a long history in the fields of computational linguistics, psychology, computer science, as well as information systems. Computing semantic relatedness has played a critical role in various situations, such as data integration and keyword recommendation. Many researchers have tried to propose more sophisticated techniques to measure semantic relatedness. However, little research has considered the change of semantic relatedness with the flow of time and occurrence of events. The authors' proposed method is validated by actual corpus data collected from a particular context over a specific period of time. They test the feasibility of our proposed method by constructing semantic networks by using the corpus collected during a different period of time. The experiment results show that our method can detect and manage the changes in semantic relatedness between concepts. Based on the results, the authors discuss the need for a dynamic semantic relatedness paradigm.

INTRODUCTION

Studies on the measurement of semantic relatedness, or similarity between concepts, have a long history in the fields of computational linguistics, psychology and computer science. Resolutions to various issues regarding semantic ambiguity often boil down to methods of relatedness or similarity between concepts.

DOI: 10.4018/978-1-7998-0951-7.ch052

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Humans are very intuitive in determining similarity or relatedness between concepts but designing a computer to perform such a task is very difficult (Sheth, Aleman-Meza, Arpinar, & Bertram, 2005). The most widely used method for calculating similarity is to employ a ready-made ontology or taxonomy. Specifically, various WordNet-based methods (Resnik, 1995; Jiang & Conrath, 1997) have produced a relatively high level of performance.

The method of using auxiliary information such as ontology or taxonomy is based on the premise that the semantic relatedness between concepts will not change as time goes by. We call it the *static semantic relatedness paradigm* (Resnik, 1995; Jarmasz & Szpakowicz, 2004). Based on the assumption of static semantic relatedness, semantic similarity is measured at a particular point in time with an ontology, taxonomy, or thesaurus. In this case, the ontology or taxonomy is directly written in advance by humans.

However, such a static semantic relatedness paradigm has been challenged by the dynamic nature of information due to the development of knowledge-sharing systems and the spread of information. As new events and the social contexts around them continue to influence how people think, the nature of relationships between concepts may change. At the same time, the events that make the nature of relatedness between concepts dynamic spread rapidly as shared knowledge via information-sharing services and technologies such as social network or news feed. In this regard, there is a need for a paradigm shift toward a *dynamic semantic relatedness paradigm*. In this paper, we use the term "*dynamic semantic relatedness paradigm*" to describe the research considering the dynamics of semantic relatedness. If a change in semantic relatedness occurs due to a particular event which can affect the human thinking process, the existing methods of the static semantic relatedness paradigm cannot fully reflect this dynamic for the following reasons: normally, the existing auxiliary information (i.e., an ontology, taxonomy, or thesaurus) is not updated frequently due to the fact that recognizing the occurrence of semantic changes and rewriting the ontology or making a new taxonomy requires considerable effort. As a result, it is difficult to maintain up-to-date auxiliary information. In addition, the existing auxiliary information refers only to the semantic relatedness relevant for general situations. For example, reaching the conclusion that "Apple" and "Music" are closely connected is difficult by using WordNet[1].

The dynamics of semantic relatedness have not received much attention that they deserve in the semantic relatedness research. Some related studies hint at the idea of incorporating dynamics of semantic relatedness and propose algorithms to address the issue but are insufficient in terms of their complexity and agility to cope with dynamics. So, we present a critical review of existing research following *static* and dynamic *semantic relatedness paradigms* and address how to deal with the dynamics in semantic relatedness better than previous works.

In this context, a new computational approach to the semantic relatedness that takes into consideration of dynamic nature of the postulated semantic relatedness is necessary. In this regard, contributions of this study can be summarized as follows:

- We propose a novel method to corpus-based semantic relatedness calculation by constructing semantic networks that reflect the time and context-dependency nature of the corpus.
- The rationality of the proposed method is validated by experimental evaluation using actual corpus data collected from a particular context over a specific period of time.
- This approach enables quantification of the change of semantic relatedness among specific concepts as well as the dynamics of a semantic network itself.
- Based on the above, the necessity of a dynamic semantic relatedness paradigm is demonstrated.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

RELATED WORKS

Previous studies on semantic similarity can broadly be classified into: those based on the topological method and those based on the statistical method. The topological method uses various forms of semantic network or taxonomy to compute the semantic relatedness between concepts. The representative semantic networks based on the topological method are WordNet and Roget's Thesaurus. On the other hand, the statistical method is a technique for calculating the statistical association between concepts by using auxiliary data, such as a dictionary or corpus that is not in the form of a network. In the following subsections, topological and statistical methods are discussed in more detail. In addition to reviews on these two traditional approaches, some studies dealing with the dynamics of semantic relatedness will also be reviewed. On the basis of comprehensive reviews, we will position our research along with the *dynamic semantic relatedness paradigm*.

Topological Methodology

The study of similarity between concepts aims to discover how a computational process can model actual human behavior and action in determining the relationship between two concepts. The reason that many studies on semantic relatedness employ semantic networks is that the system of human knowledge can be well expressed as a network. In fact, many topological methods with semantic networks have produced relatively good performances (Hale, 1998; Resnik, 1995).

Various semantic networks utilized by the topological method include a hierarchical taxonomy. The nodes in a network represent concepts, and the distance or the shared information quantity between two nodes is used for estimating the conceptual similarity between two corresponding concepts. The topological method can be categorized as either node-based or edge-based, which is also called the information content approach or the conceptual distance approach, respectively.

Node-Based Approach

The node-based approach considers how much information the two concepts share in the context of a semantic network or taxonomy. In this regard, the node-based approach is also called the information content approach. The node-based approach is mostly based on information theory (Resnik, 1995). According to information theory, the information content (IC) of concept C can be quantified as follows:

$$\mathrm{IC}(C) = \log^{-1}\mathrm{P}(C),$$

where $\mathrm{P}(C)$ is the probability of encountering an instance of concept C. Then, the similarity between two concepts can be formally defined as follows:

$$\mathrm{sim}\,(c_1, c_2) = \max\,[\mathrm{IC}(c)],\ c \in \mathrm{Sup}(c_1, c_2),$$

where $\mathrm{Sup}(c_1, c_2)$ is a set of concepts that subsume both c_1 and c_2. Resnik (1995) only considers the information content of the least common subsumer (*lcs*), which has the shortest distance from the two concepts being compared. An *lcs* is a concept in a lexical taxonomy (e.g., WordNet).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Based on Resnik's information content approach, Lin (1998) proposes a similarity theorem for measuring the distance between two concepts. The similarity between *A* and *B* is measured by the ratio between the amount of information needed to state their commonality and that needed to fully describe them:

$$\mathrm{sim}\left(\mathrm{A},\mathrm{B}\right)=\frac{\log\ \mathrm{P}\left(\mathrm{common}\left(\mathrm{A},\mathrm{B}\right)\right)}{\log\ \mathrm{P}\left(\mathrm{description}\left(\mathrm{A},\mathrm{B}\right)\right)}$$

Hence, if we know the commonality of two concepts, their similarity indicates how much additional information is required to define the two concepts.

Couto, Silva, & Coutinho (2007) use the node-based approach to calculate semantic similarity between concepts in gene ontology (GO) through a measure named Graph-Based Similarity Measure (GraSM). They use all the information in the graph structure of the GO instead of considering it as a hierarchical tree. Their measure of semantic similarity between GO terms demonstrates its greater effectiveness by obtaining a higher correlation using disjunctive common ancestors than using only the most informative common ancestor.

Gan, Dou, & Jiang (2013) apply this node-based approach to concept similarity in the domain ontology and Harispe, Sánchez, Ranwez, Janaqi, & Montmain (2014) use similar approach for the semantic relation in medical domain. Lee, Kim, & Lim (2016) also adopt this approach for semantic network visualization due to its convenience to represent graph type semantic network.

Edge-based Approach

The edge-based approach is more natural and intuitive than the node-based approach for calculating semantic similarity in a semantic network. This approach estimates the distance between the nodes that correspond to the concepts being compared. Obviously, the shorter the distance between two nodes, the greater is their similarity.

Rada, Mili, Bicknell, & Blettner (1989) as well as Rada & Bicknell (1989) adopt this approach for a hierarchical semantic network of terms used for indexing articles in the bibliographic retrieval system MEDLINE, which uses medical subject headings (MeSH). Their principal assumption for evaluating semantic similarity is that "the number of edges between terms in the MeSH hierarchy is a measure of conceptual distance between terms" (Rada, Mili, Bicknel, & Blettner, 1989, pp. 18). Based on this assumption, the distance between document *Doc* and query *Q* is defined as follows:

$$\mathrm{Distance}\left(Doc,Q\right)=\frac{1}{mn}\sum_{t_i\in Doc}\sum_{t_j\in Q}d\left(t_i,t_j\right)$$

where $d(t_i, t_j)$ is the minimal number of edges in a path from t_i to t_j within a hierarchy.

Lee, Kim, & Lee (1993) demonstrate an information retrieval method based on a conceptual distance in IS-A hierarchies. They developed a method called the Knowledge-Based Extended Boolean Model (KB-EBM), which involves a primitive distance function. Given terms t_i and t_j in an IS-A hierarchy, the distance (which is often called the primitive distance function) between the terms is as follows:

Distance (t_i t_j) = minimal number of IS-A relationships between t_i and t_j

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Statistical Method

The statistical method computes similarity by analyzing auxiliary information, such as a corpus or dictionary, in a statistical way. Applications of pointwise mutual information (PMI) or latent semantic analysis (LSA) are prime examples of the statistical method.

Use of PMI of data collected by information retrieval (PMI-IR) is proposed by Turney (2001) as an unsupervised measure of semantic similarity between words. PMI-IR is based on word co-occurrence by using counts accumulated over very large corpora (e.g., the Web). Given two words w_1 and w_2, their PMI-IR is measured as follows:

$$\text{PMI-IR}\left(w_1, w_2\right) = \log_2 \frac{\text{P}\left(w_1 \,\&\, w_2\right)}{\text{P}\left(w_1\right) * \text{P}\left(w_2\right)}$$

which indicates their degree of statistical dependence and can be used as a measure of their semantic similarity.

Another statistical measure of semantic similarity is the result of the latent semantic analysis (LSA) proposed by Landauer, Foltz, & Laham (1998). LSA is a "method for extracting and representing the contextual-usage meaning of words by statistical computations applied to a large corpus of text" (Landauer, Foltz, & Laham, 1998, pp. 259). In LSA, term co-occurrences in a corpus are captured by means of a dimensionality reduction and singular value decomposition of the term-by-document matrix representing the corpus.

LSA has been applied for many situation regarding semantic relatedness and semantic disambiguation contexts. Hoffman, Ralph, & Rogers (2013) apply the LAS for semantic disambiguation in the contextual usage of words and Martin, Martin, & Berry (2016) use this statistical approach for the evaluation of questionnaire response.

Research Dealing with the Dynamics of Semantic Relatedness

While a typical topological approach and statistical methods have not been dealing with the dynamics of semantic relatedness, some studies challenge the dynamic nature of semantic relatedness by modifying existing methods that belong to topological and statistical approaches.

Topic tracking and topic model research hint at the nature of dynamic semantic similarity by providing algorithms for detecting rising and falling topics in a target corpus. Blei & Lafferty (2006) suggest a dynamic topic model using a time series random process. They implement a probabilistic time series model to analyze the time evolution of topics in large document collections. As an extension of the topic model, they revised the latent Dirichlet allocation (LDA) considering the distribution of topic proportions as a random variable. Though they did not directly propose the concept of semantic similarity, they developed a statistical model to expect the rise and fall of topics in documents.

Rinaldi (2009) proposes an ontology driven approach for semantic information retrieval using dynamically obtained ontology to cope with the change of semantic relatedness as time changes. They also suggest a new method to calculate the semantic similarity by modifying the algorithm proposed by Rada, Mili, Bicknell, & Blettner (1989) and Wu & Palmer (1994). However, the semantic network

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

is extracted from a dictionary (i.e., WordNet) and more focused on reflecting context than dynamics with regard to the time change. If words and their relationships are not updated, the extracted semantic network might be outdated as well.

Radinsky, Agichtein, Gab rilovich, & Markovitch (2011) argue that their research is the first attempt to incorporate temporal evidence into models of semantic relatedness. They propose temporal semantic analysis (TSA) on the basis of concept vectorization. Human generated folksonomy (i.e., Flikr taggs, Wikipedia, and Del.icio.us book marks) are used for concept vectorization as it needs the relevant terms that will be used for components of concept vectors. Nevertheless, it also contains limitations as the concept of vectorization solely depends on existing folksonomies.

Milikic, Jovanovic, & Stankovic (2011) also attempt to confirm the dynamics of semantic relatedness by analyzing twitter data. They insist that most approaches consider the semantic relatedness of terms as static over time, and disregard the eventual temporal changes as imperfections. They formulated the relatedness measure using the search result of the Microposts search engine. They confirm that the semantic relatedness changes by showing possible scenarios and very simple examples without evaluation.

Though Sagi, Kaufmann, & Clark (2009) do not propose a new algorithm for detecting change of semantic relatedness, they also hint at the change of word semantic form linguistics perspective. They provide empirical evidence for change of usage pattern of specific words and phrases using latent semantic analysis (LSA). Their research objective is not exactly the same as the semantic relatedness context, but their theoretical reviews on semantics deserve to be highlighted from a perspective of the "*dynamic semantic relatedness paradigm.*"

Zhao, Wang, Cao, Zhang, Yin, & Xu (2015) introduce new approach called large-scale topic-related semantic information flow detection. This research proposes a method to track the dynamics and evolution of web information using novel graphical representation method but it lack the method for quantifying and managing the dynamics of semantic relatedness measure.

Table 1 contains a comparative summary of these approaches. Each of these studies contributed to incorporating the dynamic nature of semantic relatedness, albeit not without weaknesses. One common drawback is their dependency on external resources such as dictionaries and WordNet. As the external resource usually contains concepts and their relation in a general context, it is not appropriate to confine the context with regard to a given situation and it is difficult to reflect the change of semantic relation as time passes if resources are not updated in a timely manner. In addition, existing studies do not measure and quantify the change of semantic relatedness though they adopt a complex computation method from the literature.

METHODOLOGY FOR CONSTRUCTING A SEMANTIC NETWORK AND DETECTING ITS DYNAMICS

Dynamics of Semantic Relatedness

Most human knowledge is influenced and updated as new information or a new experience is acquired. Furthermore, with the advancement of IT and, particularly, recent widespread adoption of information sharing services based on social networks, information-acquiring methods have flourished at unprecedented speeds. Therefore, there is an even greater chance that the semantic relation between concepts recognized by humans may change.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Summary of relevant research to the dynamic semantic relatedness paradigm

Related research	Reference algorithm	Output	Resource dependency	Quantification of dynamics of semantic network and relatedness	Evaluation method	Note
Blei, & Lafferty (2006)	LDA	Topic set	N/A	N/A	Case studies	Dealing with sets of topics, not relatedness
Rinaldi (2009)	TSA	Semantic relatedness	Dictionary and WordNet, Search engine result	N/A	Case studies	
Radinsky, Agichtein, Gabrilovich, & Markovitch (2011)	Topological method	Semantic relatedness	External folksonomy	N/A	N/A	Only static semantic relatedness was evaluated
Milikic, Jovanovic, & Stankovic (2011)	Search engine-based	Semantic relatedness	Search engine result	N/A	N/A	Scenario with simple examples
Sagi, Kauffmann, & Clark (2009)	LSA, vector space models	Change of collocation pattern	N/A	N/A	Case studies	More focused on theoretical aspects of language usage patterns
Zhao, Wang, Cao, Zhang, Yin, & Xu (2015)	Vector space models	Information flow diagram	N/A	N/A	Case studies	Validated using news in Chinese

As seen in Related Works section, most studies on semantic relatedness have not considered the dynamics of semantic relatedness, despite the evolutionary nature of the semantic relatedness concept. Many studies, of course, have addressed the ontology versioning issue (i.e., ontology evolution), but have not applied it to the relatedness/similarity between concepts. Our proposed approach provides an intuitive means to estimate dynamic semantic relatedness with a simple but innovative method. By analyzing the corpus generated in a given time and context, the dynamics with regard to the time and context of the corpus can be reflected to semantic relatedness without complex computation and algorithms. The results of analyzing the corpus can be presented in the form of a semantic network and a created semantic network enables quantification of the change of the whole semantic network with regard to the time change.

Co-Occurrence Statistics of Two Concepts in a Corpus and Its Matrix Expression as Network Data

In linguistics, a corpus (plural corpora) or text corpus is defined as a large and structured set of texts (at present, these are usually electronically stored or processed) (Kilgarriff, 2003). A corpus consists of many documents that are either a set of words or concepts separated by a period. Even though there are many ways to define the co-occurrence of concepts, in this paper, we define the co-occurrence as two terms that appear in the same sentence.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Most studies based on corpus analysis (Doddington, 2002; Ohsawa, Benson, & Yachida, 1998; Roark, & Charniak, 1998) have considered each sentence as a basket, much like a collocation which analyzes an expression composed of two or more words that correspond to conventional means of colloquial and literary style. We also adopt these criteria of word co-occurrence by considering one sentence in a document as a meaning basket.

In the following, we introduce a general notation of our proposed methodology for semantic network construction using the co-occurrence statistics of two concepts in a corpus. The first step is counting frequencies of words that are in the concept set being evaluated. Based on this result, conditional frequencies between two terms are calculated. Finally, a co-occurrence matrix is obtained as a network data representation. The (i, j) component of the co-occurrence matrix represents the relatedness between i-th and j-th concept. Using this matrix, a network graph of concepts can be constructed.

Let corpus(t_1, t_2) be the corpus, which is generated during the time period from t_1 to t_2 and $\boldsymbol{C} = \{C_1, C_2, \ldots, C_N\}$ be the collection of n concepts in which we are interested[2]. For concept C_i and C_j, conditional probability of term frequency prob ($C_i \mid C_j$) is defined as:

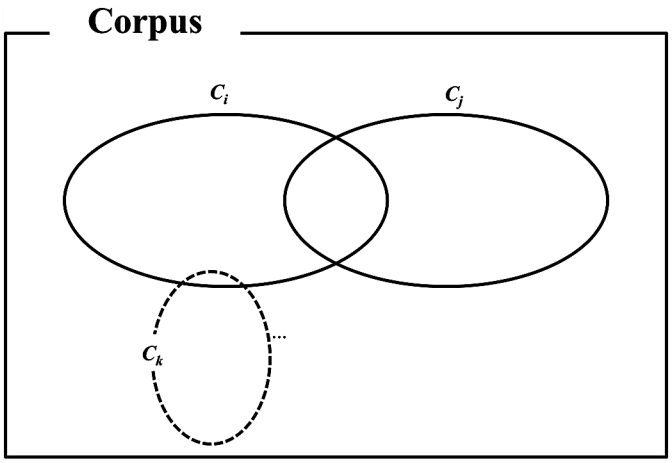
$$\text{prob}\left(C_i \mid C_j\right) = \frac{\text{freq}\left(Ci \cap Cj\right)}{\text{freq}\left(C_i\right)}$$

where:

- freq(C_j) is a term frequency of concept C_j in the corpus and
- freq($C_i \cap C_j$) is a frequency of co-occurrence of two concepts in a meaning basket in the corpus.

Generally, ratios of relative frequencies do not fit well into the hypothesis testing paradigm for collocation analysis (Manning & Schütze, 1999). However, we use this method because the objective of our study is to construct a semantic network using a specific corpus rather than simply testing a hypothesis.

Figure 1. Conceptual diagram for understanding the relationship between relative frequency and semantic relatedness



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1 illustrates the relationship between conditional probability and semantic relatedness. Assuming that concept C_i appears independently from concept C_k in the corpus, the conditional frequency of co-occurrence is similar to the unconditional frequency of concept C_i meaning that the following equation is satisfied:

$\text{prob}(C_i \mid C_k) = \text{prob}(C_i)$ for arbitrary concept C_k that is independent from concept C_i.

In contrast, if two concepts are dependent due to the influence of semantic relationship, the conditional probability is biased so that the following inequality is satisfied:

$\text{prob}(C_i \mid C_j) > \text{prob}(C_i)$ for arbitrary concept C_j that is related to concept C_i[3].

Terms that contain the biased conditional probability of co-occurrence may have an important implication in the corpus. Thus, we assume that the bias of conditional probability can be used as a surrogate of semantic relatedness.

The conditional probabilities of occurrences for given N concepts in the corpus can be expressed as a matrix. The first order co-occurrence matrix M is defined as:

$$1^{\text{st}} \text{ order co-occurrence matrix } M = \begin{pmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & \ddots & \vdots \\ m_{n1} & \cdots & m_{nn} \end{pmatrix} \quad (1)$$

where $m_{ij} = \text{prob}(C_i \mid C_j)$.

Diagonal components of M should be one by definition of conditional probability. Since conditional probability is not commutative (i.e., $\text{prob}(C_i \mid C_j) \neq \text{prob}(C_j \mid C_i)$), M may be asymmetric.

Direct and Indirect Path in the Matrix Expression and Semantic Relatedness Calculation

The reason we define matrix M as "the first order" co-occurrence matrix is that M reflects only a direct co-occurrence of two concepts in the meaning basket. On the other hand, an indirect co-occurrence cannot be ignored because if concept A co-occurs frequently with concept B and C respectively, concepts B and C might be closely related. An intuitive explanation of an indirect co-occurrence is depicted in Figure 2.

Even when the conditional probability of co-occurrence between two concepts has a very low value, there might be a non-negligible semantic relation if two concepts have common co-occurrence concepts. In other words, C_i and C_j may have a semantic relationship if both $\text{prob}(C_i \mid C_k)$ and $\text{prob}(C_j \mid C_k)$ have high values, although $\text{prob}(C_i \mid C_j)$ is low. For this reason, a new metric of indirect co-occurrence is needed.

The second order co-occurrence matrix that represents indirect co-occurrence is constructed based on the basic graph theory in which the total number of indirect links is counted. In Figure 3, X is a matrix representation of the graph of the system.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Direct and indirect co-occurrences between two concepts

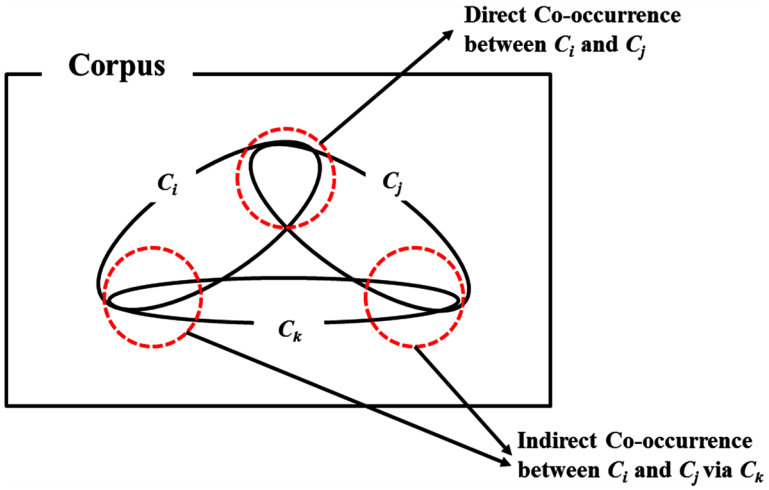
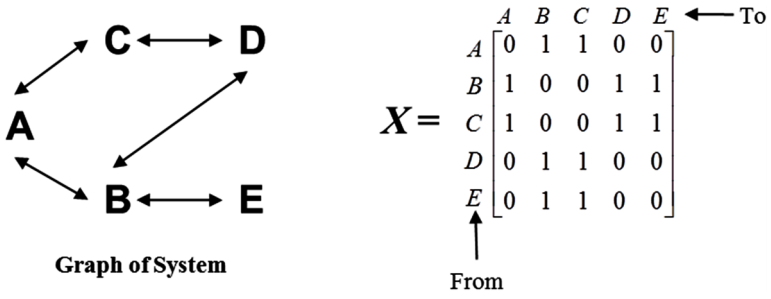


Figure 3. Sample graph of the system and its matrix representation



In matrix X, each element x_{ij} indicates a direct link between the nodes. Within this system, indirect links are assumed to exist whenever it is possible to connect two nodes through the intermediary nodes. The basic graph theory states that, under ordinary matrix multiplication procedures, the square of a direct linkage matrix, such as X, describes a graph whose paths all have a length of two steps. The sum of these paths between nodes i and j represents all the possible indirect linkages of the two steps (see Figure 4).

If the direct link matrix X is expanded and summed through successive powers until each of the elements of the summed matrix S converges ($S = X+X^2+X^3+\ldots+X^n$), then matrix S will contain elements s_{ij} representing the total direct influence (X) and indirect influence ($X^2+X^3+\ldots+X^n$) exerted between i and j over paths of increasing length up to n. This power expansion of X includes paths that may pass same nodes more than once (for the four-step path from node 1 to node 4, the sequence 1-3-1-4 could be an example).

This basic graph theory can be applied to the semantic network construction in our study. Each component of the squared first order co-occurrence matrix, M^2, is the sum of the product for indirect co-occurrence through other concepts. However, in calculating the indirect co-occurrence matrix, we use M_0, which has zero diagonal components instead of M (by definition of matrix M, all diagonal components of M should be one, while a general link matrix in graph theory or social network analysis has zero diagonal components). Thus, the second order co-occurrence matrix M_0^2 is defined as:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Squared matrix and meaning of its component (x_{23})

$$X^2 = \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} & x_{15} \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} \\ x_{31} & x_{32} & x_{33} & x_{34} & x_{35} \\ x_{41} & x_{42} & x_{43} & x_{44} & x_{45} \\ x_{51} & x_{52} & x_{53} & x_{54} & x_{55} \end{bmatrix} \times \begin{bmatrix} x_{11} & x_{12} & x_{13} & x_{14} & x_{15} \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} \\ x_{31} & x_{32} & x_{33} & x_{34} & x_{35} \\ x_{41} & x_{42} & x_{43} & x_{44} & x_{45} \\ x_{51} & x_{52} & x_{53} & x_{54} & x_{55} \end{bmatrix}$$

$$x_{23}^2 = x_{21}x_{13} + x_{22}x_{23} + x_{23}x_{33} + x_{24}x_{43} + x_{25}x_{53}$$

Sum of two step path from node 2 to node 3 through other nodes

2nd order co-occurrence matrix $M_0^2 \overset{\text{def}}{=} M_0 \times M_0 = \begin{pmatrix} m_{11}^2 & \cdots & m_{1n}^2 \\ \vdots & \ddots & \vdots \\ m_{n1}^2 & \cdots & m_{nn}^2 \end{pmatrix}$ (2)

The general order co-occurrence matrix can be defined as M_0^n

Based on these definitions, we can obtain the total co-occurrence matrix M_{total}, which contains both direct and indirect co-occurrence probabilities.

Total co-occurrence matrix[4] $M_{total} = M + M_0^2 + M_0^3 + \ldots M_0^n$ (3)

The *(i, j)* component of M_{total}, $m_{total,\ ij}$, can be used to calculate the relatedness between concept C_i and C_j in the following way:

$$rel\left(C_i, C_j\right) = \frac{m_{total,ij} + m_{total,ji}}{2} \quad (4)$$

where $rel(C_i, C_j)$ the relatedness between concepts and is a symmetric function.

We now define a 3-tuple semantic network data during the period from t_1 to t_2, $\boldsymbol{SN(t_1, t_2)}$:

$\boldsymbol{SN(t_1, t_2)} = \{corpus(t_1, t_2), C, M_{total}\}$,

where:

- $corpus(t_1, t_2)$ is the corpus generated during the time period from t_1 to t_2;
- $\boldsymbol{C} = \{C_1, C_2, \ldots, C_N\}$ is a collection of n concepts in which we are interested. This indicates the context;
- M_{total} is the total co-occurrence matrix.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

This network is constructed using a corpus generated in a given time interval with given concepts. If collection of concept C changes, the network also changes though the time interval is the same. This nature enables the constructed semantic network to reflect the dynamics of time and context. The selected concepts can create the context of a given situation by capturing the surrounding information. The notion of the context in this research is adopted from information retrieval research (Lawrence, 2000).

Detecting the Dynamics of a Semantic Network and Semantic Relatedness Between Concepts

Network dynamics, which is the study of changes in a network over time, is under study in various academic fields such as sociology, biology, computer science, and graph theory. As Wassermann & Faust (1994) points out, this is a difficult area that needs to be constantly studied in social networks. Furthermore, for time-series network data established in this study, all components of the matrix are written in floating numbers which makes it difficult to apply the methodology of an integer matrix such as the social network. This study proposes a new method to detect the changes in semantic network and demonstrates how the proposed methodology can be applied to the actual data.

Quantifying the Change of a Semantic Network

Bonacich (1972) suggests that the eigenvector of the largest eigenvalue of an adjacency matrix could make a good network centrality measure. Unlike the degree centrality, which weighs every contact equally, the eigenvector weighs contacts according to their centralities. Eigenvector centrality can also be seen as a weighted sum of not only direct connections but also indirect connections of every length. Thus, it takes into account the entire pattern in the network (Bonacich, 2007). Eigenvector centrality is defined as the principal eigenvector of the adjacency matrix defining the network. Equation (5) describes eigenvector centrality x in two equivalent ways, as a matrix equation and as a sum. The centrality of a node is proportional to the sum of the centralities of the nodes to which it is connected (Bonacich, 2007).

$$\lambda \boldsymbol{x} = \boldsymbol{A}\boldsymbol{x}, \lambda \boldsymbol{x}\boldsymbol{i} = \sum_{\boldsymbol{j}=1}^{\boldsymbol{n}} \boldsymbol{a}_{\boldsymbol{ij}} \boldsymbol{x}_{\boldsymbol{j}}, \boldsymbol{i} = 1, \ldots, \mathbf{n} \tag{5}$$

where $\boldsymbol{A}$ is the adjacency matrix of the graph, λ is the largest eigenvalue, n is the number of nodes, and x is the eigenvector.

In this study, we observe the changes of eigenvector value in a network matrix in order to detect the changes in network. Following Bonacichi's calculation of eigenvector centrality, we use the eigenvector corresponding to the dominant eigenvalue as the vector which represents the network characteristics. The importance of the eigenvector and its components has been dealt with significantly in social network research. The eigenvalue for each node and eigenvector may have been considered as important factors for detecting the dynamics of networks (McCulloh, 2009). In longitudinal social network analysis, the standard deviation of centrality, representing the change of topological connection among other nodes in a time series, can outperform other measures in terms of measurement for network dynamics (McCulloh & Carley, 2011). The change of eigenvector centrality for each concept definitely causes the change of dominant eigenvector of networks.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In this manner, from a network of N nodes, $N \times 1$ eigenvector will be extracted and the changes of this vector will represent the changes in the entire network. This $N \times 1$ column vector will measure the changes in the direction seen in the N-dimensional space to find the level of change in the network. If we denote the matrices representing networks A and B with N nodes as M_A and M_B, and the eigenvector of these two matrices as x_A, x_B, the gradient between the two networks, Θ can be expressed as follows:

$$\Theta = \cos^{-1}\left(\frac{x_A \cdot x_B}{\left|x_A\right|\left|x_B\right|}\right), \text{(unit: radian)} \tag{6}$$

If the value of Θ exceeds a particular threshold, we take it as a signal that the network characteristic has changed.

Dynamics of Semantic Relatedness between Concepts

The dynamics in the relatedness between concepts constituting the semantic network can be observed through the network matrix representation. If semantic relatedness measured by the above formula varies as time passes, components in the matrix may also change.

The semantic network with N nodes can be constructed with (N^2 - N) / 2 of $rel(C_i, C_j)$ operations in equation (4). In the next section, we present an experiment in which we track all semantic relation changes over a given time and a method of detecting changes is implemented.

EXPERIMENTAL EVALUATION

This section presents two experimental evaluations of our proposed technique, utilizing the sample corpus extracted from Yahoo News[5]. First experiments are dealing with dynamics of semantic relatedness among some political concepts about US election. The news corpus has been collected based on weekly scale during three months near the election period. Second experiment used the news corpus which has been collected throughout longer period, 2 years. We collected the corpus about Apple Company during two years and tracked the change of semantic relatedness for each month.

First Experiment: News on Political Domain

Yahoo News corpus is utilized as a sample corpus in the following way: Yahoo News Crawler is implemented using Python 2.7 and BeautifulSoup API (API for Web page parsing) to compose news corpus. The sample corpus includes Yahoo news articles on the presidential election published between June 2012 and September 2012. The detailed specifications of the corpus are shown in Table 2. Yahoo News corpus has been chosen as the sample data due to its prevalence in the online space and conjectured ease of verification of the result.

A seed set of concepts was needed in order to construct a semantic network between concepts. A total of 14 concepts used in this experiment are shown below.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Specification of the Yahoo News corpus

Data Collection Period	June 2012 ~ September 2012
Corpus Size	Over 7000 news articles (approximately 700 articles per week)
News Subject	News Articles about the Presidential Election
Extractor	Yahoo News Extractor

C = {Obama, election, Romney, tax, president, economy, healthcare, insurance, democratic, campaign, red, blue, poll, white house}

These concepts were randomly selected after reading news articles related to the U.S. presidential election. A number of studies (Leacock, Miller, & Chodorow, 1998; Wang, Du, Payattakool, Philip, & Chen, 2007) extracted seed concepts from reference taxonomy such as WordNet, which made the evaluation process of their proposed algorithms more convenient by comparing clustering results and the reference taxonomy. However, these studies deal with more comprehensive and general semantic relatedness between concepts whereas the aim of our study is to detect the dynamics of semantic relatedness that reflect either domain- or time-specific subjects. Since representative reference taxonomies such as WordNet and Roget's taxonomy describe the relation among millions of concepts under normal situations, regardless of specific context and period, they are not suitable for the purpose of evaluation in our approach.

Constructing the Co-Occurrence Matrix

A 1st order co-occurrence matrix M and its derivative matrices, including 2nd order, 3rd order and total co-occurrence matrix, were calculated from the Yahoo News corpus and the given set of concepts. In this experiment, the threshold for semantic relatedness has been set such that the two nodes within, less than, or equal to three hops are meaningfully related concepts, which is why the 4th order co-occurrence matrix and higher orders are not calculated.

Detecting the Dynamics of Semantic Relatedness and Semantic Networks

Table 3 shows the details of Yahoo News data for 13 weeks. Every Tuesday, the previous one-week's data was collected using Yahoo News Crawler. As a result, more than 700 news data related to the election in the United States were collected for every week. Using news corpus collected for every state, each weekly network was constructed, resulting in a total of 13 time-series network data.

For the 14 previously selected concepts, a total of 22 networks were formed, and each network was expressed as a 14×14 asymmetric matrix.

Using this matrix, we can calculate the semantic relatedness among concepts for each week. As we defined, formula (4) can be used for calculating semantic relatedness. Table 4 shows the semantic relatedness matrix for the first weeks. The relatedness matrix is generated from the total co-occurrence matrix, which is the grand sum of 1st, 2nd and 3rd co-occurrence matrices.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Specification of corpus used for constructing semantic network

Period	Total Number of News Articles	Real Date
Week 1	715	3rd Week of June
Week 2	705	4th Week of June
Week 3	697	1st Week of July
Week 4	721	2nd Week of July
Week 5	706	3rd Week of July
Week 6	712	4th Week of July
Week 7	698	1st Week of August
Week 8	704	2nd Week of August
Week 9	711	3rd Week of August
Week 10	708	4th Week of August
Week 11	714	1st Week of September
Week 12	709	2nd Week of September
Week 13	715	3rd Week of September
Week 14	720	4th Week of September
Week 15	732	1st Week of October
Week 16	722	2nd Week of October
Week 17	698	3rd Week of October
Week 18	720	4th Week of October
Week 19	712	1st Week of November
Week 20	730	2nd Week of November
Week 21	708	3rd Week of November
Week 22	707	4th Week of November

The semantic relatedness can be decomposed into two separated contributions; from 1st order co-occurrence and higher order (2nd and 3rd) co-occurrence matrix. This decomposition shows how indirect co-occurrence of two words can contribute to elevate the relatedness and enable the proposed relatedness to be measured more accurately and sophisticatedly. Contribution of higher order co-occurrence to semantic relatedness value is presented in Table 5.

As we can see from Table 5, most pairs have a very marginal portion from a higher order matrix. However, we can also observe that some relatedness values among important concepts are calibrated by higher order terms. For example, without considering indirect co-occurrence, the semantic relatedness between (campaign, Obama), (campaign, Romney), and (campaign, president) might have lower relatedness though they are intimately linked. Some concept pairs that are enjoying the effect of indirect co-occurrence are presented in Table 6 with the decomposition of contribution. This shows the effectiveness of proposed calculation method compared to a simple co-occurrence analysis, which has been used for corpus-based semantic related research.

For each network matrix, dominant eigenvalue and corresponding eigenvector were obtained (all operations related to the matrix were performed using R Package). All matrices had a dominant eigenvalue and the second biggest eigenvalue had a much smaller dominant eigenvalue (all the second biggest

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Semantic Relatedness (1st week, generated from the total co-occurrence matrix)

	Obama	**election**	**Romney**	**tax**	**president**	**economy**	**healthcare**	**insurance**	**democratic**	**campaign**	**red**	**blue**	**poll**
Obama													
election	0.532												
Romney	0.620	0.311											
tax	0.261	0.082	0.159										
president	0.757	0.473	0.485	0.183									
economy	0.342	0.125	0.205	0.079	0.267								
healthcare	0.379	0.060	0.180	0.080	0.257	0.056							
insurance	0.172	0.036	0.076	0.064	0.118	0.023	0.054						
democratic	0.348	0.170	0.205	0.037	0.296	0.053	0.024	0.020					
campaign	**0.663**	0.294	**0.522**	0.113	**0.481**	0.142	0.102	0.061	0.179				
red	0.354	0.174	0.257	0.143	0.294	0.122	0.109	0.110	0.130	0.250			
blue	0.423	0.120	0.263	0.028	0.275	0.033	0.007	0.010	0.065	0.162	0.199		
poll	0.504	0.134	0.314	0.068	0.330	0.071	0.085	0.026	0.074	0.185	0.161	0.035	
white house	0.357	0.132	0.176	0.050	0.314	0.075	0.055	0.018	0.069	0.141	0.136	0.033	0.071

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Contribution of higher order co-occurrence (1st week, generated from 1st and 2nd order co-occurrence matrix)

	Obama	**election**	**Romney**	**tax**	**president**	**economy**	**healthcare**	**insurance**	**democratic**	**campaign**	**red**	**blue**	**poll**
Obama													
election	0.341												
Romney	0.368	0.226											
tax	0.163	0.061	0.104										
president	0.375	0.261	0.324	0.134									
economy	0.211	0.078	0.131	0.034	0.169								
healthcare	0.221	0.060	0.131	0.027	0.171	0.030							
insurance	0.103	0.032	0.066	0.016	0.082	0.019	0.010						
democratic	0.242	0.079	0.147	0.034	0.184	0.043	0.024	0.015					
campaign	**0.393**	0.183	**0.249**	0.086	**0.339**	0.110	0.090	0.046	0.112				
red	0.223	0.133	0.166	0.061	0.190	0.080	0.085	0.038	0.084	0.161			
blue	0.264	0.070	0.151	0.028	0.211	0.033	0.007	0.010	0.030	0.113	0.087		
poll	0.282	0.098	0.173	0.043	0.236	0.053	0.029	0.021	0.049	0.147	0.107	0.035	
white house	0.223	0.082	0.141	0.036	0.170	0.044	0.028	0.018	0.044	0.111	0.081	0.033	0.052

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. Semantic Relatedness – decomposition of contribution from direct and indirect co-occurrence

	Obama, election	Romney, eletion	Obama, president	Romney, president	Obama, healthcare	campaign, Obama	campaign, Romney	campaign, president
Semantic relatedness	0.532	0.311	0.757	0.485	0.379	0.663	0.522	0.481
Contribution from 1st order	0.191	0.085	0.382	0.161	0.157	0.270	0.273	0.142
Contribution from higher order	0.341	0.226	0.375	0.324	0.221	0.393	0.249	0.339

eigenvalues were less than 1.0). Therefore, according to the process of constructing Bonacich's eigenvector centrality, these eigenvectors are enough to express the characteristics of each matrix (Resnik, 1995). The eigenvalues and eigenvectors for 22 networks are shown in Table 7.

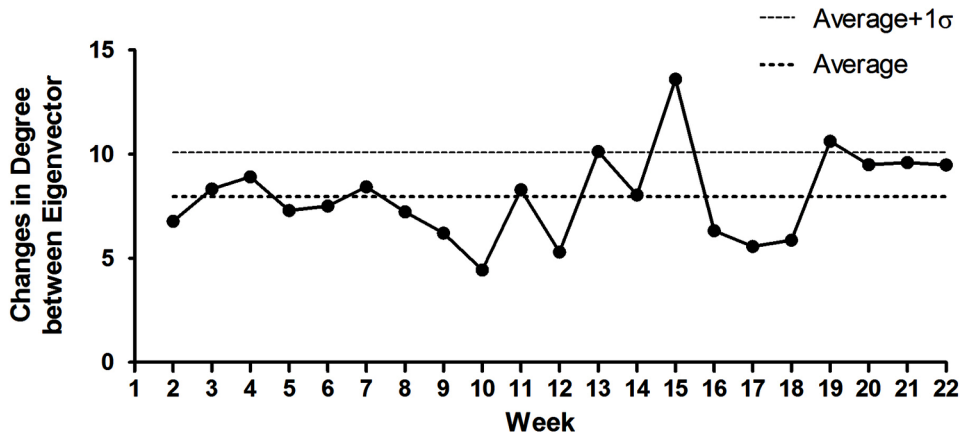
Using these eigenvectors, the gradient between networks, Θ can be obtained. The gradient can be calculated from week 2 via comparison with the previous week. The gradient Θ represents how much each network eigenvector moves in a 14-dimensional space. Table 7 exhibits the resulting eigenvector gradient, Θ.

Overall, the mean change level was 7.96 and the standard deviation was 2.12. If we assume the normal distribution for network change, a change that results in the position of 'average + 1σ' corresponds to the top 15%. In this experiment, the threshold level of network change detection is set at the top 15%. The resulting gradients and the threshold value (about 10.08) are shown in Figure 5.

From Figure 5, there is a significant change in network in weeks of 13, 15 and 19 when the change level is close to or exceeds the threshold value. Such change of network eigenvector at a significant level signifies that Bonacich's eigenvector centrality of semantic network has changed.

The dynamics of semantic relatedness between separate concepts can be verified in a simpler way. If we compute each component of total co-occurrence matrix and the average of its symmetry components, we can obtain the changes in semantic relatedness over time. Figure 6 shows the changes in semantic relatedness between concepts.

Figure 5. Detecting the significant change of network



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

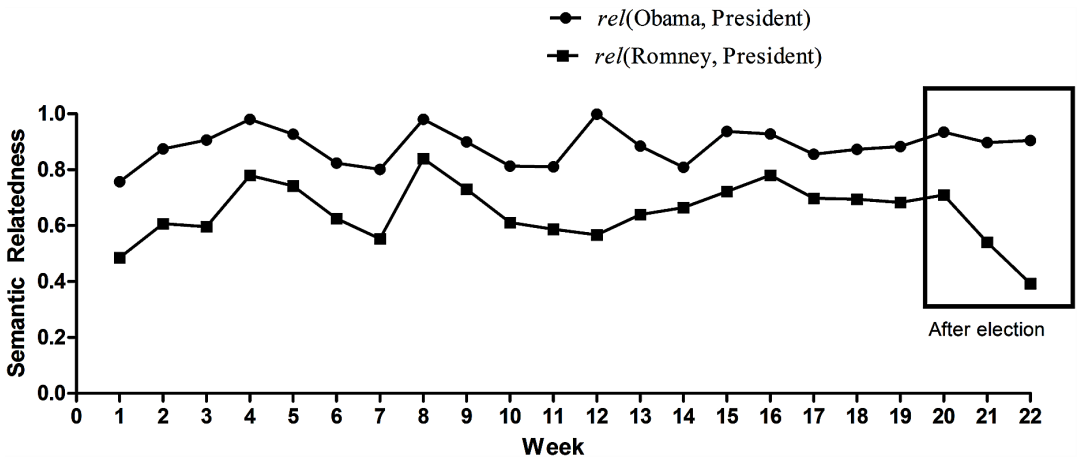
Table 7. Dominant Eigenvalue and Corresponding Eigenvector for the experiment

Week	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	
	Dominant Eigenvalue																						
	2.792	3.189	3.528	4.522	4.048	3.293	3.219	4.558	3.834	3.287	3.210	3.674	3.556	3.467	3.872	4.014	3.632	3.363	3.368	4.059	3.630	2.537	
	Dominant Eigen Vector																						**STD**
Obama	0.273	0.282	0.277	0.289	0.271	0.283	0.307	0.290	0.302	0.293	0.267	0.264	0.275	0.301	0.264	0.282	0.295	0.295	0.307	0.274	0.258	0.293	0.014
election	0.332	0.302	0.356	0.319	0.305	0.322	0.314	0.299	0.316	0.296	0.320	0.341	0.332	0.263	0.283	0.270	0.275	0.254	0.297	0.266	0.278	0.338	0.027
Romney	0.291	0.347	0.287	0.273	0.282	0.268	0.268	0.250	0.287	0.289	0.234	0.268	0.284	0.282	0.288	0.277	0.277	0.291	0.322	0.285	0.312	0.358	0.027
tax	0.169	0.170	0.172	0.262	0.243	0.196	0.242	0.265	0.274	0.234	0.243	0.187	0.241	0.227	0.275	0.243	0.205	0.237	0.171	0.196	0.189	0.242	0.034
president	0.287	0.312	0.296	0.304	0.287	0.290	0.305	0.292	0.304	0.309	0.287	0.289	0.287	0.305	0.271	0.288	0.287	0.289	0.329	0.283	0.290	0.321	0.013
economy	0.221	0.222	0.250	0.251	0.291	0.277	0.217	0.286	0.264	0.269	0.260	0.229	0.286	0.278	0.248	0.250	0.201	0.214	0.200	0.209	0.226	0.243	0.028
healthcare	0.263	0.225	0.267	0.190	0.263	0.240	0.216	0.259	0.227	0.237	0.331	0.292	0.269	0.331	0.270	0.289	0.324	0.367	0.229	0.321	0.312	0.237	0.044
insurance	0.118	0.104	0.123	0.153	0.130	0.199	0.195	0.165	0.141	0.154	0.167	0.156	0.136	0.185	0.224	0.174	0.145	0.141	0.097	0.176	0.140	0.110	0.031
democratic	0.252	0.237	0.286	0.280	0.256	0.307	0.224	0.270	0.250	0.238	0.286	0.312	0.329	0.274	0.226	0.278	0.271	0.248	0.303	0.245	0.299	0.244	0.028
campaign	0.372	0.421	0.380	0.346	0.357	0.358	0.373	0.325	0.362	0.336	0.341	0.351	0.354	0.328	0.330	0.328	0.338	0.317	0.347	0.328	0.336	0.314	0.023
red	0.173	0.150	0.146	0.187	0.155	0.162	0.181	0.182	0.174	0.182	0.182	0.195	0.139	0.176	0.160	0.205	0.185	0.175	0.170	0.185	0.207	0.157	0.017
blue	0.316	0.302	0.225	0.260	0.286	0.210	0.258	0.289	0.224	0.267	0.201	0.197	0.062	0.097	0.296	0.259	0.236	0.283	0.283	0.318	0.174	0.211	0.064
poll	0.318	0.254	0.287	0.251	0.289	0.308	0.271	0.233	0.267	0.287	0.271	0.286	0.314	0.301	0.323	0.298	0.308	0.294	0.289	0.292	0.308	0.275	0.022
white house	0.233	0.246	0.260	0.305	0.245	0.243	0.301	0.290	0.270	0.291	0.282	0.288	0.254	0.284	0.238	0.265	0.312	0.254	0.265	0.299	0.318	0.283	0.024
Changes in angle																							
	Radian	0.118	0.145	0.156	0.127	0.131	0.147	0.126	0.108	0.077	0.145	0.093	0.177	0.140	0.237	0.110	0.097	0.102	0.185	0.166	0.167	0.165	
	Degree	6.780	8.330	8.910	7.286	7.504	8.427	7.229	6.205	4.438	8.292	5.302	**10.12**	8.032	**13.59**	6.317	5.559	5.866	**10.61**	9.492	9.589	9.471	

Average of Change in Degree = 7.96, Standard Deviation of Change = 2.12 (All floating numbers are rounded)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Change of semantic relatedness between concepts – Obama and president / Romney and president



We selected two pairs of concepts, (Obama, President) and (Romney, President). Based on the methodology proposed in Section 4, semantic relatedness was calculated. As Romney was defeated in the US presidential election 2012, *rel*(Romney, President) decreased sharply after election 2012 whereas *rel*(Obama, President) maintained the high relatedness value. This result exactly corresponds with a change of human *ground truth*.

The level of dynamics of concepts in the semantic network can also be observed by checking the standard deviation of centrality, as McCulloh (2009) indicates. As we can see from the last column of table 7, some concepts have higher standard deviation than others. For example, "healthcare" and "insurance" have relatively higher dynamics as these were one of the key words in the US presidential election. This means the standard deviation of centrality can be used for detecting dynamic concepts or key topics in the semantic network by analyzing the news corpus.

In this experiment, we have tracked the changes of 91 pairs' semantic relatedness ($_{14}C_2 = 91$) and some concepts showed their dynamics in semantic relatedness while others showed very stable relatedness values regardless of time. Figure 7 below shows the change of semantic relatedness between "Romney" and other concepts.

As we can see from Figure 7, "Romney", which is one of the hottest topics in the context of this expriment, exhibits its dynamics of semantic relatedness with regard to the other concepts. Especially, semantic relatedness with concepts about election such as campaign, tax, and president fluctuates as time change. However, it is not the case that all concepts show such dynamics in semantic relatedness. Concepts which are less relevant and sensitive to the election show more stability in terms of dyanmics. Figure 8 shows different aspects of dyanmics of semantic relatedness.

In the case of semantic relatedness between "economy" and other concepts, most relatedness stay very stable in terms of change of relatedness. Only a few concepts such as campaign which are closely related to the election show higher change than other concepts.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Change of semantic relatedness between "economy" and other concepts

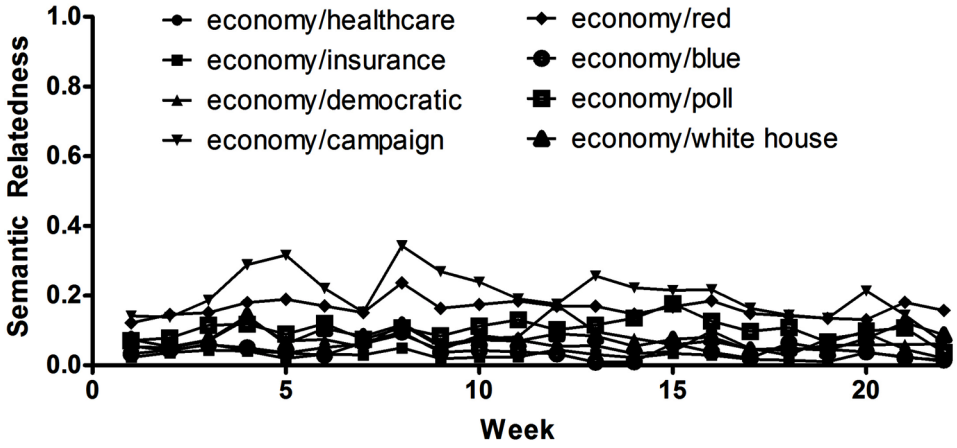
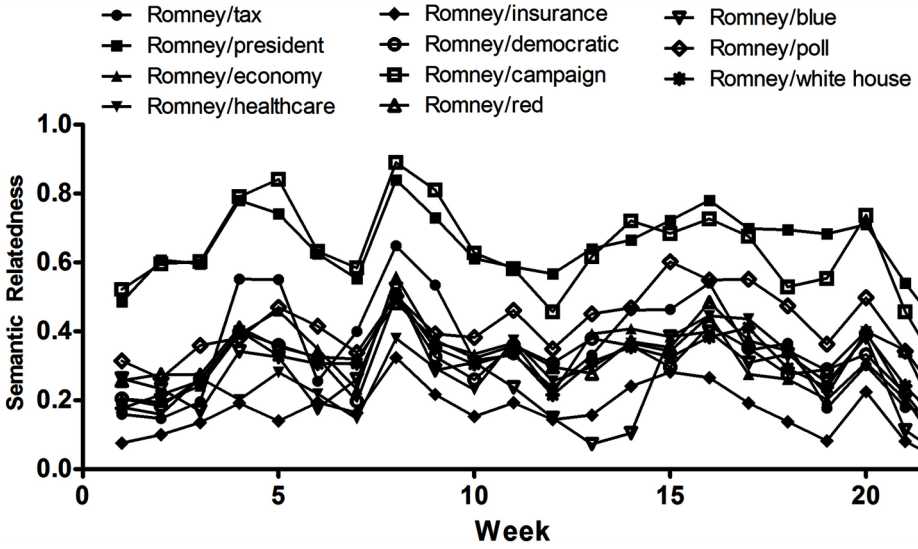


Figure 7. Change of semantic relatedness between "Romney" and other concepts



Second Experiment: News on IT Domain

For the second experiment, we collected IT news dealing with the article about Apple Company. The corpus has been composed using Nexis news database browser. The corpus includes Apple news articles on the between December 2013 and November 2015 (2 years). Nexis browser enabled us to confine the period and publication area of articles. In this experiment, we only used the article published in the U.S region during the defined period. The detailed specifications of the corpus are shown in Table 8. Apple News corpus has been chosen as the sample data due to its prevalence in the online space and conjectured ease of verification of the result.

A seed set of concepts was needed in order to construct a semantic network between concepts as we did in first experiment. A total of 10 concepts used in this experiment are shown below.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 8. Specification of the News corpus

Data Collection Period	December 2013 ~ November 2015
Corpus Size	Over 15,000 news articles (approximately 600 articles per month)
News Subject	News Articles about Apple Company
Extractor	Nexis news browser

$C = \{$Apple, Samsung, iPhone, new, Google, Android, Mobile, smartphone, technology, Yahoo$\}$

The concepts above was selected by five doctoral students who are majoring IT business in business school after they skimmed the recent articles about Apple.

Constructing the Co-Occurrence Matrix

A 1st order co-occurrence matrix M and its derivative matrices, including 2nd order, 3rd order and total co-occurrence matrix, were calculated from the Apple News corpus and the given set of concepts. In this experiment, as we did in the first experiment, the threshold for semantic relatedness has been set such that the two nodes within, less than, or equal to three hops are meaningfully related concepts, so the 4th order co-occurrence matrix and higher orders are not calculated.

Detecting the Dynamics of Semantic Relatedness and Semantic Networks

For each month, more than 600 articles published in the U.S. were collected. Using news corpus collected for every state, each monthly network was constructed, resulting in a total of 24 time-series network data. Table 9 shows the semantic relatedness among apples and other concepts in first month (December 2013) and how the results can be decomposed into direct and indirect components. We can see that some relevant concepts such as iPhone, Google, and Android are strongly compensated by contribution of higher order.

Figure 9 shows the change of semantic relatedness between Apple and other relevant concepts. Some relatedness has very stable value regardless of time flow but some concepts such as Samsung and iPhone show significant change in relatedness. As we can see in Figure 9, relatedness between Apple and Samsung/Google has very high value through March - April 2014 and May – June 2015. In this period, Apple had been suing over Samsung (March - April 2014) and its result of litigation has been released (May – June 2015).

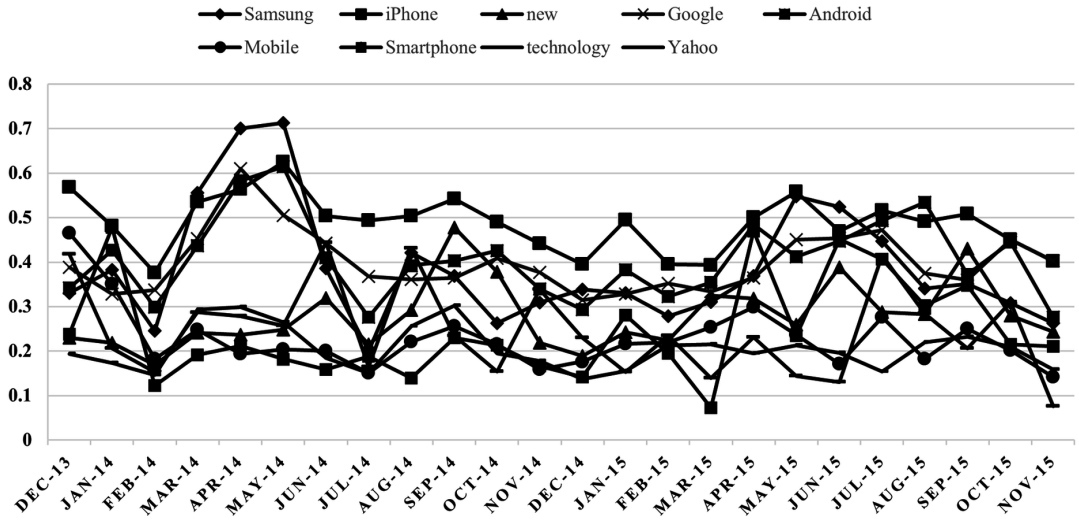
These anecdotal influence on semantic relatedness have an impact on overall semantic network as well. Table 10 shows the change of whole semantic network in terms of its eigenvectors. During these two periods when the lawsuit has been going on, the semantic network shows significant change due to these events.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 9. Semantic Relatedness – decomposition of contribution from direct and indirect co-occurrence

Apple &	Samsung	iPhone	new	Google	Android	Mobile	Smartphone	technology	Yahoo
Semantic relatedness	0.330	0.568	0.230	0.388	0.342	0.465	0.238	0.194	0.418
Contribution from 1st order	0.199	0.327	0.117	0.197	0.150	0.243	0.073	0.105	0.243
Contribution from higher order	0.131	**0.242**	0.112	**0.191**	**0.192**	0.222	0.166	0.089	0.175

Figure 9. The change of Semantic Relatedness between "Apple" and other concepts



DISCUSSION

The proposed approach to manage the dynamics and temporal properties of the semantic relatedness was verified through two experimental evaluations. The evaluations not only revealed the dynamics of semantic relatedness between specific words, but also reaffirmed the needs for dynamic semantic relatedness paradigm. In this regard, this research has both theoretical and practical implications and contributions.

The proposed method and evaluation result provide the empirical evidence to dynamic nature of semantic relatedness, which has its basis on the theory of cognitive semantics. The basic assumption of cognitive semantics is that the language spoken by human being is the part of the cognition (Saeed, 2003) and this assumption can hint the dynamic nature of relatedness among the concepts as human cognition can be varied with their experience and knowledge (Löbner, 2013). These experience and knowledge has been called *semantic knowledge* in traditional semantics and this semantic knowledge has been considered as a main factors for human to understand the world (Holland and Quinn, 1987). Our research has very significant contributions to traditional semantics and cognitive semantics, which provide the theoretical background to dynamic semantic relatedness paradigm, by providing clear empirical evidence of changes in semantic relatedness and its management strategies.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 10. Dominant Eigenvalue and Corresponding Eigenvector for second experiment

Month	Dec-13	Jan-14	Feb-14	Mar-14	Apr-14	May-14	Jun-14	Jul-14	Aug-14	Sep-14	Oct-14	Nov-14	Dec-14	Jan-15	Feb-15	Mar-15	Apr-15	May-15	Jun-15	Jul-15	Aug-15	Sep-15	Oct-15	Nov-15	
			Dominant Eigenvalue																						
	3.69	4.079	2.799	3.842	4.22	4.218	3.941	2.953	3.72	3.86	3.511	3.186	3.876	3.272	3.48	3.221	4.13	3.757	4.208	4.135	3.931	3.873	3.562	2.96	
			Dominant Eigen Vector																						**STD**
Apple	0.546	0.509	0.532	0.544	0.553	0.572	0.517	0.542	0.563	0.536	0.548	0.57	0.561	0.555	0.49	0.552	0.535	0.543	0.536	0.553	0.509	0.517	0.559	0.564	0.021
Samsung	0.227	0.286	0.262	0.329	0.376	0.387	0.253	0.23	0.291	0.281	0.234	0.243	0.248	0.28	0.233	0.299	0.281	0.353	0.368	0.296	0.265	0.264	0.258	0.265	0.046
iPhone	0.382	0.329	0.35	0.321	0.319	0.354	0.298	0.391	0.337	0.398	0.384	0.356	0.324	0.388	0.295	0.319	0.321	0.357	0.268	0.309	0.339	0.37	0.324	0.358	0.033
new	0.205	0.199	0.271	0.231	0.182	0.173	0.296	0.288	0.259	0.366	0.349	0.28	0.31	0.268	0.249	0.338	0.282	0.253	0.31	0.245	0.284	0.36	0.291	0.294	0.051
Google	0.315	0.295	0.365	0.362	0.387	0.333	0.398	0.363	0.341	0.274	0.325	0.351	0.343	0.293	0.417	0.339	0.286	0.35	0.347	0.333	0.34	0.297	0.356	0.371	0.035
Android	0.279	0.358	0.369	0.344	0.387	0.388	0.345	0.36	0.317	0.327	0.382	0.327	0.353	0.371	0.346	0.369	0.357	0.335	0.339	0.345	0.4	0.314	0.359	0.315	0.027
Mobile	0.333	0.28	0.249	0.212	0.135	0.15	0.189	0.181	0.183	0.208	0.194	0.166	0.170	0.21	0.247	0.261	0.23	0.198	0.154	0.205	0.187	0.209	0.189	0.201	0.043
Smartphone	0.244	0.429	0.192	0.186	0.151	0.15	0.177	0.261	0.145	0.216	0.216	0.197	0.218	0.291	0.281	0.057	0.382	0.254	0.352	0.308	0.28	0.331	0.221	0.278	0.082
technology	0.148	0.13	0.186	0.204	0.186	0.174	0.147	0.154	0.201	0.212	0.161	0.172	0.159	0.143	0.221	0.213	0.134	0.175	0.137	0.119	0.178	0.182	0.174	0.182	0.027
Yahoo	0.305	0.128	0.224	0.263	0.207	0.174	0.353	0.176	0.326	0.157	0.117	0.293	0.248	0.136	0.274	0.149	0.163	0.138	0.085	0.264	0.233	0.156	0.247	0.094	0.074
	Changes in angle																								
	Radian	0.290	0.286	0.100	0.132	0.289	0.083	0.228	0.221	0.234	0.114	0.204	0.147	0.217	0.240	0.180	0.298	0.347	0.169	0.223	0.121	0.156	0.184	0.175	
	Degree	**16.60**	**16.35**	5.744	7.568	**16.53**	4.761	13.09	12.67	13.41	6.526	11.66	8.436	12.40	13.72	**10.31**	**17.09**	**19.88**	9.709	12.75	6.910	8.933	10.56	10.04	

Average of Change in Degree = 11.55, Standard Deviation of Change = 3.92 (All floating numbers are rounded

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Also, the research can be applied for various practical situation where the management of dynamic semantic relatedness is needed. The keyword recommendation and service discovery in virtual community can be an application example of proposed method (Crasso, Zunino, & Campo, 2011). As the main interest and topic among the virtual community users may vary, the relatedness among keywords and query can be varied as well. In this case, the proposed dynamic network model can cope with and manage the dynamics of relatedness among keywords and the quality of topic recommendation can be improved.

These contributions are very meaningful in the context of rapidly changing IT society with the active knowledge sharing and information diffusion. While the structure of human cognition can vary very fast with IT infrastructure, its theoretical consideration and method for dealing with dynamics of semantic relatedness has been lacking. Therefore, this research can be the very significant starting point for *dynamic semantic relatedness paradigm* research.

CONCLUSION AND FUTURE RESEARCH

This research focuses on dynamics in semantic relatedness and semantic networks under the *dynamic semantic relatedness paradigm*. For this purpose, new semantic relatedness measures can be defined based on the dynamic semantic network construction using a corpus created in a specific time period and context. Not only the change of semantic relatedness among concepts, but also the change of the whole semantic network can be measured. Specifically, the change of a semantic network consisting of given concepts has been quantified based on the fundamental concept of network dynamics in social network research.

Compared to previously dominant approaches following the *static semantic relatedness paradigm*, the proposed approach has significant advantages. First, our approach does not have any dependency on external resources such as WordNet, dictionaries, and thesauri. As most dictionaries contain limited vocabulary, newly generated concepts and proper nouns cannot be referred from dictionaries. Considering that content and contexts are changing rapidly due to technological advances, extracting semantic networks using only a given corpus can be more flexible in terms of reflecting the real word situation.

Second, the proposed method is simpler than other statistical approaches adopted by studies dealing with dynamics of semantic relatedness (Blei & Lafferty, 2006; Radinsky, Agichtein, Gabrilovich, & Markovitch, 2011; Sagi, Kaufmann, & Clark, 2009) such as LSA, TSA and LDA. By simply constructing semantic networks based on conditional probability, we can measure semantic relatedness. The simplicity of measurement is very critical for implementing real-time systems. Though the examples in this research were weekly-based and monthly-based detection of semantic relatedness, our approach can be applied to a large amount of data generated within a short period of time. In such case, simple algorithms are highly desirable for reducing the calculation cost.

Third, this research proposed the calculation method of semantic relatedness, but also sought to quantify the dynamics of whole semantic networks. Implied by social network research (McCulloh, 2009; McCulloh & Carley, 2011), the dynamics of one concept in the network can also be observed via the standard deviation of centrality. To the best of our knowledge, this is the first attempt to quantify the dynamics of semantic networks. The quantification is a more advanced technique than finding the existence of dynamics. With the proposed technique, we can compute the degree of dynamics of a given semantic network so that we can compare the dynamics with respect to the time change.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

These strengths are verified with the experimental evaluation using real corpus data gleaned from Yahoo News. The evaluation is an important aspect of semantic relatedness research. Methods of previous research following the *static semantic relatedness paradigm* evaluate their approach with benchmark testing by comparing human ratings (Resnik, 1995). These evaluation methods usually deal with the relatedness among very common concepts whose relationships can be estimated using common sense and ground truth. However, evaluation of semantic relatedness reflecting the context or dynamics can be more difficult as it cannot be evaluated by using only the human ratings. In this case, only humans who are involved in the given context in a specific time have the knowledge to estimate the relatedness among relevant concepts. Due to such difficulties in evaluation of dynamic semantic relatedness, previous research dealing with dynamic semantic relatedness has evaluated their methods using the assessment method for static semantic relatedness evaluation; only common sense by human is used for evaluation without reflecting the knowledge generated by context during specific time periods. For example, Radinsky, Agichtein, Gabrilovich, & Markovitch (2011), Rinaldi (2009) and Sagi, Kaufmann, & Clark (2009) evaluate their methods by selecting common words that hardly change their semantics and compare with the human rating even though they propose temporal semantic analyses. Other works do not even contain the evaluation results and only provide the list of topics (Blei & Lafferty, 2006).

The evaluation method of this research is hinted from those of information retrieval. To evaluate the correctness of the retrieval system output, the output must be compared with ground truth (Rinaldi, 2009). Similarly, the dynamics of semantic relatedness can be evaluated with the change in *ground truth*. For this reason, we selected concepts with the change in *ground truth*; the change of president can cause the change in *ground truth* in human recognition. Our experiment results exactly correspond with ground truth that has been changed by presidential election.

Throughout the experiment, we tracked the change of relatedness of 92 pairs' concepts to get more concrete reliability of proposed semantic relatedness calculation method where the impact of indirect co-occurrence and significant change have been confirmed. By showing multiple cases of dynamic change of semantic relatedness, we tried to verify the general applicability of proposed method to other contexts.

The experiment results show applicability of a proposed method to another area. As changes in various network measures such as centrality, closeness, and between-ness in longitudinal social network analysis have important meanings (McCulloh & Carley, 2011), these network measures in semantic networks might be valuable in terms of dynamics of semantic relatedness. We have already discovered the potential of centrality variation for finding the concepts that have the most dynamic nature in a given context.

Our proposed approach can be used for rising topic detection and tracking large streams of data such as newswire and broadcast news since news topics and relationships among topics change as time passes. Existing studies of the static semantic relatedness paradigm cannot capture the dynamic nature of semantic relatedness between topics in news and media. Keyword and tag recommendation is one applicable research area. Tagging has become a standard way of organizing information on the Web, particularly in various types of data created by communities of users. Tags and keywords attached to each resource create a bridge between data and users accustomed to keyword-based search and browsing. Our research can be utilized to provide users with a set of user-defined tags. As the relatedness between tags/keywords changes easily as time passes, our approach can be an excellent alternative to existing tag/keyword recommendation algorithms.

There are at least three ways to improve our approach towards the dynamic semantic relatedness paradigm. First, a statistically significant level of network change should be detected through corpus analysis on diverse subjects. By testing the proposed algorithm with various types of corpus and keyword sets,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

one can derive more generalized statistics representing dynamics of semantic relatedness. Second, the ordinary statistical features of semantic network dynamics should be studied through data collection over a longer period of time. Compared to the ordinary social network, the semantic network lacks studies on the properties of the network itself. Third, the efficiency of proposed algorithms should be assessed for real-time detection of semantic relatedness dynamics. For real-time processing, more efficient ways of calculating semantic relatedness and its dynamics need to be in place.

ACKNOWLEDGMENT

This study was supported by research grants from the Institute of Management Research at the Seoul National University.

REFERENCES

Blei, D. M., & Lafferty, J. D. (2006). Dynamic topic models. *Proceedings of the 23rd International Conference on Machine learning* (pp. 113-120). ACM.

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification. *The Journal of Mathematical Sociology*, *2*(1), 113–120. doi:10.1080/0022250X.1972.9989806

Bonacich, P. (2007). Some unique properties of eigenvector centrality. *Social Networks*, *29*(4), 555–564. doi:10.1016/j.socnet.2007.04.002

Couto, F. M., Silva, M. J., & Coutinho, P. M. (2007). Measuring semantic similarity between Gene Ontology terms. *Data & Knowledge Engineering*, *61*(1), 137–152. doi:10.1016/j.datak.2006.05.003

Crasso, M., Zunino, A., & Campo, M. (2011). A Survey of Approaches to Web Service Discovery in Service-Oriented Architectures. *Journal of Database Management*, *22*(1), 102–132. doi:10.4018/jdm.2011010105

Doddington, G. (2002). Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. *Proceedings of the Second International Conference on Human Language Technology Research* (pp. 138-145). Morgan Kaufmann Publishers Inc. 10.3115/1289189.1289273

Gan, M., Dou, X., & Jiang, R. (2013). From ontology to semantic similarity: Calculation of ontology-based semantic similarity. *TheScientificWorldJournal*, 2013. PMID:23533360

Hale, M. M. (1998). A comparison of WordNet and Roget's taxonomy for measuring semantic similarity. arXiv preprint cmp-lg/9809003

Harispe, S., Sánchez, D., Ranwez, S., Janaqi, S., & Montmain, J. (2014). A framework for unifying ontology-based semantic similarity measures: A study in the biomedical domain. *Journal of Biomedical Informatics*, *48*, 38–53. doi:10.1016/j.jbi.2013.11.006 PMID:24269894

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hoffman, P., Ralph, M. A. L., & Rogers, T. T. (2013). Semantic diversity: A measure of semantic ambiguity based on variability in the contextual usage of words. *Behavior Research Methods*, *45*(3), 718–730. doi:10.375813428-012-0278-x PMID:23239067

Holland, D., & Quinn, N. (1987). *Cultural models in language and thought*. Cambridge University Press. doi:10.1017/CBO9780511607660

Jarmasz, M., & Szpakowicz, S. (2004). Roget's Thesaurus and Semantic Similarity1. In N. Nicolov, K. Bontcheva, G. Angelova, & R. Mitkov (Eds.), *Recent Advances in Natural Language Processing III: Selected Papers from RANLP 2003, Current Issues in Linguistic Theory* (Vol. 260, pp. 111–120). John Benjamins. doi:10.1075/cilt.260.12jar

Jiang, J. J., & Conrath, D. W. (1997). Semantic similarity based on corpus statistics and lexical taxonomy. arXiv preprint cmp-lg/9709008

Kilgarriff, A., & Grefenstette, G. (2003). Introduction to the special issue on the web as corpus. *Computational Linguistics*, *29*(3), 333–347. doi:10.1162/089120103322711569

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2-3), 259–284. doi:10.1080/01638539809545028

Lawrence, S. (2000). Context in web search. *IEEE Data Eng. Bull.*, *23*(3), 25–32.

Leacock, C., Miller, G. A., & Chodorow, M. (1998). Using corpus statistics and WordNet relations for sense identification. *Computational Linguistics*, *24*(1), 147–165.

Lee, D., Kim, H., & Lim, M. (2016, January). A hierarchical model for real-time massive reaction visualization based on anatomical semantics. *Proceedings of the 2016 International Conference on Big Data and Smart Computing (BigComp)* (pp. 157-162). IEEE.

Lee, J. H., Kim, M. H., & Lee, Y. J. (1993). Information retrieval based on conceptual distance in IS-A hierarchies. *The Journal of Documentation*, *49*(2), 188–207. doi:10.1108/eb026913

Lin, D. (1998, July). *An information-theoretic definition of similarity* (Vol. 98, pp. 296–304). ICML.

Löbner, S. (2013). *Understanding semantics*. Routledge.

Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. Cambridge, Massachusetts: MIT Press.

Martin, D. I., Martin, J. C., & Berry, M. W. (2016). The Application of LSA to the Evaluation of Questionnaire Responses. In *Unsupervised Learning Algorithms* (pp. 449–484). Springer International Publishing. doi:10.1007/978-3-319-24211-8_16

McCulloh, I. (2009). *Detecting changes in a dynamic social network*. ProQuest.

McCulloh, I., & Carley, K. M. (2011). *Detecting change in longitudinal social networks. Military Academy West Point NY Network Science Center*. NSC.

Milikic, N., Jovanovic, J., & Stankovic, M. (2011). Discovering the Dynamics of Terms' Semantic Relatedness through Twitter. In # MSM (pp. 57-68).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ohsawa, Y., Benson, N. E., & Yachida, M. (1998). KeyGraph: Automatic indexing by co-occurrence graph based on building construction metaphor. *Proceedings of the IEEE International Forum on Research and Technology Advances in Digital Libraries ADL '98* (pp. 12-18). IEEE.

Rada, R., & Bicknell, E. (1989). Ranking documents with a thesaurus. *Journal of the American Society for Information Science*, *40*(5), 304–310. doi:10.1002/(SICI)1097-4571(198909)40:5<304::AID-ASI2>3.0.CO;2-6 PMID:10303917

Rada, R., Mili, H., Bicknell, E., & Blettner, M. (1989). Development and application of a metric on semantic nets. *Systems, Man and Cybernetics*. *IEEE Transactions on*, *19*(1), 17–30.

Radinsky, K., Agichtein, E., Gabrilovich, E., & Markovitch, S. (2011). A word at a time: computing word relatedness using temporal semantic analysis. *Proceedings of the 20th International Conference on World Wide Web* (pp. 337-346). ACM. 10.1145/1963405.1963455

Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxonomy. arXiv preprint cmp-lg/9511007

Rinaldi, A. M. (2009). An ontology-driven approach for semantic information retrieval on the web. *ACM Transactions on Internet Technology*, *9*(3), 10. doi:10.1145/1552291.1552293

Roark, B., & Charniak, E. (1998). Noun-phrase co-occurrence statistics for semiautomatic semantic lexicon construction. *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics* (Vol. 2, pp. 1110-1116). Association for Computational Linguistics.

Saeed, J. I. (2003). Semantics (Introducing Linguistics). In *Semantics-introducing linguistics*.

Sagi, E., Kaufmann, S., & Clark, B. (2009). Semantic density analysis: Comparing word meaning across time and phonetic space. *Proceedings of the Workshop on Geometrical Models of Natural Language Semantics* (pp. 104-111). Association for Computational Linguistics. 10.3115/1705415.1705429

Sheth, A., Aleman-Meza, B., Arpinar, I. B., Bertram, C., Warke, Y., Ramakrishanan, C., ... Kochut, K. (2005). Semantic association identification and knowledge discovery for national security applications. *Journal of Database Management*, *16*(1), 33–53. doi:10.4018/jdm.2005010103

Turney, P. (2001). Mining the web for synonyms: PMI-IR versus LSA on TOEFL. In L. De Raedt, & P. Flach, (Eds.), *Proceedings of the 12th European Conference on Machine Learning ECML '01* (pp. 491-502). Berlin, Heidelberg: Springer-Verlag.

Wang, J. Z., Du, Z., Payattakool, R., Philip, S. Y., & Chen, C. F. (2007). A new method to measure the semantic similarity of GO terms. *Bioinformatics (Oxford, England)*, *23*(10), 1274–1281. doi:10.1093/bioinformatics/btm087 PMID:17344234

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge, United Kingdom: Cambridge University Press. doi:10.1017/CBO9780511815478

Wu, Z., & Palmer, M. (1994). Verbs semantics and lexical selection. *Proceedings of the 32nd Annual Meeting of the Associations for Computational Linguistics (ACL '94)* (pp. 133-138). Las Cruces, New Mexico. 10.3115/981732.981751

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Zhao, M., Wang, H., Cao, L., Zhang, C., Yin, H., & Xu, F. (2015, December). LSIF: A System for Large-Scale Information Flow Detection Based on Topic-Related Semantic Similarity Measurement. *Proceedings of the 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology* (WI-IAT) (Vol. 1, pp. 417-424). IEEE. Chicago.

ENDNOTES

1. WordNet is a large lexical database of English. Nouns, verbs, adjectives and adverbs are grouped into sets of cognitive synonyms, each expressing a distinct concept (http://wordnet.princeton.edu)
2. We do not intend to limit the number of concepts in $\boldsymbol{C}$. N can be either the total number of concepts in the corpus or the number of given concepts.
3. The difference between conditional and unconditional probability can be verified using a statistical test, such as *t*-test.
4. Diagonal values of the total co-occurrence matrix are meaningless since they represent the relatedness between same concepts.
5. http://news.yahoo.com

This research was previously published in the Journal of Database Management (JDM), 27(2); edited by Keng Siau; pages 1-26; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 53
Emotion Mining Using Semantic Similarity

Rafiya Jan
Central university Of Kashmir, India

Afaq Alam Khan
Central University of Kashmir, India

ABSTRACT

Social networks are considered as the most abundant sources of affective information for sentiment and emotion classification. Emotion classification is the challenging task of classifying emotions into different types. Emotions being universal, the automatic exploration of emotion is considered as a difficult task to perform. A lot of the research is being conducted in the field of automatic emotion detection in textual data streams. However, very little attention is paid towards capturing semantic features of the text. In this article, the authors present the technique of semantic relatedness for automatic classification of emotion in the text using distributional semantic models. This approach uses semantic similarity for measuring the coherence between the two emotionally related entities. Before classification, data is pre-processed to remove the irrelevant fields and inconsistencies and to improve the performance. The proposed approach achieved the accuracy of 71.795%, which is competitive considering as no training or annotation of data is done.

INTRODUCTION

Social networks have become a popular interaction platform. Social Networking Sites like Facebook, Twitter has changed the standard of social interaction among the online users. Social Networking Sites enables online users to share their ideas, experiences, feelings, and emotions related to their daily routines. Social Networks are considered the most abundant sources of affective information for emotion analysis, affective computing and sentiment analysis. Sentiment analysis deals with determining the appraisals, feelings, opinions about different entities. Emotion analysis deals in determination and identification of feelings or emotions like Anger, Disgust, Fear, Happiness, Sadness, and Surprise. The goal of this

DOI: 10.4018/978-1-7998-0951-7.ch053

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

research is to create an automated system able to detect emotions in text using Semantic similarity between two textual elements. Our system creates a knowledge base for classification and then uses that knowledge base to infer classes of unseen data.

Emotion is defined as the mental state observed by behavioral or developmental changes. Emotions are universal despite the vast differences between different individuals and their cultures. The social conditions also induce emotions. Emotions are not bounded by the person's inner conscience but also afflicted by the person's experiences, goals and most importantly by the society. Emotions are a fundamental part of human nature possessed by every human. Emotions can be broadly classified into two types: Basic emotions and Complex Emotions. Basic emotions include Anger, Disgust, Fear, Happiness, Sadness, and Surprise. Complex emotion encompasses two or more emotions felt by a person in a particular situation. An accurate emotion detection system can be thus very beneficial for social orientation. Therefore, the need of the hour is a reliable and scalable emotion detection system capable of mapping human wrote social media posts into a different set of emotions for tracking the well-being of the society and provide benefits to the business establishments and political systems. The creation of automatic emotion detection system would thus allow us to evaluate and identify the people's emotion by automatically analyze the social media content like twitter posts and Facebook comments. The automation emotion detection system would thus be having one module for analyzing the social web for the social content and tracking the well-being of the individual. Social media posts also contain the geographical and temporal information; the system will hence be able to track the well-being state of a social group at some particular place, in some specific temporal area.

The relevance of this paper is found as the vast number of applications that are being used to detect the emotional states of people and the benefit they are providing to the business establishments, political parties, individuals and society at large. Its relevance can also be seen in the e-learning environment (Rodriguez, Ortigosa, & Carro, 2012), suicide prevention (Suero Montero, & Suhonen, 2014), depression detection (Desmet, & Hoste, 2013), detection of cyber-bullying (Vaassen, 2014) or tracking the well-being of user community (Dadvar, Trieschnigg, Ordelman, & de Jong, 2013).

The primary challenge for the automatic emotion detection systems is that they do not capture the semantic meaning of the textual systems; instead, they focus on the statistical features present in the text. Traditionally, these systems employ either lexical resources or use statistical measures to build the automatic emotion detection systems. However, none of these systems would employ the measures to focus on the semantics of the text. Moreover, with the employment of these resources the emotion detection systems would have bias and thus classification of emotions is not accurate and thus the emotion detection systems are not scalable, robust and are thus error-prone.

For this reason, we propose an Automatic emotion detection system which is capable of exploiting the semantic similarity between textual elements and thus classify text based on semantic relatedness measure. The proposed system consists of three main steps: 1) creation of seed documents using NRC Emotion Lexicon (Mohammad, & Turney, 2013) which is used to evaluate semantic association of the textual element to be classified and emotion class; 2) Generation of Big-Vectors of text using Word2Vec (Mikolov, Chen, Corrado, & Dean, 2013); 3) Using semantic relatedness measure to calculate the semantic association between seed documents and textual element to compute emotion class of the textual element.

We measured the effectiveness of the proposed automatic emotion classifier using pre-labeled standard datasets and found comparable results. Thus, proposed automatic emotion classifier be built keeping the semantics of text in consideration which was lacking in the previous start of art research conducted in the area of automatic emotion classification. Moreover, the results of our proposed technique confirm the efficacy of the proposed technique of emotion detection.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

RELATED WORK

This section summarizes the state of art approaches used in the emotion detection systems, their features, how they have built, their accuracies and their deficiencies. Also, this section presents the overview of the emotion models, emotion lexicons used by the automatic emotion detection systems. Majority of the automatic emotion detection and sentiment analysis systems are built around the statistical properties of the text and ignore the semantics and contextual feature which are essential for any emotion detection and sentiment analysis systems. Supervised learning algorithms are majorly used to detect and analyze emotions without taking into account the semantics and context into account.

Emotion Models

Ekman (1992) proposed the basic emotion model. This model identified six types of emotions: Anger, Disgust, Fear, Happiness, Sadness, and Surprise that are assumed as the basic emotion types. Plutchik, & Kellerman (1980) proposed the emotion model with eight primary emotion including Ekman's six basic emotions with two additional emotions Trust and Anticipation. He assembled these eight emotions in a Wheel. The radius of the wheel specifies the intensity of the emotions. The emotion closer to center indicates the sharp intensity of emotions. Plutchik also proposed that these emotions form 4 variant pairs like anticipation &surprise, anger, and fear, joy, and sadness, trust and disgust. Parrot's hierarchy model has categorized emotions into six classes of anger, fear, joy, love, sad, and surprise (Bradley, M. M., & Lang, P. J. (1999)). Izard in his model divides emotions into 12 classes viz. anger, contempt, disgust, fear, joy, interest, self-hostility, shame, shyness surprise, sadness and guilt (Hancock, Landrigan, & Silver, 2007). PANA (Positive Activation-Negative Activation) model (Watson, & Tellegen, 1985) resolved the system into positive affect and negative affect. The vertical axis depicting the positive affect and the horizontal axis depicting the negative affect. Russell's Circumflex Model (Russell, 1980) suggests that emotions are distributed in a 2D circular space in which the vertical axis represents the subjectivity and horizontal axis represent the intensity and the origin represents neutral as shown in Figure 2. The vertical dimensions specify how much PLEASANT and UNPLEASANT an emotion is. The horizontal dimensions distinguish ACTIVATION and DEACTIVATION. PAD (pleasure, arousal, and dominance) model (Mehrabian, A. (1996).) model is a superset of the circumflex model with the addition of the dominance dimension, which signifies whether impression of subject is in of a particular situation or not.

Emotion Lexicons

Emotion lexicon is a database of words labeled with labels of emotion according to their emotional connotations. Labels are generally emotional categories assigned to the word in the lexicon. In literature we find the following standard lexicons:

1. ANEW (Bradley, & Lang, 1999): Affective Norms for English Words is a set of standard emotional ratings of 1,034 words in English language including adjectives, nouns, and verbs.720 members participated for rating these words. The words are assigned scores from 1 to 9 on the basis of scope of 3 elements viz arousal, dominance, and valence by employing the SAM (Self-Assessment Manikin).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. DAL (Mehrabian, 1996): DAL is the lexicon of emotion words composed 8,742 words weighted by people for their imagination, awareness, interpretations, and intelligence and personification purposes.
3. WordNet Affect (WNA) (Strapparava, & Mihalcea, 2008): WordNet-Affect is an extension of WordNet realm that is composed of synsets that are relevant to express the affective concepts associated with emotion words. The synsets that are labelled with A-label distinguishes the emotive concepts. Other a-labels depicts the emotive responses, affects evoked by different situations and behaviour.
4. NRC Word-Emotion Association Lexicon (EmoLex) (Mohammad, & Turney, 2013):The NRC Emotion Lexicon is composed of 14,182 unigrams English words categorized into eight basic emotions (anger, anticipation, disgust, fear, joy, trust, sadness and surprise) and two sentiments (positive and negative).In this experiment, only 6 emotions (anger, disgust, fear, joy, sadness, and surprise) are used. Crowdsourcing on Mechanical Turk is used for manual word level annotation of these lexicons.
5. NRC Hashtag Emotion Lexicon (Mohammad, & Kiritchenko, 2015): NRC is a Twitter dataset composed of 16,000 words that are automatically annotated with Plutchik's eight emotions by employing the words bearing emotion hashtags.
6. EmoSenticNet (ESN) (Poria, SGelbukh, Hussain, Howard, Das, & Bandyopadhyay, 2013): ESN is the morphological asset of 13,171 words that ascribes subjective emotion labels and significant polarity scores to SentiNet approaches. WordNet Affect labels comprised of Ekman's emotions is the set of emotions that is used for annotating the concepts.
7. DepecheMood (Staiano, & Guerini, 2014): DepecheMood is like a dictionary of nearly 37,000 words in which every word is conjoined with emotion scores of 8 emotion categories viz afraid, amused, angry, annoyed, inspired, don't care, happy and sad. It was assembled by manipulating the crowd-sourced emotion labelling scheme implemented by the news readers of Rappler website.

Table 1 shows summarized emotion lexicons developed for automatic emotion detection systems.

Computational Approaches

Machine Learning Based Approach

Machine learning based approaches use statistical algorithms on syntactic and linguistic features for building automatic emotion detection systems. These algorithms are either supervised or unsupervised. Some examples of supervised machine learning based approaches are: (De Choudhury, Gamon, Counts & Horvitz, 2013) used supervised learning algorithms to detect depressive disorders; they measured behavioral attributes like language, emotion, social engagement, linguistic style and mentions of anti-depressant medication to build their classifier. Moreover, they used crowd-sourced twitter dataset for classification and achieved 70% accuracy in the classification task. Purver & Battersby (2012) used supervised learning algorithms for their emotion detection system. They trained their supervised classifier using twitter dataset labeled with automatic annotations of Ekman's (Ekman, 1992) classes which include anger, fear, happiness, sadness, surprise, and disgust and achieved 60% accuracy. Qadir & Riloff (2013) employed a bootstrapping approach using emotion hashtags to train their supervised classifier.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Related work

Authors	Dataset	Emotions	Annotations
Alm, Roth, & Sproat, 2005	Children stories (185)	Ekman's Classes	Manual
Aman, & Szpakowicz, 2007	Blogs	Ekman's Classes	Manual
Strapparava, & Mihalcea, 2008	News headlines (WordNet Affect 4,787 terms)	Emotion valence and categories and Affective labels.	Manual
Wiebe, Wilson, & Cardie, 2005	10,000 sentences of news articles	Ekman's Classes	Manual
Liu, Lieberman, & Selker, 2003	OMCS (Open Mind Common Sense)	Ekman's Classes	Manual
Balabantaray, Mohammad, & Sharma, 2012	8,150 tweets	Anger, Disgust, Fear, Joy, Sad, Surprise, Neutral	
Bradley, & Lang, 1999	ANEW (1034 words)	Valence, Arousal and Dominance	Manual
Mohammad, & Turney, 2013	14,182 words	Plutchiks Emotions	Manual
Qadir, & Riloff, 2013	Hashtags	Joy, anger, affection, anxiety and sadness	Automatic
Mehrabian, 1996	DAL (8,742 words)	Activation, Evaluation and Imaginary	Manual
Staiano, & Guerini, 2014	DepecheMood (37,216 words)	8 Emotion categories	Automatic
Suttles, & Ide, 2013	Twitter dataset	Plutchiks eight emotion labels	Automatic
Hasan, Rundensteiner, & Agu, 2014	Twitter dataset	Valence and Arousal	Automatic
Neviarouskaya, Prendinger, & Ishizuka, 2007	Blog posts	Ekman's Classes	Manual
Purver, & Battersby, 2012	Twitter dataset	Ekman's Classes	Manual
Ortigosa, Martín, & Carro, 2014	Facebook Posts	Positive, Negative and Neutral	No Annotation

They used candidate emotion hashtags for five emotion classes of anger, affection, anxiety, joy, and sadness and achieved 68% accuracy in their classification task. Wilson, Wiebe & Hwa (2004) classified the text into Subjective and Objective texts. This approach used the Supervised Learning approach for classification of text. Subjectivity determines whether a sentence is a positive or negative sentence and Objectivity determines the intensity (relationship strength) of the emotions related to that sentence. This system requires proper sentence structure for determining syntactic relations. Thelwall, Wilkinson & Uppal (2010) analyzed the different texts from online social media MySpace and implemented emotion mining on these texts. They concluded 2-Dimensional models provide better accuracy for emotion analysis. Balabantaray, Mohammad, & Sharma (2012) used sentence level annotation of 8,150 tweets collected from the web to train their supervised classifier. They have used contextual, syntactic and linguistic features in their classification task and achieved 73.05% accuracy.

The supervised automatic emotion detection systems require an affective corpus. Affective corpus is a considerable set of structured sentences where each sentence is annotated with emotion tags. This corpus is built by employing human annotators who analyze each sentence and assign a corresponding label to the sentence. The corpus acts as a heart of the supervised learning algorithms as they are trained on this corpus to infer class to the unseen data. However, the corpus is difficult and arduous to create as human judgment is expensive and error-prone.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Emotion Corpus is generally annotated with Ekman's basic emotion categories, there are several such works, like Alm, Roth, & Sproat (2005) have created sentence level annotation of 185 children stories with six emotional classes; Aman & Szpakowicz (2007) created a corpus with emotional intensity and categories of blog posts acquired directly from the web; Strapparava & Mihalcea (2008) labelled news headlines with emotion intensities and categories.

Some recent advances in the area of emotion annotation task have been recorded in the literature where the focus on the task of annotating the sentences is automated; for example, Bifet & Frank (2010) used bootstrapping to automatically classify tweets by expanding seed lists using Word2Vec models and LSA algorithms. Their evaluation results were also promising. Go, Bhayani & Huang (2009) use distant-supervision using emoticons to create positive and negative labeled dataset for the supervised learning algorithm s. Suttles, J., & Ide, N. (2013, March) uses distant supervision using hash-tags, emoticons, and emojis to automatically annotate the twitter dataset using Plutchiks eight emotion labels; Hasan, Rundensteiner & Agu (2014) used circumplex model (Russell, 1980) where emotional states characterizing affect experience composed of valence and arousal is captured using twitter hash-tags to annotate twitter dataset automatically.

Unsupervised learning algorithms do not stand in need of any training set for learning and are thus less expensive than supervised algorithms for emotion detection process. Some examples of unsupervised learning algorithms for emotion detection include Agrawal & An (2012) used the context-based approach based on a procedure of exclusive emotion lexicon, and thus their model can classify prior to Ekman 's six classes. Calvo & Mac Kim (2013) proposed an unsupervised learning algorithm for emotion detection using dimensional models and used ANEW (Bradley & Lang, 1999) database for emotion detection. Neviarouskaya, Prendinger & Ishizuka (2007) proposed a rule-based technique to determine emotions: Anger, Disgust, Fear, Happiness, Sadness, and Surprise (Ekman, 1992) from the sentences of different blog posts.

Lexicon Based Approach

In the lexicon-based approach lexical resources like a set of words, ontologies and lexicons are employed for emotion detection. These methods usually start with the small set of words called seed set and then usually built around it and expand the seeds using several resources like WordNet (Staiano & Guerini, 2014) and build a more extensive set of words for classification purpose. Liu, Lieberman & Selker (2003) proposed a technique to extricate emotions from the sentence. This technique is premised on the affective common sense in which real-world knowledge is used to draw sentiments or emotions from sentences. The method is based on the assumption that all the individuals have the same feeling towards a particular event. This technique categorizes the emotions into the Ekman's six emotions (Ekman, 1992). Yassine & Hajj (2010) presents a framework to determine the affective interactions. They used Lebanese Facebook dataset as a case study and used the unsupervised technique to classify texts based on Friendship and subjectivity. They used k-means for classification of text by subjectivity: neutral, moderately subjective and subjective. The classification results reported an accuracy of 87% based on the subjective content of comments shared by online Users.

Ortigosa, Martín & Carro (2014) presented a novel approach for sentiment analysis of Facebook posts. They employed a hybrid approach of lexical-based and machine-learning techniques using SentBuk for Classification of Facebook Status posts.SentBuk obtained Facebook posts from Facebook users performs classification of the posts and presents the results of classification by the interface. They used

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

J48, Naïve-Bayes and Support Vector Machines to build the classifier. The performance of this classifier was excellent with an accuracy of 83.27%. This approach can be used in e-learning systems and recommended systems where a user can freely express his thoughts and emotions.

Aman & Szpakowicz (2007) Implemented the technique of annotating sentiments to determine emotion bearing sentences or words, emotion categories, and their intensities. They proposed an annotation scheme for recognition of different types of emotions, the strength of emotions, and phrases bearing emotion in the text. They worked on self-annotated blog posts having happy and sad mood labels. They conducted their work on two specific emotions happiness and sadness. Their complete processed work is used for automatic sentiment and affective experiments. Their classification techniques produced an accuracy of 73.89%.

Alm, Roth & Sproat (2005) described a technique of supervised learning for automatic classification of sentences into the basic six emotions (Ekman emotions) in children's fairy tales. They used Snow learning architecture for classification of sentences.

Wiebe, Wilson & Cardie (2005) used the MPQA corpus composed of 10,000 sentences of news articles for manual labeling of opinions, emotions, and sentiment. As mentioned previously the task of emotion detection is the particularly hot topic in the sentiment detection and analysis field with its applications ranging from human-computer interaction to suicide detection, depression detection, mood swings to welfare tracking and cyberbullying. However, not much work focuses on semantic of text instead the focus is towards building a corpus for supervised earning and lexical features of the text. Supervised learning although producing excellent results is costly to implement as it involves the creation of emotional corpora which is created by human annotators and is thus expensive to do and can lead to bias due to inter-annotator agreements.

In this work, we propose an automatic emotion detection system which captures the semantic features of the text. It uses an amalgam of deep learning and unsupervised learning to capture the semantic similarity between textual elements to classify the text for emotion classification purpose. We use the NRC Emotion Lexicon (EmoLex) (Mohammad & Turney, 2013) as our lexical resource and classify the text into Ekman's six basic emotion classes.

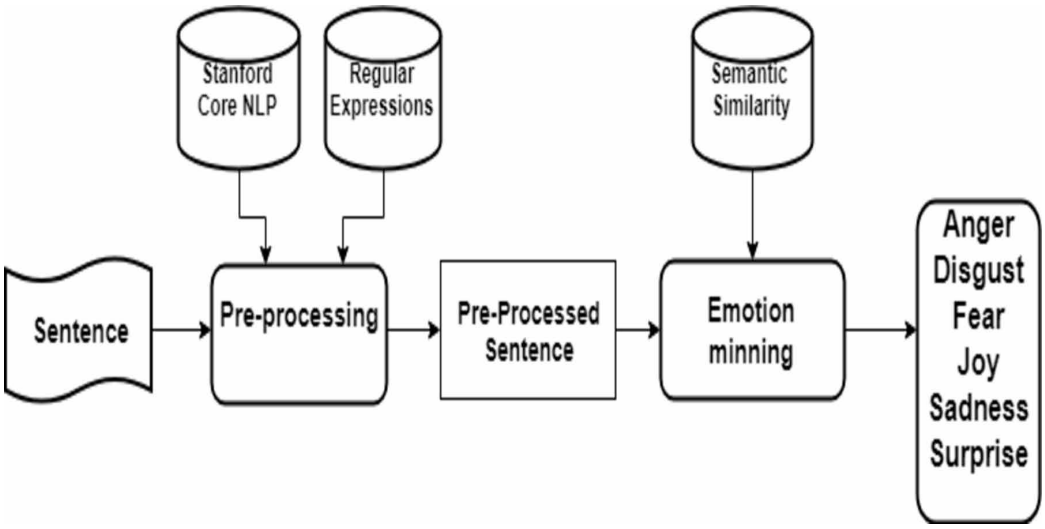
As an outcome of the compilations implicated from the analysis of the previous work and a consideration on the undetermined issues, our proposed semantic automatic emotion classifier is described in detail in the next section.

SEMANTIC AUTOMATIC EMOTION CLASSIFICATION

The objective of the automatic semantic emotion classification is to classify unlabeled sentences conveying some emotions. To address this, we have designed an automatic emotion classifier capable of classifying a textual sentence into predefined classes taking semantic connotation of the sentence to be classified into consideration. The classifier designed is thus capable of classifying an English sentence into one of the predefined classes based on the emotion orientation of the sentence. Structure of our emotional classifier consists of 1) Generate lexicon for classification using NRC emotion Lexicon (EmoLex) (Mohammad & Turney, 2013); 2) Use distributional Semantic (Marelli, Bentivogli, Baroni, Bernardi, Menini & Zamparelli, 2014) algorithms for classification. Precisely, our sematic automatic emotion classification approach consists of unsupervised and deep learning algorithms. Our classification process is thus an amalgam of unsupervised learning and deep learning. For generating of our lexicon,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Overall methodology



we have used EmoLex (Mohammad & Turney, 2013) in its reduced form which is then extended using WordNet (Staiano & Guerini, 2014). For classification, we do the preprocessing of data followed by using distributional semantic algorithms to capture the semantics of text and then vectorizing the data for classification. Lexicon is also vectorized, and then a mapping function is applied to calculate the emotional score of the sentence against each class of the emotional categories. The maximum score for the emotional category will give the label for the sentence. The overview of our Semantic emotion classifier is shown in the diagram Figure 1.

The process acquires a collection of unlabeled phrases or sentences and a set of emotional classes as input; exclusively this paper works with the Ekman's six basic emotions and produces as output an emotional label among the six emotional labels of Ekman's classes for the sentence if it has an emotional affinity towards any of the six categories or neutral otherwise. This section is divided into three subsections where the fundamental task performed by each is described in detail.

LEXICON GENERATION

In lexicon generation, we discuss the technique of building the initial seed by using the NRC Word-Emotion Association Lexicon (Emolex) (Mohammad & Turney, 2013). Emolex consists of general English language unigrams annotated using Amazon Mechanical Turk. It is annotated using Plutchik's emotion classes plus two sentiment labels (positive and negative). Emolex has 14000 English unigrams. Since our emotional labels are Ekman's six emotional labels thus, we need to reduce the Emolex to cover only six Ekman's labels (Figure 2), and hence our reduced Emolex has 3,462 emotional words annotated with six Ekman's classes.

Table 2 shows the distribution of reduced Emolex for Ekman's six classes.

However, due to the reduction, there was a need to extend Emolex. Hence, we extended the reduced Emolex with synonyms. More precisely we extended reduced Emolex to build a better and bigger seed set using WordNet (Staiano, J., & Guerini, M. (2014).) and Oxford (Marelli, Bentivogli, Baroni, Bernardi, Menini & Zamparelli, 2014) synonyms.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Lexicon generation

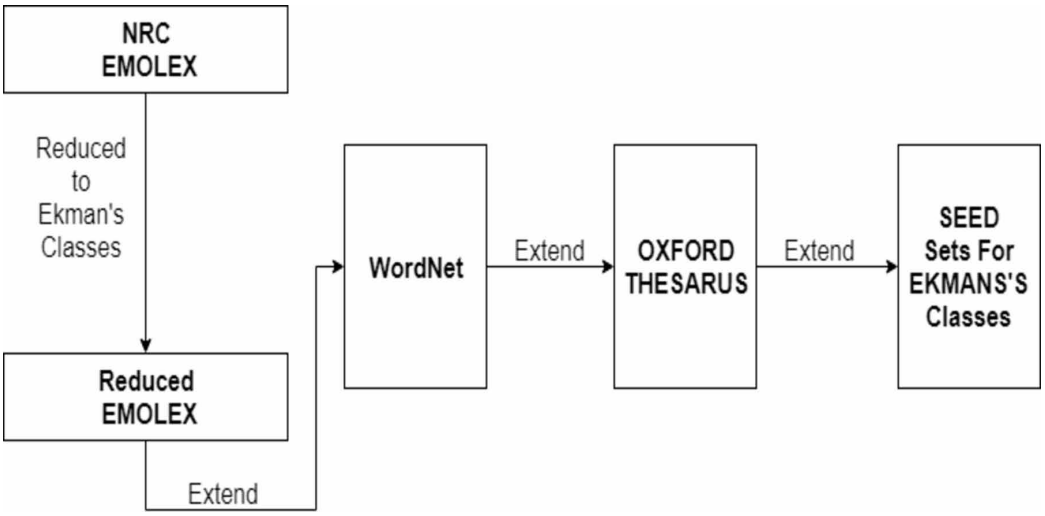


Table 2. Reduced Emolex

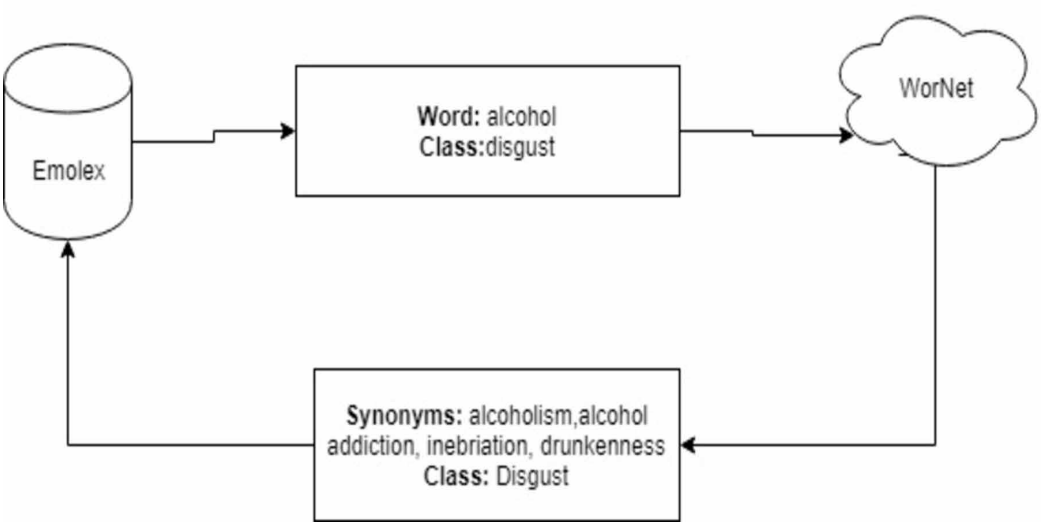
	Emolex (Ekman's emotions) # of Words
Anger_words	1,247
Disgust_words	1,058
Fear_words	1,476
Joy_words	689
Sadness_words	1,191
Surprise_words	1,247

SEED EXTENSION USING WORDNET

In seed expansion process, each word in the reduced Emolex is looked for the synsets using WordNet (Staiano & Guerini, 2014). For each word in the Emolex, we obtain synonyms of its most appropriate sense using WordNet and annotate the retrieved synonym using the same annotation as that of the word in the reduced Emolex. Figure 3 shows the example of extracting synonyms from WordNet. The word alcoholism is looked up for synonyms in WordNet and has class "disgust," from WordNet the process retrieves synonyms "alcoholism, alcohol addiction, inebriation, drunkenness" which are appended to the Emolex and annotated with the same class label, i.e., disgust. After this process, Emolex was extended with 5219 words resulting in a new lexicon of 8681 words.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Extracting synonyms from WordNet



EXTENSION USING OXFORD THESAURUS

This process is carried out to extend the Emolex further. The extended Emolex obtained in the previous step is feed to the Oxford Thesaurus (Marelli, Bentivogli, Baroni, Bernardi, Menini & Zamparelli, 2014), and synonyms are retrieved for each word. Synonyms are added to the Emolex if they are not already present in it with the emotion tag of the synonym being the same as that of the word that was probed for.

Figure 4 shows the example of lexicon enhancement using Oxford thesaurus.

The word "appreciation" belonging to the class joy is probed for synonyms in the Oxford thesaurus and synonyms "appreciation, grasp, hold, taste, discernment, perceptiveness, admiration" are retrieved which if not present in the lexicon are added with the same annotation, i.e., joy is added to the lexicon.

After this process, the number of words added to the lexicon is 10,811 making the total number of words in the lexicon equal to 19,452. Lexicon is then divided into six sets. Each set is associate with the six Ekman's classes. Figure 3 shows the process of seed set creation.

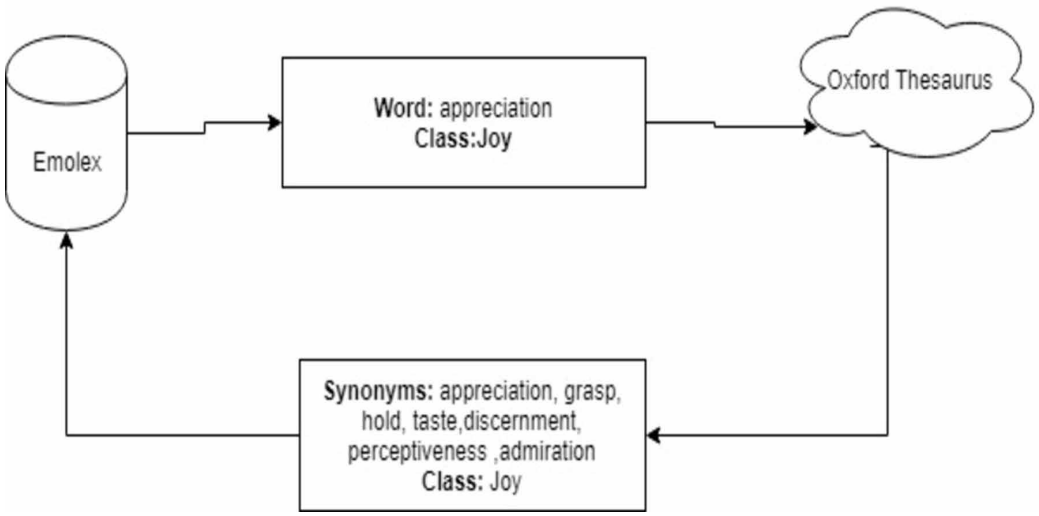
AUTOMATIC CLASSIFICATION

After the creation of Lexicon, the next step is the classification process. The input to the automatic classifier is the set of sentences and a set of labels; the output is the sentence along with the classified label. Our classifier consists of following steps:

1. Data pre-processing
2. Capturing semantics using semantic model
3. Vectorization
4. Semantic similarity

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Seed set creation



Data Pre-Processing

The data is pre-processed to remove the irrelevant fields and inconsistencies. Data is refined at each step of pre-processing and then passed to the automatic classifier for classification. The pre-processing of data is done by the following steps as shown in Figure 5.

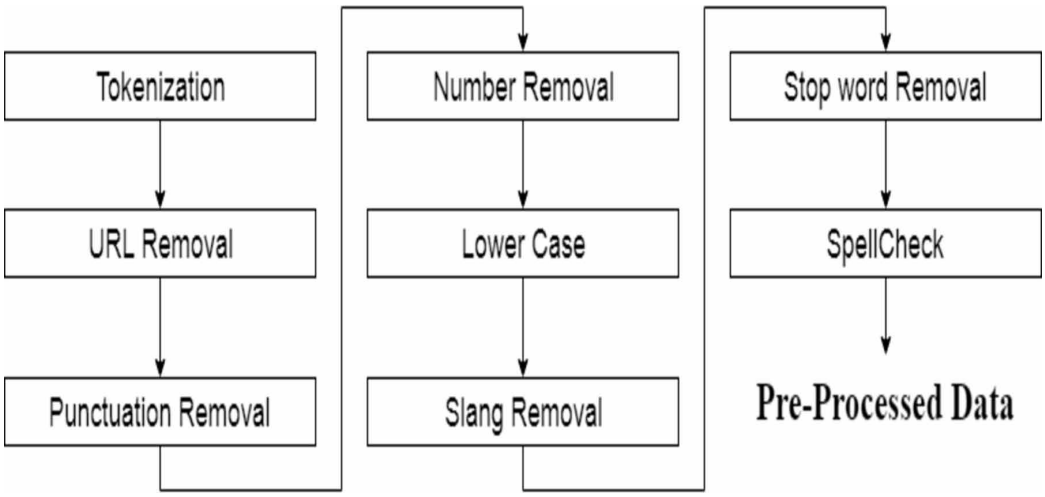
1. **Tokenization**: The data is tokenized to extract the tokens using Stanford Core NLP, and in this step, only the white spaces are used to separate the tokens.
2. **Removing URLs**: Data may contain different URLs and links which are removed.
3. **Punctuation Removal**: Punctuation marks do not have any impact on affect, and thus they are removed.
4. **Number Removal**: Numbers are removed from the data
5. **Lower Case**: This step converts the entire token into lower case.
6. **Slang Removal**: The slangs are replaced by their respective expansions like “lol” is replaced by “laugh out loud.”
7. **Stopword Removals**: Stanford can be used for identification of stop words which then can be removed easily.
8. **Spellcheck**: For spell checking, J Spell is used.

Capturing Semantics Using Semantic Model

Semantic models characterize the semantics of entire sentences. Semantic models work on the principle of distributional hypothesis which demonstrates that words appearing in the same circumstances tend to have similar associations. This leads to the derivation that semantic similarity models predict semantic cohesiveness between the textual elements. For example, “amazing” and “wonderful” are similar could be extended to predict that if “amazing” has positive sentiment so will wonderful. So is the true

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Steps of pre-processing

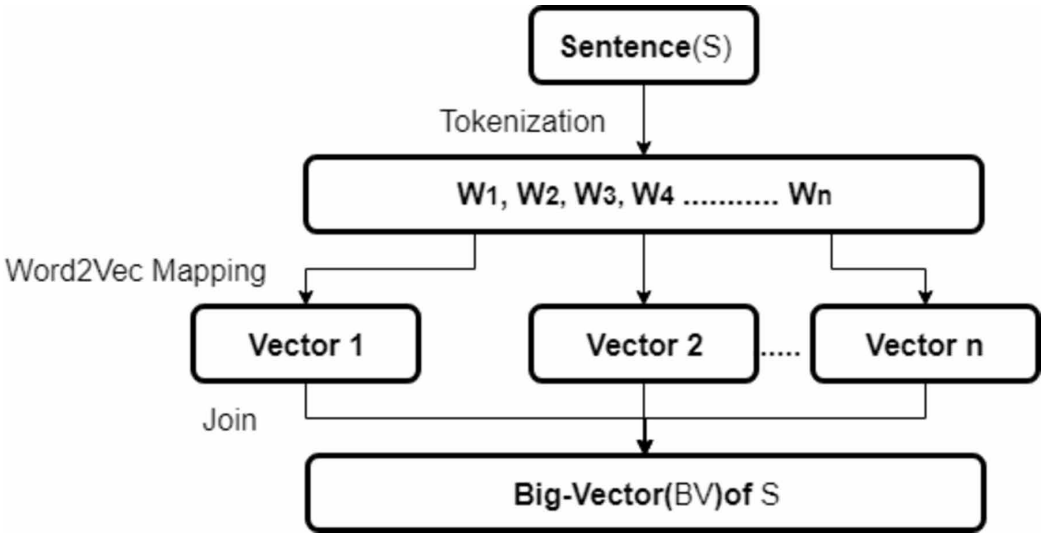


for opposite like words terrible and awful. Semantic models are built using semantic representations (high-dimensional semantic vector spaces) using statistical analysis of the contexts in which words occur. More precisely, each word is represented by a real-valued vector known as word embedding or word vector along with the geometrical properties. These geometrical properties of word vectors or word embeddings prove to be semantically meaningful. Word2Vec algorithm (Mikolov, Chen, Corrado & Dean, 2013) incorporates this intuition. Word2Vec (Mikolov, Chen, Corrado & Dean, 2013) creates a vector representation of words by using a two-layer neural network. The two models of Word2Vec are continuous bag of words (CBOW) and Skip-gram model (SKIP). Both models describe how neural networks learn the underlying principle of word representations for each word. CBOW portends the current word given its context and SKIP portends context given the word. For our classifier, we have used CBOW model train on Wikipedia dump. Wikipedia provides free copies of all available content to interested users. We trained our Word2Vec model using the same data. Each sentence is then fed to the Word2vec to retrieve its Word vector. Thus, for each sentence, we provide all words in it to the word2Vec and retrieve its vectors, and all vectors of the sentence are then combined to form the Big-vector of the corresponding sentence. Hence each Big-vector represents the semantically coherent information of the sentence. Figure 6 shows the Big-vector formation of the sentence.

To further augment big-vectors we used WordNet. WordNet is a lexical reference library for the English language created at Princeton University. It is a lexical database comprised of adjectives, adverbs, nouns, and verbs that are aggregated into related sets called synsets. Synsets have a relation of synonymy. Each synset is related to specific context. WordNet also provides the count of synsets of a particular word by frequency score. Each word in the Big-vector is feed to the WordNet to retrieve its synonyms which are then added to the big-vector to expand it further. After retrieving all the synonyms from the WordNet, we have the final Big-vector which is used for the classification process. To further augment big-vectors we used WordNet. WordNet is a lexical reference library for the English language created at Princeton University. It is a lexical database comprised of adjectives, adverbs, nouns, and verbs that are aggregated into related sets called synsets. Synsets have a relation of synonymy. Each synset is related to specific context. WordNet also provides the count of synsets of a particular word by frequency score.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Big-Vector formation



VECTORIZATION

For Vectorization, TF_IDF is employed. TF_IDF stands for Term Frequency-Inverse Document Frequency. It is used for unsupervised clustering. TF_IDF provides us the importance of terms in the corpus. Term Frequency gives us the frequency of terms in a document. Mathematically, TF is given by c (w) where

$$c(w) = \frac{\text{Frequency of word } \mathbf{w} \text{ in a document}}{\text{Frequency of words in the document}}.$$

IDF provides us importance of the terms in a document. Mathematically, IDF is given by d (w) where

$$d(w) = \log e \frac{\text{Total Frequency of documents}}{\text{Frequency of documents with word w}}$$

Both the seed sets and the sentence are vectorized for the classification purpose. More precisely, both seed sets and documents are represented as vectors in the vector space model (VSM), and the similarity is measured using semantic relatedness between them.

Semantic Similarity

For obtaining an emotional class of the sentence semantic relatedness is calculated between the sentence and the seed sets and then scoring function is used to determine the class of the sentence to be classified. Seed sets created in the lexicon generation and the sentence after preprocessing and application of semantic models is converted into Big-vector are both converted into vectors using Vectorization. Next semantic similarity is found between them to determine the emotional class of the sentence.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Seed words in the lexicon are transformed into categories and sentence to be classified are represented as documents. Then categories are matched against documents using our scoring function. Both categories and documents are represented as vectors in the vector space model (VSM), and the similarity is measured by the cosine similarity. Then our scoring function is used to calculate the class of the sentence.

The scoring function then finds the maximum of the cosine similarity score with the seed sets and assigns the class of the sentence to that of the seed set with which it has the maximum score.

Mathematically,

$$Score(s) = sim_{VCM}(id_c, \mathrm{s}) = \cos\left(\overrightarrow{id_c}, \vec{s}\right)$$

Where id_c is the seed set, s is the sentence, $\overrightarrow{id_c}$, is vector representation of the seed set and $\vec{s}$, is vector representation of Sentence.

ALGORITHM

The algorithm of our classifier is given as under:

Step 1: Seed lexicon is generated from NRC Emolex and extended using WordNet and Oxford thesaurus
Step 2: Data is pre-processed to remove inconsistencies and noise from data.
Step 3: Distributional semantic model (Word2vec) is used to capture the semantics of the pre-processed sentence.
Step 4: Big-Vector is generated using Word2Vec of the sentence.
Step 5: WordNet is used to extract synonyms of the Big-vector to enhance its semantic content.
Step 6: Background Corpus is used to get TF_IDF of Big-vector, and seed lexicon is obtained to transform data and lexicon into vectors.
Step 7: Semantic relatedness of sentence and seed lexicon is measured.
Step 8: The class for which the semantic similarity score is maximum is chosen to be the emotion class of the sentence. Mathematically, emotion class of the sentence is given by

$$emotion = \max(\cos(Big - vector, seed\ sets))$$

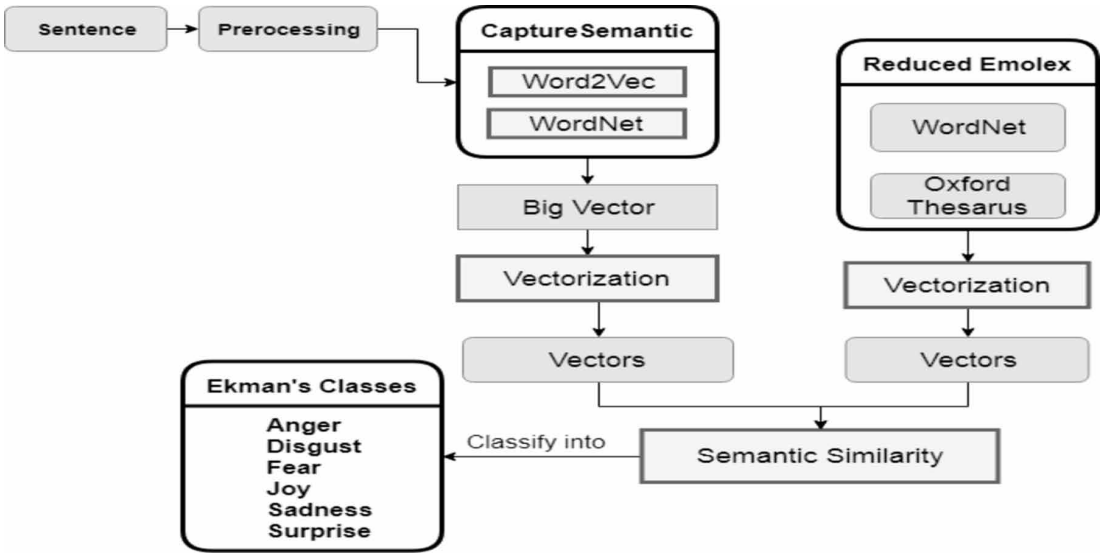
Overall working of semantic emotion classifier is explained in Figure 7.

4. DATASETS

The objective of this research is to develop an automatic emotion classifier which captures the semantics of the text and access its viability to classify the text into different emotional classes. To achieve that in this paper two evaluations on two different data sets have been done, explained in Section a and Section b gives the overview of the datasets used for evaluation.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Overall working of semantic emotion classifier



Our proposed emotion classifier is applied on two emotion corpora annotated at the sentence level: (1) Aman Corpus (Aman & Szpakowicz, 2007); and (2) Affective Text Corpus (Strapparava & Mihalcea, 2007). We selected these corpora because these corpora are annotated by humans and will allow us to compare the results of our classifier with the human annotations; also they are most used corpora for the emotion detection tasks like (Keshtkar & Inkpen, 2010; Chaffar & Inkpen, 2011; Mohammad, 2012). Also, the source of these datasets range from web blogs to news headline, and thus the applicability of our Automatic emotional classifier can be tested for the different genre.

Aman Corpus

This dataset is collected from web blogs and has around 4000 sentences in it. Sentences are manually annotated with Ekman's six basic emotional categories.

Affective Text Corpus

Affective Text Corpus contains 1,250 short text stories collected from the news headlines of the newspapers like the New York Times, CNN (Cable News Network), BBC as well as the news from the Google News. Ekman's six basic emotion classes are used for the annotation which is done manually by the human annotators.

Aman Corpus

Aman corpus is manually annotated dataset that does not require any training data, although annotated sentences are used to determine the different annotations that are needed. In this corpus, Ekman's emotion classes along with (i) mixed emotions and (ii) no emotion are employed resulting in 8 emotion classes. Cohen's kappa (Cohen, 1960) was employed for annotation metrics.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Affective Text Corpus

The second dataset used in this experiment is taken from semeval-2007: affective text (Strapparava & Mihalcea, 2007). Affective text deals with the categorization of emotions and emotion intensities to determine the association between the emotions and lexical semantics. Emotional content is expressed by the words; even the neutral words can convey different acquaintances due to semantic differences. The automatic exploration of emotions is becoming popular due to its practical applications like e-learning, personalized marketing, opinion mining and affective computing. This dataset is composed of the news headlines drawn from different newspapers and news websites like Google news search engine. New York Times, CNN (Cable News Network), and British Broadcasting Corporation news. The inter-annotator agreement was implemented using Pearson correlation measure for six emotions: Anger, Disgust, Fear, Happiness, Sadness, and Surprise (Ekman, 1992) and intensity annotations. The results of Pearson correlation are shown in Table 3.

Table 4 gives the distribution of different emotions in the two datasets.

5. EVALUATION AND RESULTS

For classification, we used Semantic similarity which is popularly employed in Sentiment and Emotion analysis. Semantic similarity provides us the semantic relationship between the vectors concerning the emotions. Accuracy, precision, recall, and F1- score is used to evaluate the performance of Semantic Similarity. These parameters are estimated using the confusion matrix. A confusion matrix (Table 5) is a table that is used to measure the efficiency of a Classifier.

Table 3. Inter-Annotator Agreement

Emotions	Evaluation Measures (Pearson correlation)
Anger	49.55
Disgust	44.51
Fear	63.81
Joy	59.91
Sadness	68.19
Surprise	36.07
Valence	78.01

Table 4. Emotion distribution of datasets

Dataset	Happiness	Sadness	Anger	Disgust	Surprise	Fear
Affective Text corpus	334	214	175	28	131	166
Aman Corpus	536	173	115	115	172	179

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Confusion matrix

		Predicted class	
		P	**N**
Known class	**P**	Tr_P	Fa_N
	N	Fa_P	Tr_N

Accuracy

Accuracy is a perceptive evaluation method for measuring performance. It is given by the number of true classifications to the total classifications.

$$\text{Accuracy} = \frac{\text{Tr_P}+\text{Tr_N}}{\text{Tr_P}+\text{Fa_P}+\text{Fa_N}+\text{Tr_N}}$$

Where Tr_P signifies True positive, Fa_P signifies false Positive, Tr_N signifies True Negative, and Fa_N signifies False Negative.

Precision

Precision also called Positive Predictive Value is equal to the correctly measured positive classifications divided by the total predicted positive classifications. Higher the precision, the lower the rate of false positives.

$$\text{Precision} = \frac{\text{Tr_P}}{\text{Tr_P}+\text{Fa_P}}$$

Where Tr_P signifies True positive, Fa_P signifies false Positive.

Recall

Recall or Sensitivity is equal to the correctly predicted positive classifications divided by all classifications in the actual class.

$$\text{Recall} = \frac{\text{Tr_P}}{\text{Tr_P}+\text{Fa_N}}$$

Where Tr_P signifies True positive, Tr_N signifies True Negative, and Fa_N signifies False Negative

F1_Score

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

F1_score is the harmonic mean of precision and recall.

$$\text{F1 Score}=2*\frac{\text{Recall*Precision}}{\text{Recall+Precision}}$$

Our semantic automatic emotion classifier achieved the results on both datasets as described in the tables: Table 6 and Table 7.

Figure 8 and Figure 9 shows the performance of our classifier graphically on two datasets Aman Corpus and Affective Text Corpus respectively.

Table 6. 1st dataset

Aman corpus			
	Precision	**Recall**	**F1_Score**
Anger	0.536	0.456	0.49277419
Disgust	0.86	0.63	0.72724832
Fear	0.845	0.56	0.67359431
Joy	0.745	0.458	0.56726517
Sadness	0.568	0.541	0.55417133
Surprise	0.871	0.456	0.59860739
Neutral	0.678	0.714	0.69553448
Macro average	0.729	0.545	0.61559931
Accuracy		**74.56%**	

Table 7. 2nd Dataset

Affective Text Corpus			
	Precision	**Recall**	**F1_Score**
Anger	0.489	0.785	0.60261381
Disgust	0.823	0.74	0.77929623
Fear	0.744	0.654	0.696103
Joy	0.854	0.581	0.69153171
Sadness	0.652	0.885	0.7508393
Surprise	0.851	0.562	0.67694551
Neutral	0.745	0.56	0.63938697
Macro average	0.73685714	0.681	0.6909595
Accuracy		**69.03%**	

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Performance of classifier on Aman Corpus

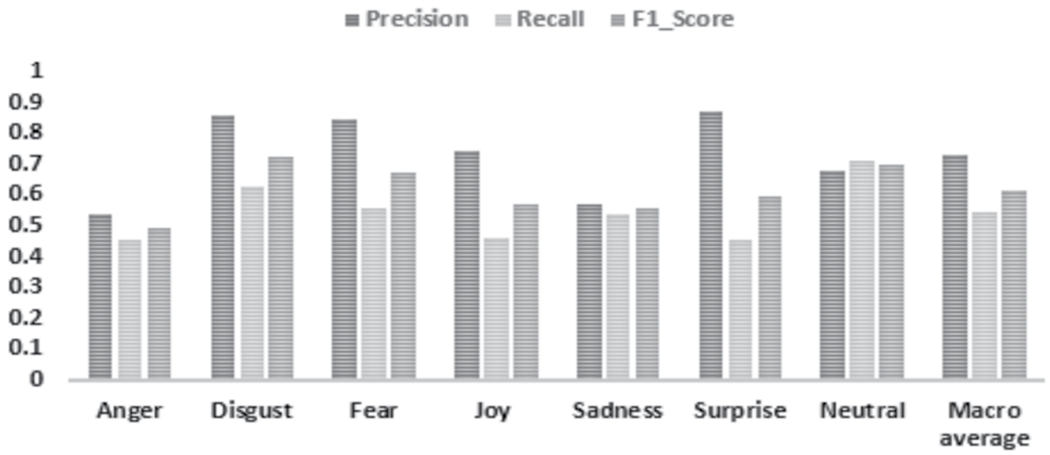
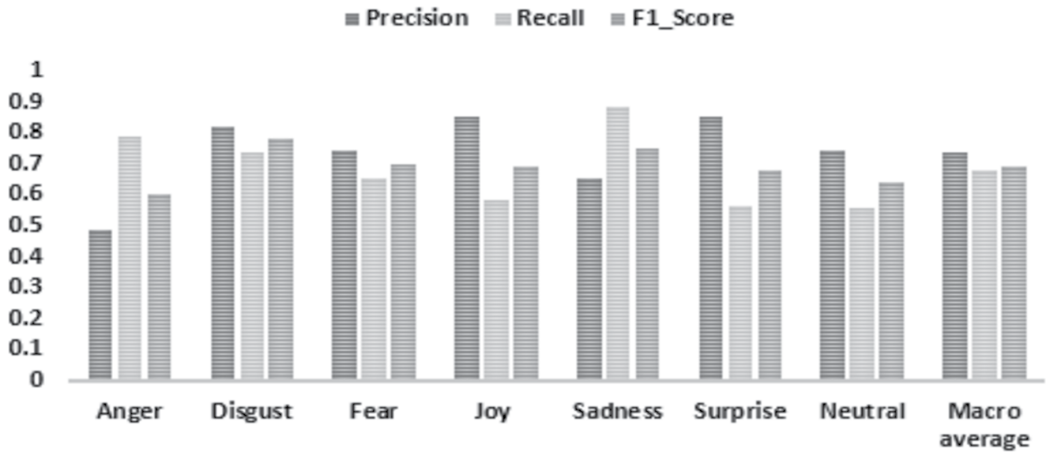


Figure 9. The overall process of semantic similarity



6. COMPARATIVE STUDY

According to our awareness, no allied research has been performed on utilizing semantic relatedness to create an emotion classifier for textual data streams. However, some work has been carried out utilizing Aman Corpus, NRC Emolex lexicon, and Affective Text Corpus. Table 8 compares our work with the existing work.

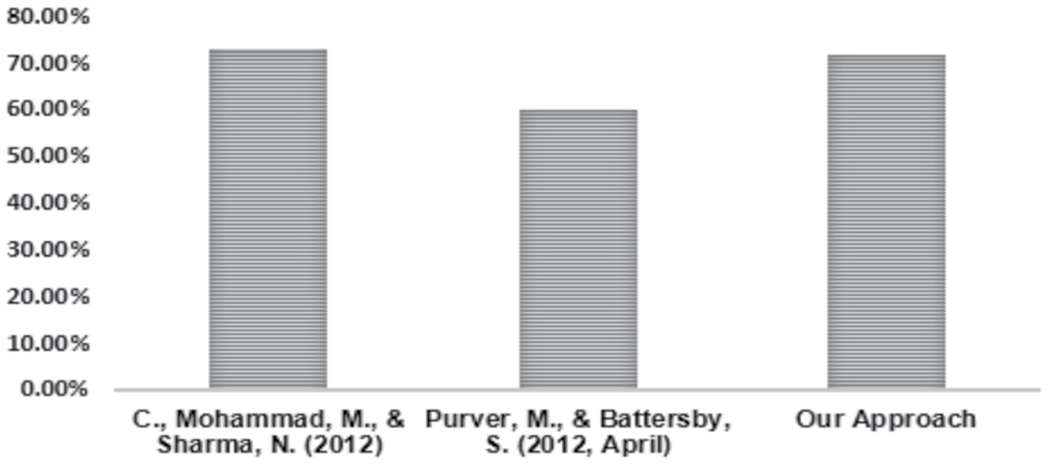
Figure 10 show the comparison of our work with the existing work graphically.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 8. Comparative study

Authors	Accuracy
Aman & Szpakowicz, 2007	73.89%.
Balabantaray, Mohammad & Sharma, 2012	73.05%.
Qadir & Riloff, 2013	68%.
De Choudhury, Gamon, Counts & Horvitz, 2013	70%.
Balabantaray, Mohammad & Sharma, 2012	73.05%
Purver & Battersby, 2012	60%
Our Approach	**71.795%**

Figure 10. Comparison of our Approach with existing approaches



7. CONCLUSION AND FUTURE SCOPE

As presented in the introduction of the paper, the motive behind this research is to capture semantic relatedness while classifying during the automatic emotion classification task. This research aims to design an automatic emotion classifier which computes the semantic relationship between two textual elements into consideration while classifying a textual element into different emotional classes. This paper presents a semantic emotion classifier premised on the principle of distributional semantic models and uses three main steps 1) creation of lexicon using reduced NRC Emolex 2) using Distributional semantic models to capture semantic relatedness between textual elements and 3) using scoring function to classify textual elements into Ekman's six basic emotion classes.

Based on the evaluations performed, the appropriateness and accuracy of our approach are corroborated. Thus, the main conclusions of our research are:

1. The viability of Semantic automatic emotion classification as a process to automatically classify the emotional sentences is demonstrated
2. Distributional semantic models act as a good source to capture the semantics of the text

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

3. NRC EmoLex as a lexicon to automatically classify text into several classes performs satisfactorily taking into consideration that it is a general-purpose lexicon and has been employed for different domains of text.
4. The classification results allow us to conclude that the semantic automatic emotion classifier is an efficient classification scheme for emotion classification which considers semantic orientations of the textual streams.

Our future research will deal with

1. Exploitation of large affective datasets like DepecheMood, EmoSenticNet.
2. Contemplation of different multi-dimensional models and testing of these approaches with in their specified domain.
3. Analysing applicability of the other lexicons to extend the seed.
4. Experimenting the technique in more corpora with emotion lexicon annotated by more emotion classes.

REFERENCES

Ortigosa, A., Martín, J. M., & Carro, R. M. (2014). Sentiment analysis on Facebook and its application to e-learning. *Computers in Human Behavior*, *31*, 527–541. doi:10.1016/j.chb.2013.05.024

Aman, S., & Szpakowicz, S. (2007, September). Identifying expressions of emotion in text. In *International Conference on Text, Speech, and Dialogue* (pp. 196-205). Springer, Berlin, Heidelberg.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, *6*(3-4), 169–200. doi:10.1080/02699939208411068

Plutchik, R., & Kellerman, H. (1980). Theories of Emotion (Vol. 1).

Liu, H., Lieberman, H., & Selker, T. (2003, January). A model of textual affect sensing using real-world knowledge. In *Proceedings of the 8th international conference on Intelligent user interfaces* (pp. 125-132). ACM. 10.1145/604045.604067

Neviarouskaya, A., Prendinger, H., & Ishizuka, M. (2007, January). Analysis of affect expressed through the evolving language of online communication. In *Proceedings of the 12th international conference on Intelligent user interfaces* (pp. 278-281). ACM. 10.1145/1216295.1216346

Wilson, T., Wiebe, J., & Hwa, R. (2004, July). Just how mad are you? Finding strong and weak opinion clauses. In AAAI (Vol. 4, pp. 761-769).

Thelwall, M., Wilkinson, D., & Uppal, S. (2010). Data mining emotion in social network communication: Gender differences in MySpace. *Journal of the Association for Information Science and Technology*, *61*(1), 190–199.

Balabantaray, R. C., Mohammad, M., & Sharma, N. (2012). Multi-class twitter emotion classification: A new approach. *International Journal of Applied Information Systems*, *4*(1), 48–53. doi:10.5120/ijais12-450651

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Yassine, M., & Hajj, H. (2010, December). A framework for emotion mining from text in online social networks. In *2010 IEEE International Conference on Data Mining Workshops (ICDMW)* (pp. 1136-1142). IEEE.

Alm, C. O., Roth, D., & Sproat, R. (2005, October). Emotions from text: machine learning for text-based emotion prediction. In *Proceedings of the conference on human language technology and empirical methods in natural language processing* (pp. 579-586). Association for Computational Linguistics. 10.1109/ICDMW.2010.75

Wiebe, J., Wilson, T., & Cardie, C. (2005). Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*, *39*(2-3), 165–210. doi:10.100710579-005-7880-9

Alm, E. C. O. (2008). *Affect in* text and speech*. University of Illinois at Urbana-Champaign.

Strapparava, C., & Mihalcea, R. (2007, June). Semeval-2007 task 14: Affective text. In *Proceedings of the 4th International Workshop on Semantic Evaluations* (pp. 70-74). Association for Computational Linguistics. 10.3115/1621474.1621487

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, *29*(3), 436–465. doi:10.1111/j.1467-8640.2012.00460.x

Rodriguez, P., Ortigosa, A., & Carro, R. M. (2012, July). Extracting emotions from texts in e-learning environments. In *2012 Sixth International Conference on Complex, Intelligent and Software Intensive Systems (CISIS)* (pp. 887-892). IEEE. 10.1109/CISIS.2012.192

Suero Montero, C., & Suhonen, J. (2014, November). Emotion analysis meets learning analytics: online learner profiling beyond numerical data. In *Proceedings of the 14th Koli calling international conference on computing education research* (pp. 165-169). ACM. 10.1145/2674683.2674699

Desmet, B., & Hoste, V. (2013). Emotion detection in suicide notes. *Expert Systems with Applications*, *40*(16), 6351–6358. doi:10.1016/j.eswa.2013.05.050

Vaassen, F. (2014). *Measuring emotion* (Doctoral dissertation, Ph. D. dissertation, Universiteit Antwerpen).

Cherry, C., Mohammad, S. M., & De Bruijn, B. (2012). Binary classifiers and latent sequence models for emotion detection in suicide notes. *Biomedical Informatics Insights*, 5.

Dadvar, M., Trieschnigg, D., Ordelman, R., & de Jong, F. (2013, March). Improving cyberbullying detection with user context. In *European Conference on Information Retrieval* (pp. 693-696). Springer.

Mohammad, S. M., & Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, *29*(3), 436–465. doi:10.1111/j.1467-8640.2012.00460.x

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Strapparava, C., & Mihalcea, R. (2008, March). Learning to identify emotions in text. In *Proceedings of the 2008 ACM symposium on Applied computing* (pp. 1556-1560). ACM. 10.1145/1363686.1364052

Bifet, A., & Frank, E. (2010, October). Sentiment knowledge discovery in twitter streaming data. In *International conference on discovery science* (pp. 1-15). Springer. 10.1007/978-3-642-16184-1_1

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. (CS224N Project Report).

Suttles, J., & Ide, N. (2013, March). Distant supervision for emotion classification with discrete binary values. In *International Conference on Intelligent Text Processing and Computational Linguistics* (pp. 121-136). Springer. 10.1007/978-3-642-37256-8_11

Hasan, M., Rundensteiner, E., & Agu, E. (2014). Emotex: Detecting emotions in twitter messages.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. doi:10.1037/h0077714

De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. (2013). Predicting depression via social media. *ICWSM*, *13*, 1–10.

Purver, M., & Battersby, S. (2012, April). Experimenting with distant supervision for emotion classification. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 482-491). Association for Computational Linguistics.

Qadir, A., & Riloff, E. (2013). Bootstrapped learning of emotion hashtags# hashtags4you. In *Proceedings of the 4th workshop on computational approaches to subjectivity, sentiment and social media analysis* (pp. 2-11).

Agrawal, A., & An, A. (2012, December). Unsupervised emotion detection from text using semantic and syntactic relations. In *Proceedings of the 2012 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technology* (Vol. 1, pp. 346-353). IEEE Computer Society. 10.1109/WI-IAT.2012.170

Calvo, R. A., & Mac Kim, S. (2013). Emotions in text: Dimensional and categorical models. *Computational Intelligence*, *29*(3), 527–543. doi:10.1111/j.1467-8640.2012.00456.x

Bradley, M. M., & Lang, P. J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings* (Technical report C-1). The center for research in psychophysiology, University of Florida.

Hancock, J. T., Landrigan, C., & Silver, C. (2007, April). Expressing emotion in text-based communication. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 929-932). ACM. 10.1145/1240624.1240764

Izard, C. E., Libero, D. Z., Putnam, P., & Haynes, O. M. (1993). Stability of emotion experiences and their relations to traits of personality. *Journal of Personality and Social Psychology*, *64*(5), 847–860. doi:10.1037/0022-3514.64.5.847 PMID:8505713

Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, *98*(2), 219–235. doi:10.1037/0033-2909.98.2.219 PMID:3901060

Mehrabian, A. (1996). Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament. *Current Psychology*, *14*(4), 261–292. doi:10.1007/BF02686918

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Mohammad, S. M., & Kiritchenko, S. (2015). Using hashtags to capture fine emotion categories from tweets. *Computational Intelligence*, *31*(2), 301–326. doi:10.1111/coin.12024

Poria, S., Gelbukh, A., Hussain, A., Howard, N., Das, D., & Bandyopadhyay, S. (2013). Enhanced SenticNet with affective labels for concept-based opinion mining. *IEEE Intelligent Systems*, *28*(2), 31–38. doi:10.1109/MIS.2013.4

Cambria, E., Poria, S., Bajpai, R., & Schuller, B. (2016). SenticNet 4: A semantic resource for sentiment analysis based on conceptual primitives. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers* (pp. 2666-2677).

Staiano, J., & Guerini, M. (2014). Depechemood: a lexicon for emotion analysis from crowd-annotated news. arXiv:1405.1605

Marelli, M., Bentivogli, L., Baroni, M., Bernardi, R., Menini, S., & Zamparelli, R. (2014). Semeval-2014 task 1: Evaluation of compositional distributional semantic models on full sentences through semantic relatedness and textual entailment. In *Proceedings of the 8th international workshop on semantic evaluation (SemEval 2014)* (pp. 1-8). 10.3115/v1/S14-2001

D'Alessandro, P. (2004). The Oxford American writer's thesaurus.

Keshtkar, F., & Inkpen, D. (2010, June). A corpus-based method for extracting paraphrases of emotion terms. In *Proceedings of the NAACL HLT 2010 Workshop on Computational approaches to Analysis and Generation of emotion in Text* (pp. 35-44). Association for Computational Linguistics.

Keshtkar, F., & Inkpen, D. (2010, June). A corpus-based method for extracting paraphrases of emotion terms. In *Proceedings of the NAACL HLT 2010 Workshop on Computational approaches to Analysis and Generation of emotion in Text* (pp. 35-44). Association for Computational Linguistics.

Chaffar, S., & Inkpen, D. (2011, May). Using a heterogeneous dataset for emotion analysis in text. In *Canadian Conference on Artificial Intelligence* (pp. 62-67). Springer. 10.1007/978-3-642-21043-3_8

Mohammad, S. (2012, June). Portable features for classifying emotional text. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 587-591). Association for Computational Linguistics.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. doi:10.1177/001316446002000104

This research was previously published in the International Journal of Synthetic Emotions (IJSE), 9(2); edited by Amira S. Ashour and Nilanjan Dey; pages 1-22; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 54
Development of Part of Speech Tagger for Assamese Using HMM

Surjya Kanta Daimary
Punjabi University, India

Madhumita Barbora
Tezpur University, India

Vishal Goyal
Punjabi University, India

Umrinderpal Singh
Punjabi University, India

ABSTRACT

This article presents the work on the Part-of-Speech Tagger for Assamese based on Hidden Markov Model (HMM). Over the years, a lot of language processing tasks have been done for Western and South-Asian languages. However, very little work is done for Assamese language. So, with this point of view, the POS Tagger for Assamese using Stochastic Approach is being developed. Assamese is a free word-order, highly agglutinate and morphological rich language, thus developing POS Tagger with good accuracy will help in development of other NLP task for Assamese. For this work, an annotated corpus of 271,890 words with a BIS tagset consisting of 38 tag labels is used. The model is trained on 256,690 words and the remaining words are used in testing. The system obtained an accuracy of 89.21% and it is being compared with other existing stochastic models.

1. INTRODUCTION

Part-of-Speech (POS) tagging is the process where every word in a natural language sentence is marked with its corresponding part of speech category like noun, verb, adjective, adverb, etc. based on both its definition and context. Besides words, punctuation characters and symbols are also labeled accordingly. It is a very important process because it resolves the ambiguity of words in a sentence by assigning accurate POS label to a word depending on the context. As Assamese is morphologically rich and agglutinative language, several words have more than one POS category that makes the word ambiguous. There is an inflection of noun and verb in a sentence in accordance with the grammatical characteristics

DOI: 10.4018/978-1-7998-0951-7.ch054

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

as well. Therefore, POS tagging becomes a challenging task for Assamese. POS Tagger tries to assign the accurate POS labels to ambiguous words in a sentence according to the context and it has a vital role in various NLP applications as because the POS tagged data is used in many other NLP tasks (Jurafsky & Martin, 2000), e.g., in Parsing, the tagged data helps in finding out noun and verb groups, in Named Entity Recognition, it helps in determining the proper names like the name of a person, place or a thing, in Information Retrieval, it helps in selecting the proper nouns or other important word classes from a given text, in Speech Recognition, it helps in modeling a language, in Machine Translation, it helps in generating the probability for word translation of the source language into the target language, as well as it is useful for many other NLP applications. Thus, it is considered as an initial step of the language processing task. As POS Tagger has a great impact on other NLP systems, a tagging result with high accuracy is always encouraging.

There are several methods of POS tagging and basically there are three main approaches which are Rule Based Approach, Stochastic Approach and Hybrid Approach. Rule Based POS tagging is the most primitive approach where hand-written linguistic rules are used for tagging. These rules identify the appropriate tag for an ambiguous word. This method is dependent on dictionary or lexicon to generate the possible POS tags for every word in input text. The Stochastic Approach is based on the probabilities of words that occur for a particular tag. The tag which occurs most repeatedly in the training data is assigned to unknown or ambiguous word. The probability of a given sequence of tags is calculated from the frequency of words from the annotated training corpus. Hybrid Approach is the combination of more than one method which usually contains rule- based and statistical methods. This model uses the essential feature of statistical approaches and uses the rules for better efficiency. The developed POS Tagger for Assamese follows the Stochastic Approach. A bigram Hidden Markov Model (HMM) is used which is one of the processes in this technique. It is a probabilistic model that uses an annotated training corpus. The tagging process is done by computing the tag sequence probability and the word likelihood probability of the corpus. This method is called supervised learning method. Therefore, HMM requires a large amount of annotated corpus to obtain high accuracy. On the other hand, unsupervised learning method does not use the annotated corpus and it calculates the probabilities by using automatic word groupings.

This paper is further divided into five more sections in which second section provides the related work and next section shows the morphological characteristics of Assamese. Fourth section describes the approach then fifth section gives the evaluation of the system. Finally, the paper is concluded in sixth section.

2. RELATED WORK

Till recent years, many research works have been accomplished on Part of Speech tagging for different languages using different approaches. In case of the Indian languages like Hindi, Malayalam, Bengali, etc. which are morphologically rich in nature, a good number of POS taggers have been developed using Stochastic Approach with varied accuracies. This approach takes less effort in implementation and involves very little knowledge about the language but requires large training data to obtain high accuracy. (Dalal et al., 2006), developed POS Tagger using Maximum Entropy Markov Model for Hindi with a corpus of 15562 words and using 27 POS tags and obtained an accuracy of 94.81%. (Joshi et al., 2013), developed HMM based POS Tagger for Hindi using IL POS tag set and obtained an accuracy of 92%.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

(Kalam & Hasan, 2014), developed various stochastic POS Taggers for Bengali where hybrid HMM model with morphological analyzer gives best accuracy of 96.3%. (Saharia et al., 2009), developed HMM based POS Tagger for Assamese using text corpus (Corpus ASM) of nearly 3,00,00 words where 10,000 words were manually tagged for training. 172 tags are used and obtained average accuracy is 87%. (Shambhavi et al., 2012), developed POS Tagger for Kannada language using second order Hidden Markov Model and Conditional Random Fields (CRF) with training data that consist of 51,269 words and test data consists of around 2932 tokens and obtained accuracy is 79.9% for HMM model and 84.58% for CRF model respectively. (Sharma et al., 2011), developed a system using HMM to improve the accuracy of Punjabi POS Tagger. The system was evaluated over a corpus of 26,479 words and achieved the accuracy of 90.11%. (Shrivastava et al., 2008), developed HMM based POS Tagger for Hindi using stemmer and 18 tags and achieved the accuracy of 93.12%. (Singha et al., 2012), developed POS Tagger for Manipuri using HMM and used annotated test set data with 97 morph-syntactic tag categories and obtained an accuracy of 92%.

3. MORPHOLOGICAL CHARACTERISTICS

Assamese is a complex language which is morphologically rich and agglutinative in nature. It is an official language of Assam, one of the North-eastern states in India and spoken by nearly 20 million people in the region. Morphology is the study and description of formation of words in a language. A sentence is formed by a Subject and a Predicate and therefore the Subject, Object and Verb are important elements of a sentence. In a text, Assamese words are inflected with various grammatical features such as Noun is inflected with Case, Number and Gender, Pronoun changes according to Number and Honorary features, Verb is inflected with Number, Person, Tense, Aspect and Mood, etc. Some of the examples are as follows:

Example 1: A proper noun *(Ram)* inflects according to Case (Nominative).

English: Ram is eating rice.

Assamese: Ram-*e* b^{h}at k^{h}a-*i* as-*e* ।

Gloss: Ram-*nom* rice eat-*nf-prog* exist-*3rd P*

Example 2: A proper noun *(Ram)* inflects with Number.

English: Where are Ram and others going?

Assamese: Ram-*hɔt-e* k.*ɔt* go-*i* as-*e* ?

Gloss: Ram-*plu-nom* where.*loc* go-*nf* exist-*3rd P*

Example 3: A common noun *manuh 'human'* inflects with Gender.

Masculine manuh-*zon*

man-cl.mg

Feminine manuh-*zon-i*

woman-*cl-fg*

Example 4: A pronoun *tɔi 'you'* (2nd Person) inflects according to Honorary feature and Number.

Singular

apuni '*you*(*honorific*)'

tumi '*you*(*familiar*)'

tɔi '*you*(*inferior*)'

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Plural

apuna-lʊk '*honorific'*
you.hon-plu
tʊma-lʊk '*familiar'*
you.fam-plu
tɔ-hɔt '*inferior'*
you.inf-plu

Example 5: An Interrogative pronoun *kun-tu 'which'* inflects with Number.

Singular

kʊn-*tʊ*
which- *classifier*

Plural

kʊn-*bʊr*
which-*plural marker*

Example 6: A word *goi ase 'going'* is the continuous form of verb 'go' that inflects with Person.

English: I am going.
Assamese: mɔi go-*i* as-*u* ।
Gloss: I go-*nf-prog* exist-*1st P*
English: You are going.
Assamese: tumi go-*i* as-*a* ।
Gloss: You go-*nf-prog* exist-*2nd P*
English: They are going.
Assamese: xi- *hɔt -e* go-*i* as-*e* ।
Gloss: He-*plu-nom* go-*nf-prog* exist-*3rd P*

Example 7: A verb *kɔ 'say'* inflects according to Aspect (Perfective).

Present: mɔi kɔ *-is -ʊ* । (*I have said*)
Gloss: I say *–perf -1st P*
Past: mɔi kɔ *-is -il -ʊ* । (*I had said*)
Gloss: I say *–perf –past -1st P*

4. THE APPROACH

A supervised learning approach is being followed and therefore, annotated corpus is required for training the system. The corpus is provided by TDIL-DC, Indian Language Technology Proliferation & Deployment Centre and BIS (Bureau of Indian Standards) tagset is used with a total number of 38 annotation tag labels. The corpus contains the sets from three domains that are Agriculture, Entertainment and Science and Technology field which include 271890 words. From the set of corpus, 256690 words are used in training the system and the rest of 15200 words are used for testing. A bigram HMM is used for developing the POS Tagger and then Viterbi algorithm is used as a decoder. The methodology is described as follows:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.1. Hidden Markov Model

A POS Tagger based on HMM (Jurafsky & Martin, 2000) assigns the best tag sequence to a given word sequence or sentence. Normally, the most probable tag sequence is assigned to each sentence by means of the Viterbi algorithm. The task of Part-of-Speech (POS) tagging is to find the best POS tags sequence, $T = \{t_1, t_2, t_3, \ldots, t_n\}$ for a given word sequence $W = \{w_1, w_2, w_3, \ldots, w_n\}$. By following the Bayes' Law, the tagger looks for the maximum value which is equal to $argmax_T P(T) \times P(W|T)$.

The probability of the tag, $P(T)$ can be calculated by Markov assumption which informs that the probability of a tag depends on n number of previous tags. The bigram model is used where the probability of a tag is dependent on its one previous tag, and as a result;

$$P(T) = P(t_1) \times P(t_2|t_1) \times P(t_3|t_2) \times P(t_4|t_3) \ldots \times P(t_n|t_{n-1}) .$$

A generative HMM is defined as:

$$f(s_1 \ldots s_n) = argmax_{t_n}^{t_1} P(s_1 \ldots s_n, t_1 \ldots t_n) \quad (1)$$

where $s_1 \ldots s_n$ are source language phrases and $t_1 \ldots t_n$ are target language phrases. By giving the inputs $s_1 \ldots s_n$, the highest probability phrase sequence is taken as the output of target language. Hence the bigram HMM is calculated as:

$$P(s_1 \ldots s_n, t_1 \ldots t_n) = \pi_i \, q(t_n|t_{n-1}) \, \pi_i \, e(s_i|t_i) \quad (2)$$

where $q(t_n|t_{n-1})$ is the probability of a current tag given the previous tag called transition probability and $e(s_i|t_i)$ is the probability of the sequence of words given the tag sequence called word likelihood probability.

The transition probability is calculated as:

$$q\left(t_n|t_{n-1}\right) = \frac{freq\left(t_{n-1}, t_n\right)}{freq\left(t_{n-1}\right)} \quad (3)$$

Each tag transition probability is computed by calculating the frequency count of two tags seen together in the corpus divided by the frequency count of the previous tag seen independently in the corpus. This is done as because it is more likely for some tags to precede the other tags.

The word likelihood probabilities is calculated using $e(s_i|t_i)$ i.e., the probability of a word given a current tag. This probability is computed as:

$$e\left(s_i|t_i\right) = \frac{freq\left(t_i \rightarrow s_i\right)}{freq\left(t_i\right)} \quad (4)$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Here, the probability of word provided a tag is computed by calculating the frequency count of the tag in the corpus and the word occurring together is divided by the frequency count of the occurrence of the tag alone in the corpus.

4.2. Viterbi Algorithm

According to (Singh et al., 2016), decoding finds the most likely state sequence from given observation $O = o_1, o_2, o_3, \ldots o_n$. The sequence of states is backtracked after decoding the whole sequences. The algorithm is defined in Algorithm 1:

4.3. System Architecture

The Stochastic Approach is followed using Hidden Markov Model that trains the system and for the purpose of decoding, a Viterbi algorithm is used. In the developed model, initially a system corpus is activated that is stored in a system file and then the system is trained. The training corpus is in Unicode encoding text format. Next, the input is provided to the system for tagging. The system displays the output in a list view where in each row a word is displayed along with its corresponding POS tag, description of the tag and the probability measure. The generalized architecture of the system is shown in Figure 1.

The initial process of the developed POS Tagger is a Sentence Detector. Maximum Entropy Sentence Detector from the Open NLP tools is used that splits the input sentences. Then by following the Greedy approach, a Tokenizer separates the words and punctuation marks from the input text. After tokenization, HMM computes the transition and emission probability for each pair of words and tag sequences from the training corpus. Next, a Viterbi Decoder generates the state graph and calculates the state transition probability for each transition in the matrix. Finally, the Viterbi matrix is backtracked by a Tag Sequence Generator that scans the path with the highest probability and then POS tags are assigned accordingly to each and every word in a given input text. As a result, a list of words along with its tag is displayed by the system.

5. EVALUATION

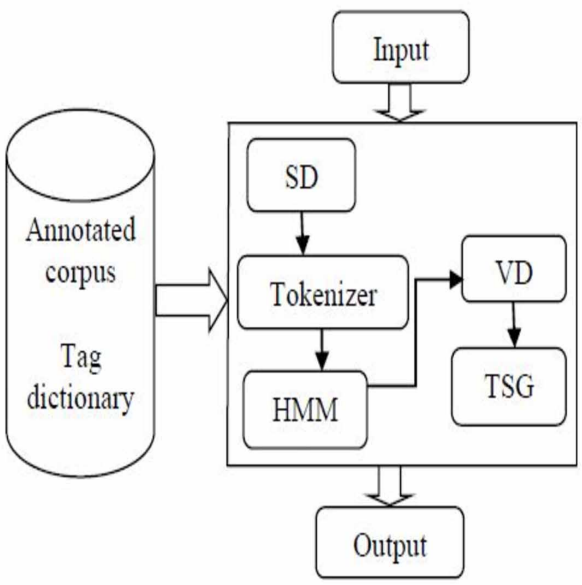
The annotated corpus is divided into two parts where 256690 words are used in training the system and the remaining portion is used for testing. The POS labels of every word in the evaluation set are removed and the text from resulting raw corpus is provided as an input to the system. Finally, the output

Algorithm 1

Input: - A Sentence $x_1 \ldots x_n$ and Parameters $q(t_n|t_{n-1})$, $e(s_i|t_i)$
Define K to set of all tags $K_{-1} = K_0 = (start)$
Π (0, *start*, *start*) = 1
For $k = 1 \ldots n$
For $a \in K_{k-1}$, $b \in K_k$
$\Pi(k,a,b) = argmax$ (π (k-1,a,b) $\times$ $q(b|a)$ $\times$ $e(x_k|b)$)
Return $argmax$ (π (n,b) $\times$ $q(stop|b)$)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Architecture of the developed Assamese POS Tagger



is compared with the data from the evaluation set. Therefore, based on these comparisons, the system's performance is evaluated by computing Precision, Recall and F-Measure respectively. Hence the result is obtained as follows:

No. of tokens in evaluation set = 15200
No. of tokens tagged by system = 15200
No. of correctly tagged tokens by system = 13560

$$Precision\left(P\right) = \frac{No.of\ correctly\ tagged\ tokens\ by\ system}{No.of\ tokens\ tagged\ by\ system} = \frac{13560}{15200} = 0.8921$$

$$Recall\left(R\right) = \frac{No.of\ correctly\ tagged\ tokens\ by\ system}{No.of\ tokens\ in\ evaluation\ set} = 0.8921$$

$$F \rightarrow Measure = \frac{2 \times P \times R}{P + R} = 0.8921$$

The system is capable of assigning tags to every unknown word and thus the Precision and Recall values are equivalent and accordingly the F-Measure gives the same result. Then the Accuracy of the system is computed as:

$$Accuracy\left(A\right) = \frac{No.of\ correctly\ tagged\ tokens\ by\ system}{No.of\ tokens\ in\ evaluation\ set} \times 100 = \frac{13560}{15200} \times 100 = 89.21$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Therefore, accuracy of the developed POS Tagger is 89.21% which is quite promising. It is being observed that the accuracy level of the system progresses as the number of words in a training corpus increases and this could be described by Figure 2.

Next, the performance of the system is compared with other existing HMM based models developed for Indian languages and the records are shown in Table 1.

It is found that the model developed by Kalam and Hasan in 2014 has obtained the highest accuracy of 96.3% where hybrid HMM method is used with a Morphological Analyzer. In the developed model, language analysis method or linguistic rules are not applied. Hence the accuracy of the system is less but it shows little improvement in comparison to existing Assamese POS Tagger and it is achieved by using a large training corpus that includes the sets from three different domains.

Figure 2. Line-graph of performance of the developed POS Tagger

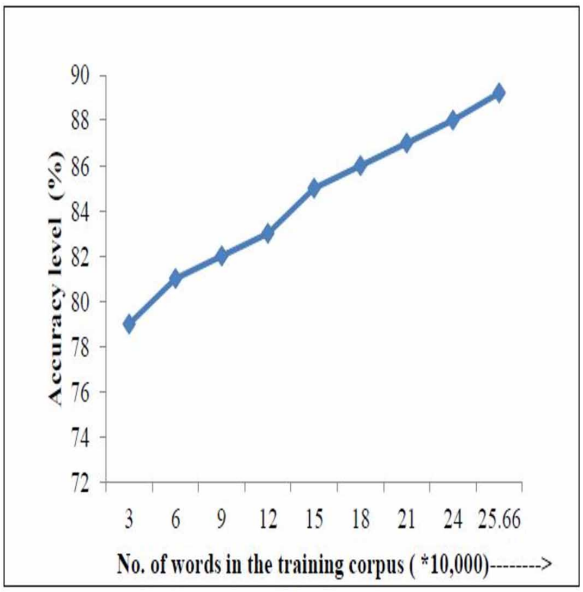


Table 1. Comparison of Developed Model with other HMM based models

Author	Language	Accuracy
Joshi et al. 2013	Hindi	92%
Kalam et al. 2014	Bengali	96.3%
Saharia et al. 2009	Assamese	87%
Shambhavi et al. 2012	Kannada	79.9%
Sharma et al. 2011	Punjabi	90.11%
Shrivastava et al. 2008	Hindi	93.12%
Singha et al. 2012	Manipuri	92%
Developed Model	Assamese	89.21%

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

6. CONCLUSION

The HMM based Assamese POS Tagger is developed by using a training corpus of 256690 words and after testing it on a corpus of 15200 words, the system obtained an accuracy of 89.21%. As Assamese is a fairly free word-order, morphologically rich and agglutinative in nature, POS tagging of Assamese is quite a challenging work. With respect to HMM model, a problem arises when there are less annotated training corpuses because when number of unknown words is greater then it tends to reduce the performance of the tagger. The performance of the developed POS Tagger is still lagging behind other stochastic models and the occurrence of many ambiguous words in test corpus has affected its performance. The result of the system is quite notable although linguistic rules or morphological analyzer is not applied in it. Hence it is assured that the Stochastic Approach performs well with the substantial amount of corpus for morphologically rich Assamese language.

In view of the morphological richness of the Assamese, future development of the system could be incorporating the morphological analyzer along with the HMM to upgrade the performance of tagging. Further enhancement of the system could be achieved by adding specific linguistic rules, suffixation method and procedure for handling unknown words and compound words considering the agglutinative nature of the language. If the present model could be trained using tri-gram HMM then there will be significant increase in the performance of tagging. Finally, it is concluded that Part of Speech tagging is the important field of NLP as because other language processing applications are dependent on the performance of POS Tagger.

ACKNOWLEDGMENT

We express our sincere gratitude to TDIL, Indian Language Technology Proliferation & Deployment Centre for providing the annotated Assamese Text Corpus.

REFERENCES

Dalal, A., Nagaraj, K., Sawant, U., & Shelke, S. (2006). Hindi Part-of-Speech Tagging and Chunking: A Maximum Entropy Approach. In *Proceedings of the Natural Language Processing Artificial Intelligence Machine Learning Competition*. Retrieved from https://www.researchgate.net/publication/241211496

Indian Institute of Technology Guwahati. (IITG). (n.d.). Assamese Design Guide. Retrieved from http://www.iitg.ernet.in/rcilts/phaseI/newassamesedesign.pdf

Joshi, N., Darbari, H., & Mathur, I. (2013). HMM based POS Tagger for Hindi. In *Proceedings of the International Conference on Artificial Intelligence, Soft Computing* (pp. 341-349).

Jurafsky, D., & Martin, J. H. (2000). *Speech and Language Processing* (pp. 285–317). Englewood Cliffs, NJ: Prentice Hall.

Kalam, A., & Hasan, Md. R. (2014). Review of Stochastic POS tagging techniques used in Bengali. *International Journal of Computers and Applications*, *102*(8), 35–39.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Natural Language Processing. (n.d.). (blog site). Retrieved from http://language.worldofcomputing.net

Saharia, N., Das, D., Sharma, U., & Kalita, J. (2009). Part of Speech Tagger for Assamese Text. *In Proceedings of the Association for Computational Linguistics-International Joint Conference on Natural Language Processing, Conference Short Papers* (pp. 33–36).

Shambhavi, B.R. & Ramakanth, P.K. (2012). Kannada Part-Of-Speech Tagging with Probabilistic Classifiers. *International Journal of Computer Applications, 48*(17), 26-30.

Sharma, S. K., & Lehal, G. S. (2011). Using Hidden Markov Model to Improve the Accuracy of Punjabi POS Tagger. In Proceedings of *Computer Science and Automation Engineering*, *IEEE International Conference* (pp. 697-701).

Shrivastava, M., & Bhattacharyya, P. (2008). Hindi POS Tagger Using Naive Stemming: Harnessing Morphological Information without Extensive Linguistic Knowledge. *In Proceedings of the International Conference on Natural Language Processing*. Retrieved from https://www.cse.iitb.ac.in/pb/papers/icon08-hindi-pos-tagger.pdf.

Singh, U., Goyal, V., & Lehal, G. S. (2016). Urdu to Punjabi Machine Translation: An Incremental Training Approach. *International Journal of Advanced Computer Science and Applications*, *7*(4), 227–238. doi:10.14569/IJACSA.2016.070428

Singha, K. R., Purkayastha, B. S., & Singha, K. D. (2012). Part of Speech Tagging in Manipuri with Hidden Markov Model. *International Journal of Computer Science Issues*, *9*(6), 146–149.

The Language Gulper. (n.d.). Assamese. Retrieved from http://languagesgulper.com/eng/Assamese.html

Wikipedia. (n.d.). Natural language processing. Retrieved from https://en.wikipedia.org/wiki/Natural_language_processing

Wikipedia. (n.d.). Part of speech tagging. Retrieved from https://en.wikipedia.org/wiki/Part-of-speech_tagging

This research was previously published in the International Journal of Synthetic Emotions (IJSE), 9(1); edited by Amira S. Ashour and Nilanjan Dey; pages 23-32; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 55
A Survey on Supervised Convolutional Neural Network and Its Major Applications

D. T. Mane
SGGS IE&T, India

U. V. Kulkarni
SGGS IE&T, India

ABSTRACT

With the advances in the computer science field, various new data science techniques have been emerged. Convolutional Neural Network (CNN) is one of the Deep Learning techniques which have captured lots of attention as far as real world applications are considered. It is nothing but the multilayer architecture with hidden computational power which detects features itself. It doesn't require any handcrafted features. The remarkable increase in the computational power of Convolutional Neural Network is due to the use of Graphics processor units, parallel computing, also the availability of large amount of data in various variety forms. This paper gives the broad view of various supervised Convolutional Neural Network applications with its salient features in the fields, mainly Computer vision for Pattern and Object Detection, Natural Language Processing, Speech Recognition, Medical image analysis.

1. INTRODUCTION

With the advent of deep learning, the world has proceeded into the new era of machine learning. With the main intention of getting closer to the original goal of machine learning, that is, Artificial Intelligence, deep learning has opened up new avenues to explore. Artificial Neural Networks (ANNs) are biologically motivated machine learning algorithms applied to solve problems, where conventional approach fails, such as computer vision. It takes in the input, let it be an image or an audio signal, extracts features which describe the input and generalizes these features so that the results obtained can be replicated for other examples of the input. This paper gives an overview of a particular type of ANN, known as supervised

DOI: 10.4018/978-1-7998-0951-7.ch055

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Convolutional Neural Network (CNN) and gives information of its development and results in various fields. Initially, we see the history of CNN followed by its architecture and results of its applications. The references of the few used papers have been mentioned here.

2. HISTORY OF CNN

The idea of modeling brain networks has been a research question even before the advent of computers. In initial stages, neural networks were evaluated with propositional logic. Then with the discovery of concepts such as convolution and back propagation applied to neural networks, NN got better. Until the emergence of GPUs, the computers were not fast enough to implement multi-layer neural networks. So NN were not commercially viable. With the power of GPUs and more efficient algorithms, CNNs can be applied to real life applications (see Table 1).

3. OVERVIEW OF CNN

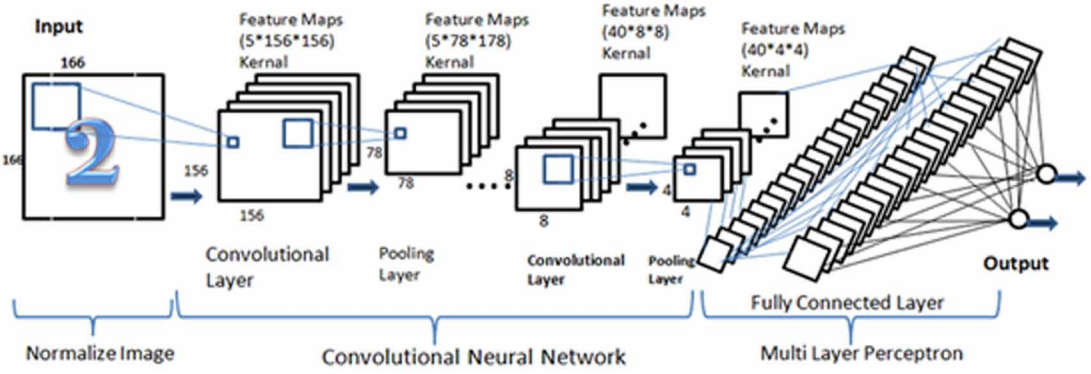
Convolutional Neural Network is a supervised machine learning algorithm where a group of neurons are responsible for recognizing a specific feature and classifying them into categories. These algorithms perform convolution on the input to reduce the size of the input features without adversely affecting the output. This feature could be an edge in an image, or speech subsample. Certain convolution kernels are used so as to highlight the essence of the input especially in case of images, where a lot of unnecessary pixels are present which do not represent any feature. CNN consists of multi-layered neurons, which extract more high level features with consecutive increase in layer. A typical CNN architecture is as shown in Figure 1 (LeCun, Kavukcuoglu, & Farabet, 2010).

Table 1. Time line chart of CNN

Period	Progress	Year	Activity
1940 - 1979	Emergence of NN	1943	McCulloch, Pitts compare neural activity with propositional logic.
		1949	Hebb proposes cell assembly theory.
		1962	Hubel and Wiesel model the visual system in Cat brain.
1980 - 1998	Concept of Convolution	1980	Fukushima proposed a self-learning NN which preserves the intrinsic geometric representation.
		1989	LeCun et al demonstrated the use of back propagated CNN for real life Application.
1999-2010	More Efficient CNN	1999	Poggio et al proposed max pooling.
		2006	Ranzato et al proposed max pooling for CNN.
2011-Present	GPGPU acceleration of CNN	2011	Ciresan et al put forth the concept of CNN on GPUs.
		2012	Hinton et al demonstrated Drop Out for CNN.
		2013	Lecun et al proposed Drop Connect for better CNN.
		2014	Min Lin et al designed Network in Network concept for CNN.
		2015	Google releases different open source libraries for CNN

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Architecture of convolutional neural network



4. APPLICATIONS OF CNN IN VARIOUS FIELDS

Deep convolutional neural network systems are given very good results in some applications like as pattern recognition, natural language processing, computer vision and speech recognitions etc. Figure 2 represent the different areas applications of deep convolutional neural network.

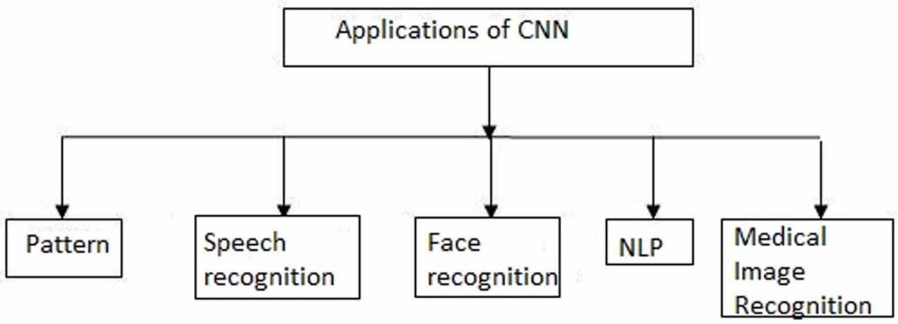
4.1. Pattern Recognition System

Convolutional Neural network has played a very important role in the domains like handwritten digit recognition, logo detection, Vehicle plate detection, etc.

4.1.1. Character Recognition of License Plate Number (2011)

This paper presents machine printed character recognition obtains from license plate with deep convolutional neural network. The numbers of training and test images used in this research are 750 and 434, and its accuracy 98.79%and 1.21% accuracy in misclassification (Radzi & Mohamed, 2011). As compared to previous techniques everything is same except the use of geometrical rules to extract the characters. From observations this work gives better results in terms of reduced number of samples.

Figure 2. Application of convolutional neural network in various fields



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.1.2. Recognizing characters from Natural Scene Images (2012)

This application uses the CNN along with the bimodal image enhancement which helped in improving the CNN classifier. Natural scene images have many variations because of low resolution of image, less clarity, noise, distortion, light reflections, etc. To overcome all these issues some methods have been applied. First maximum separability based color to gray, then normalization based on histogram alignment and finally character image normalization containing shape holding gray-scale is adopted (Zhu, Sun, & Naoi, 2012). As a result, the performance of CNN classifier has increased to 85.96% on ICDAR dataset, more than any previous results.

4.1.3. Detecting and Recognizing Food Images (2014)

Detection and recognition of food images has been performed using CNN classifier along with optimization of parameters. The previous classification algorithms such as SVM which learns from handcrafted features failed to give good results due to large amount of variations in food items. On the contrary CNN learns distinctive features through its internal evaluation (Cagayan, Kiyoharu, & Makoto, 2014). Thus, this paper has shown results 93.8% on food dataset of most common food substance in publically accessible food-logging system. Also, the result shows that color features played a very important role in recognizing the food images. Hence, CNN proved to be an efficient method for food classification.

4.1.4. Pedestrian Detection (2015)

Pedestrian detection is one of the important applications in driving assistant systems. Mainly it is achieved through vision-based methods. So, to achieve high accuracy in this domain CNN along with Random Dropout and Ensemble Interference Network (EIN) has been used (Fukui, Chubu, Takayoshi, Yuji, & Fujiyoshi, 2015). Random Dropouts require a technique in which dropout units are selected on the basis of flexible rate. Along with this, multiple networks with variant structures are constructed in EIN in fully connected layers. As compared to previous techniques the results provided are very accurate.

The summary of the research papers mentioned above is documented in Table 2.

4.2. Face Detection and Recognition

4.2.1. Head Detection (2013)

Head detection has an importance in counting and pedestrian detection. Previously, head detection had low recognition rate as it is based on outline, color and template mainly. But, by the use of CNN with multi-stage weighted feature which combines local and global features, accuracy has been increased tremendously. The recognition rate 0f 99.70% is achieved (Ke, Thuc, Lee, Hwang, Yoo, & Choi, 2013).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Summary of pattern detection

SN	Title	Salient Features
1	Digit recognition from License Number Plate (Radzi & Mohamed, 2011)	Technique used: CNN Idea: CNN has been applied for recognizing digits in license number plates. Several factors to improve the system performance have been discussed like Appropriate learning rate, Activation function, Second order back propagation, Momentum and weight decay, Cross validation. Dataset: Images collected from Malaysian license plates for training 750 images and for testing: 434 images Results: Accuracy: 98.79%
2.	Digit Recognition (Ciresan, Meier, Gambardella, & Schmidhuber, 2010)	Technique used: CNN with back propagation Idea: Extract feature using CNN and GPUs is used for training Dataset: MNIST Results:Error rate: 0.35%
3.	Handwritten Character Recognition (TELUGU) (Soman, Ashakranthi, & Chakravarthy, 2013)	Technique Used: CNN Idea: Multi-class classification of telugu handwritten characters using CNN Dataset: Telugu character dataset for consonant with 36 classes and vowel modifier with 15 classes. Results: Consonants: 92.26% and Vowel modifier: 92%
4	Food Detection and Recognition (Cagayan, Kiyoharu, & Makoto, 2014)	Technique used: CNN Idea: The success of CNNs for food image recognition and detection has been attended to through parameter optimization. Dataset: Constructed from the most frequent food items in a publicly available food-logging system. Results: Testing Accuracy:93.8%
5.	Pedestrian Detection System (Fukui, Chubu, Takayoshi, Yuji, & Fujiyoshi, 2015)	Technique Used: CNN with Ensemble inference network (EIN) Idea: Two CNN-based methods were introduced Random Dropout: for training process EIN for classification process Results: Caltech Pedestrian Dataset: 10.5%.Daimler Mono Pedestrian: 2.0% .Daimler Mono: 31.32%
6.	Recognizing natural scene characters (Zhu, Sun, & Naoi, 2012)	Technique: Convolutional neural network and bimodal image enhancement. This application uses the CNN along with the bimodal image enhancement which improving the CNN classifier. Dataset: ICDAR Result: Performance of CNN classifier increased to 85.96%

4.2.2. Age Estimation (2014)

Human face as biometric characteristics can be used for obtaining lots of information about identity, age, expression, etc. Hence CNN can be used to extract the facial features. By constructing multi-level CNN model based on training data the age groups are prepared and different classifiers can be used for performing classification (Yan, Lang, Wang et al., 2014). Using CNN for extracting aging features from images helps to provide superior results 76.64% than previous methods in age detection.

4.2.3. Multifaceted learning using DCNN for Person Related Categories in Photos (2015)

Multi facet learning has been proposed to improve training of DNN (Deep Neural Networks). In case of existing studies there required a lot of labeled images to train the DNN. The idea is to explore multiple datasets and learning effective representations to learn related visual concepts. It enforces to share features

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

for image recognition. Four types of datasets were used for experimental evaluation. They proved that results related to multi-facet give better performance than person related dataset (Cao, Yan, & Smith, 2015). The first purpose was to use existing deep neural network as an off the shelf feature extractor. Another purpose was to train convolutional neural network on human facet data directly. Error rate is very less as compared to other techniques.

4.2.4. Simultaneous Classification of Several Features of a Person's Appearance Using a Deep Convolutional Neural Network (2015)

A model of CNN has been described for automatic simultaneous extraction of several features of the person's appearance from an image. In an image one can find many features of appearance like sex, face, age, the presence of a moustache or beard, etc. It has various applications like identifying a person's age from a photograph, video surveillance systems, an interactive user interface, and data collection systems for marketing research, etc. The features information can be used to improve the quality of searching for images on the Internet by content. LWF dataset has been used for experimental evaluation. As features of person's appearance it uses sex, moustache and beard. Accuracy for features such as sex-95.13, Mustache-91.58 and beard: 93.80 were given. The algorithm has a high accuracy for the problem of identifying the person's sex and the presence of a moustache and a beard from a photograph (Kukharenko & Konushin, 2015).

The summary of the research papers mentioned above is documented in Table 3.

Table 3. Summary of face detection and recognition

SN	Title	Salient Features
1.	Head Detection (Ke, Thuc, Lee, Hwang, Yoo, & Choi, 2013)	Technique Used: Multistage weighted CNN Idea: Human head detection is major with respect to pedestrian detection. It combines global features with local features. Dataset: Collected large number of positive and negative images from video of passenger flow. Results: Accuracy:99.7%
2.	Age Estimation (Yan, Lang, Wang, Xuetao, & Zhang, 2014)	Technique Used: CNN Idea: CNN used to extract facial features and age has been approximated from these features Dataset: Web Face database Results: Accuracy: 76.64%.
3.	Face Verification (Mohamed & Liew, 2014)	Technique used: Siamese CNN Idea: Siamese architecture with two CNNs has been introduced. Also explored various optimizations that can be applied on Siamese CNN architecture to reduce overall computing time. Dataset: AT&T face database Results: EER (equal error rate):3.33%
4.	Person-Related Categories in Photos (Cao, Yan, & Smith, 2015)	Technique used: CNN with multi-facet Learning Idea: Learning from multiple facets introduces need to share features for image recognition. Dataset: Collected total 33536 images in different category i.e. Affiliation facet: 7424 images, Age facet: 10752 images, Body view: 4352 images, gender facet: 11008 images Results: Recognition error of Multi-facet CNN: Affiliation:0.451 Age:0.373, Bodyview:0.320 Gender:0.332
5.	Classification based on Person Appearance (Kukharenko & Konushin, 2015)	Technique Used: Deep CNN Idea: The projected model can be used for synchronized identification of a better number of features of person's look without a major increase in operating system. Dataset: LWF dataset: 13233 images, Sex, Moustache and beard were used as features of Appearance. Results: Accuracy for each feature: Sex:95.13 Mustache:91.58 beard:93.80

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

4.3. Speech Recognition

4.3.1. A Deep Convolutional Neural Network Using Heterogeneous Pooling for Trading Acoustic Invariance with Phonetic Confusion (2013)

A novel deep convolutional neural network architecture i.e. an HP-CNN followed by a DNN has been proposed, where heterogeneous pooling was used to provide constrained frequency-shift invariance in the speech spectrogram. The combined features of both heterogeneous pooling and dropout regularization were used intelligently (Li, Hamid, & Yu, 2013). TIMIT phonetic recognition task, achieved an 18.7% phone error rate. As a result, by using this for TIMIT core test sets phone detection, error rate has condensed to 19.3% from 20.4% obtained with the optimal but single fixed pooling size. Also after regularizing the CNN using a variant of the dropout technique, the error rate of the HP-CNNDNN dropped further to 16.7%, from 19.9% with the same dropout but without heterogeneous pooling.

4.3.2. Speech Recognition using CNN (2014)

A CNN was proposed for speech recognition. Basically CNN was proposed to increase the performance of ASR (automatic speech recognition). The aim of automatic speech recognition is the record of human speech into spoken words. Also initiate a new method of limited weighted allocation scheme that can hold speech features in a better way than the full weight sharing which was standard in previous CNN architectures (Abdel-Hamid, Abdel-Rahman, Jiang et al., 2014). CNN technique has been used in such a way that it can directly accommodate some types of speech variability. The CNN has been applied on two ASR tasks: TIMIT phone recognition and a large-vocabulary voice search. An improved performance was given by CNN over standard DNN. The error rate was reduced by 8% as compared to previous techniques.

4.3.3. Combining Time and Frequency-Domain Convolution in CNN Based Phone Recognition (2014)

CNN has proved very successful in image recognition because of their tolerance to small translations. Hence it is used for speech recognition by using a spectral representation of input. In the past activities translations along frequency axis and translation along time axis were used separately for speech recognition (Laszlo, 2014, pp 190-194). CNN with combined translation method along frequency and time axes has been proposed here. This method has been used for TIMIT Phone recognition task. And by combining advantages of both frequency domain and time domain achieved the error rate of 16.7% which was better performance over the traditional methods.

4.3.4. CNN based Learning Salient Features for Speech Emotion Recognition (2014)

Speech is one of the most significant and natural means for humans to speak their emotion to others. Therefore, significant way of human emotion understanding Speech Emotion Recognition has attracted great deal of attention. It has applications in HCI (Human Computer Interaction) and effective computing (Mao, Dong, & Huang, 2014).

The summary of the research papers mentioned above is documented in Table 4.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Summary of speech recognition

SN	Title	Salient Features
1.	Trading Acoustic Invariance With Phonetic Confusion (Li, Hamid, &Yu, 2013)	Technique used: DCNN Idea: DCNN is used for minimizing phonetic confusion using dropout method. Results: TIMIT core test set: 19.3% error rate and using dropout HP-CNN-DNN goes down to 18.7%.
2.	Speech Recognition (Abdel-Hamid, Abdel-Rahman, Jiang, & Li, 2014)	Techniques used: CNN Idea: The CNN has been applied on two ASR tasks: TIMIT phone recognition and a large vocabulary voice search. Performance was given by CNN over standard DNN. Result: Error rate was reduced by 8%
3.	Combining T & F Domain convolution in Phone Recognition (Laszlo, 2014, pp 190-194)	Technique Used: CNN Idea: Combination of frequency(F) and time(T) domain translations Dataset: TIMIT phone recognition task Results: Error rate:16.7%
4.	Learning Salient Features for Speech Emotion Recognition (Mao, Dong, & Huang, 2014)	Technique Used: CNN Idea: Learning salient, discriminative features is an important research issue for SER. CNN is used for recognizing these features in speech Dataset: Surrey Audio Visual Expressed Emotion Database (SAVEE), Berlin Emotional Database(Emo-DB) Results: SAVEE-89% Emo-DB-93%

4.4. Natural Language Processing

4.4.1. Sentence Modeling (2014)

The ability to represent sentences correctly is the basic need for understanding any language. The technique known as dynamic CNN can be used efficiently for this. The DCNN apply the dynamic k-max pooling concept as a non-linear subscripting activation function. The high performance was achieved by the network on the question and sentiment classification. The accuracy obtained is 93% on the TREC Questions dataset and 87.4% on the Twitter Sentiment dataset (Kalchbrenner, Grefenstette, & Blunsom, 2014)

4.4.2. Twitter Sentiment Analysis (2015)

Due to the informal nature of language, sentiment analysis in social networking websites provides lots of challenges. Hand crafted features were used for previous NLP methods which resulted in over-specified and incomplete features. This architecture uses multiple filters and non-linear layers for recognizing the relevant features from twitter posts and classifying them. The data used for training contains both Twitters user as well as non-Twitter user data. The result is presented in the form of F1 score of 64.85% Twitter2015 dataset (Severy & Moschitti, 2015).

4.4.3. Semantic Role Labeling (2015)

Semantic Role Labeling is very important for understanding relationships between words. CNN can be used in this problem for extracting semantic roles in texts. SENNA (Semantic Extraction using a Neural Network Architecture) was introduced for semantic relation extraction (Foland & Martin, 2015). This introduces automated semantic role labeling. The result was an f-measure of 0.55 on BC-PPI dataset.

The summary of the research papers mentioned above is documented in Table 5.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Summary of natural language processing

SN	Title	Salient Features
1	Multitask Learning in Natural Language Processing (Collobert & Weston, 2008)	Technique used: CNN Idea: CNN does learning on multiple tasks in a semi-supervised manner and does not rely on parse tree features. Dataset: Prop Bank dataset: Results:14.30% error rate
2	Semantic Role Labeling for Automated Relation Extraction from Biomedical Texts (Barnickel, Weston, Collobert, Mewes, & Stumpflen, 2009)	Technique used: CNN Idea: Authors of the paper use CNN for labeling semantic roles SENNA (Semantic Extraction using a Neural Network Architecture) was introduced for semantic relation extraction Datasets and Results: BC-PPI dataset: F-measure = 0.55 LLL'05 dataset: F-measure = 0.51
3	Sentence Modeling (Kalchbrenner, Grefenstette, & Blunsom, 2014)	Technique used: DCNN with dynamic max pooling Idea: The ability to represent sentences correctly is the basic need for understanding of any language. The authors have described DCNN which uses dynamic k-max pooling operators as a non-linear sub sampling function. Datasets and Results: Twitter sentiment dataset87.4% and TREC questions dataset: 93%
4.	Twitter Sentiment Analysis (Severy & Moschitti, 2015)	Technique used: CNN Idea: Sentiment analysis is extraction of emotion from sentences. This paper shows that CNN could be applied for sentiment analysis Datasets and Results: Twitter2015 dataset: 64.85% F1-score
5	Dependency-Based Semantic Role Labeling (Foland & Martin, 2015)	Technique used: CNN Idea: The authors describe various CNN architectures for semantic role labeling Datasets and Results: Brown Dataset: Sense Accuracy:87.40%

4.5. Medical Image Recognition

4.5.1. Classification of Histopathological images using CNN (2014)

The classification of cellular structure in high resolution histopathological images is a very exhaustive and time-consuming process. The classification of these images provides the easy analysis of cellular structure. Also it is useful in the CAD (Computer Aided Diagnostic) systems by preventing pathologists from wasting their time on image regions where decisions can be made in straightforward manner. The CNN in this domain has been proposed with different spatial window sizes in RGB color spaces. The different pixel neighborhoods can be evaluated with the help of spatial dependencies of cell and non-cell pixels (Hatipoglu & Bilgin, 2014). By using CNN, the cell classification gives better result than SVM and k-Nearest Neighbors. Further, it is also possible to extend the accuracies obtained with convolutional networks by adapting to other domain-specific features.

4.5.2. Region Segmentation in Histopathological Breast Cancer Images using CNN (2015)

Breast cancer is one of the leading concern worldwide and main causes of deaths among women diagnosed with cancer. The images of tumor called histopathology images and their analysis plays very significant role in analysis, prognosis and therapy scheduling for breast cancer. But their manual inspection is time-consuming as well as subjective. Hence CAD systems should be used for this purpose. The

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

system with fast scanning has been used to perform ROI (Region of Interest) segmentation scheme. Each pixel sensible classification is adopted along with optimized forward propagation network. The features are learned by network in supervised manner from raw-pixel values. The efficiency can be obtained by avoiding redundancy in convolutional and max pooling layers (Su, Liu, Xie, Xing, Meyyappan, & Yang, 2015). This method can outperform the various texture classification methods in its segmentation accuracy time cost. In (Su, Liu, Xie, Xing, Meyyappan, & Yang, 2015), proposed method is used to segment a high resolution image (1000*1000) done in only 2.3 seconds. This shows the scalability and accuracy of CNN in biomedical image classification area.

The summary of the research papers mentioned above is documented in Table 6.

5. CONCLUSION

Through Deep Learning is an old concept, due to excessive time required for the extensive computational work, improvements in this particular field have been limited. The revolution occurred in 2009 when NVIDIA introduced its GPU cards which enabled extensive computational work in a shorter period of time. Following this, many researchers used deep learning and neural networks to achieve better results in various fields than expected for e.g. Pattern Recognition, NLP, Computer vision and speech. The convolutional neural network has the structure which has resemblance to human brain neurons and its functionality is same as that of it. It has the ability to extract the features on its own rather than using any handcrafted features. Deep Convolutional Neural Network means Back Propagation algorithm with many hidden layers. Like a human brain, it has deep understanding structure with detailed knowledge

Table 6. Summary of medical image analysis

SN	Title	Silent Features
1	CNN with embedded Fourier Transform for EEG classification (Cecotti & Graeser, 2008)	Technique used: CNN based on Fourier Transform Idea: In this work, authors have proposed a Fourier transform based CNN architecture for EEG activity classification Result: 97.17%
2	Medical Image Classification (Li, Cai, Wang, Zhou, Feng, & Chen, 2014)	Technique used: CNN Idea: Here they have proposed a customized CNN architecture to classify HRCT lung image patches of ILD (interstitial lung disease) patterns Results: Average accuracy: 73%.
3	Classification of Histopathological Images Using Convolutional Neural Network (Hatipoglu & Bilgin,2014)	Technique used: CNN Idea: This work shows classification of cells using CNN. They used different spatial window sizes in RGB color spaces Datasets and Results: Bio-segmentation dataset: 86.88% .
4	Detection of Sclerotic Spine Metastases (Holger, Yao, Lu, Stieger, Burns, & Summers, 2015)	Technique used: CNN Idea: This work shows that CNN can be generalized to tasks in medical image analysis. The result shows that random set of CNN classification can be used to reduce FP.
5	Region Segmentation in Histopathological breast cancer images using CNN (Su, Liu, Xie, Xing, Meyyappan, & Yang, 2015)	Technique used: fast-scanning CNN along with optimized-forward propagation Idea: Region segmentation is important in treatment of breast cancer done by using CAD systems. Dataset: 92 images from TCGA breast cancer dataset Result: Performance accuracy 85%

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

of abstraction layer, we can do machine learning. This paper gives a brief overview on supervised Convolutional Neural Network, its history and the various applications and of CNN. It has limitless applications and can be used on many real life problems. The results of many different applications of deep CNN are good, whose detailed information can be seen one by one in above section 4 and 5. After going through the summary mentioned in Tables 2,3,4,5,6, we observe how recently CNN can be used in various applications. The summary has enabled us to understand advantages of CNN, reasons behind selecting datasets which researchers have used, the technique employed and hence effectiveness of results of a particular CNN. If anyone wishes to work on a particular application using supervised CNN in the future, the above summary can provide valuable directions. We have learnt how researchers, using modified CNN, have found astonishing results in various applications. As computers become more and more powerful, CNN proves to be one of the best algorithms for pattern recognition as well as classification problems.

REFERENCES

Abdel-Hamid, O., Abdel-Rahman, M., Jiang, H., & Li, D. (2014). Convolutional neural networks for speech recognition. *IEEE/ACM Transactions on Audio*, *Speech, and Language Processing*, *22*(10), 1533–1545.

Barnickel, T., Weston, J., Collobert, R., Mewes, H. W., & Stumpflen V. (2009). Large scale application of neural network based semantic role labeling for automated relation extraction from biomedical texts. *Open access on PLoS One*, 4(7), e6393.

Cagayan, H., Kiyoharu, A., & Makoto, O. (2014). Food Detection and Recognition Using Convolutional Neural Network. *Proceedings of the 22nd ACM International Conference on Multimedia,* Orlando, FL, USA (pp. 1085-1088).

Cao, L., Yan, Z., & Smith, J. R. (2015). Multi-facet Learning using Deep Convolutional Neural Network for Person-Related Categories in Photos. *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval*, China (pp. 567-570).

Cecotti, H., & Graeser, A. (2008). Convolutional neural network with embedded Fourier transform for EEG classification. *Proceedings of the 19th International Conference on Pattern Recognition,* Florida, USA (pp. 1-4). 10.1109/ICPR.2008.4761638

Ciresan, D. C., Meier, U., Gambardella, L. M., & Schmidhuber, J. (2010). Deep big simple neural nets excel on handwritten digit recognition. *Neural Computation*, *22*(12), 3207–3220. doi:10.1162/NECO_a_00052 PMID:20858131

Collobert, R., & Weston, J. (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. *Proceedings of the 25th ACM international conference on Machine learning*, NY, USA (pp. 160-167). 10.1145/1390156.1390177

Foland, W. R., & Martin, J. M. (2015). Dependency-Based Semantic Role Labeling using Convolutional Neural Networks. *Proceedings of the Fourth Joint Conference on Lexical and Computational Semantics*, Colorado, USA (pp. 279–288). 10.18653/v1/S15-1033

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Fukui, H., Chubu, U., Takayoshi, Y., Yuji, Y., & Fujiyoshi, H. (2015). Pedestrian detection based on deep convolutional neural network with ensemble inference network. *Proceedings of the 2015 IEEE Intelligent Vehicles Symposium (IV)*, Korea (pp. 223 -228). 10.1109/IVS.2015.7225690

Hatipoglu, N., & Bilgin, G. (2014). Classification of histopathological images using convolutional neural network. *Proceedings of the 4th International Conference on Image Processing Theory, Tools and Applications (IPTA),* Paris, France (pp. 1-6). 10.1109/IPTA.2014.7001976

Holger, R. R., Yao, J., Lu, L., Stieger, J., Burns, J. S., & Summers, R. M. (2015). *Detection of sclerotic spine metastases via random aggregation of deep convolutional neural network classifications. In Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*. Springer International Publishing.

Kalchbrenner, N., Grefenstette, E., & Blunsom, P. (2014). A Convolutional Neural Network for modeling sentences. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, Baltimore, USA (pp. 655-665).

Ke, S. R., Thuc, H. L. U., Lee, Y. J., Hwang, J. N., Yoo, J. H., & Choi, K. H. (2013). A review on video-based human activity recognition. *Open access Computers* (Vol. 2, pp 88-131). Doi:10.3390/computers2020088

Kukharenko, A. I., & Konushin, A. S. (2015). Simultaneous classification of several features of a persons appearance using a deep convolutional neural network. *Pattern Recognition and Image Analysis*, *25*(3), 461–465. doi:10.1134/S1054661815030128

Laszlo, T. (2014). Combining time-and frequency-domain convolution in convolutional neural network-based phone recognition. *Proceedings of the 2014 IEEE International Conference on Acoustics, Speech and Signal Processing,* Florence, Italy (pp. 190- 194).

LeCun, Y., Kavukcuoglu, K., & Farabet, C. (2010). Convolutional networks and applications in vision. *Proceedings of 2010 IEEE International Symposium on In Circuits and Systems (ISCAS),* Paris, France (pp. 253–256). 10.1109/ISCAS.2010.5537907

Li, D., Hamid, O. A., & Yu, D. (2013). A deep convolutional neural network using heterogeneous pooling for trading acoustic invariance with phonetic confusion. *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, BC, Canada (pp. 6669 -6673).

Li, Q., Cai, W., Wang, X., Zhou, Y., Feng, D. D., & Chen, M. (2014). Medical image classification with convolutional neural network. *Proceedings of the 13th International Conference on Control Automation Robotics & Vision (ICARCV),* Singapore (pp. 844 – 848). 10.1109/ICARCV.2014.7064414

Mao, Q., Dong, M., Huang, Z., & Zhan, Y. (2014). Learning Salient Features for Speech Emotion Recognition Using Convolutional Neural Networks. *IEEE Transactions on Multimedia, 16*(8), 2203–2213. doi:10.1109/TMM.2014.2360798

Mohamed, K., & Liew, S. S. (2014). A convolutional neural network approach for face verification. *International Conference on High Performance Computing & Simulation*, Bologna, Italy (pp. 707-714).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Radzi, S. A., & Mohamed, K. (2011). Character recognition of license plate number using convolutional neural network. *Proceedings of the Second international conference on Visual informatics: sustaining research and innovations* (pp. 45-55). Berlin Heidelberg: Springer. 10.1007/978-3-642-25191-7_6

Severy, A., & Moschitti, A. (2015). Twitter Sentiment Analysis Using Deep Convolutional Neural Network. *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, Santiago, Chile (pp. 959-962).

Soman, S. T., Ashakranthi, N., & Chakravarthy, V. S. (2013). An efficient multiclassifier system based on convolutional neural network for offline handwritten Telugu character recognition. *Proceedings of the IEEE National Conference of Communications (NCC),* New Delhi, India (pp. 1-5). 10.1109/NCC.2013.6488008

Su, H., Liu, F., Xie, Y., Xing, F., Meyyappan, S., & Yang, L. (2015). Region segmentation in histopathological breast cancer images using deep convolutional neural network. *Proceedings of the IEEE 12th International Symposium* on *Biomedical Imaging (ISBI),* NY, USA (pp. 55 -58).

Yan, C., Lang, C., Wang, T., Xuetao, D., & Zhang, C. (2014). Age Estimation Based on Convolutional Neural Network. *Proceedings of the 15th Pacific-Rim Conference on Advances in Multimedia Information Processing,* NY, USA (pp. 211-220). 10.1007/978-3-319-13168-9_22

Zhu, Y., Sun, J., & Naoi, S. (2012). Recognizing natural scene characters by convolutional neural network and bimodal image enhancement. *Proceedings of the 4th international conference on Camera-Based Document Analysis and Recognition* (pp. 69-82). Springer Berlin Heidelberg.

This research was previously published in the International Journal of Rough Sets and Data Analysis (IJRSDA), 4(3); edited by Nilanjan Dey; pages 71-82; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 56
From Citizens to Decision-Makers:
A Natural Language Processing Approach in Citizens' Participation

Eya Boukchina
University of Carthage, Tunisia

Sehl Mellouli
Université Laval, Canada

Emna Menif
University of Carthage, Tunisia

ABSTRACT

Citizens' participation is a form of democracy in which citizens are part of the decision-making process with regard to the development of their society. In today's emergence of Information and Communication Technologies, citizens can participate in these processes by submitting inputs through digital media such as social media platforms or dedicated websites. From these different means, a high quantity of data, of different forms (text, image, video), can be generated. This data needs to be processed in order to extract valuable data that can be used by a city's decision-makers. This paper presents natural language processing techniques to extract valuable information from comments posted by citizens. It applies the Latent Semantic Analysis on a corpus of citizens' comments to automatically identify the subjects that were raised by citizens.

INTRODUCTION

In the new era of democracy, governments are encouraged to align with the open government principles that refer, among others, to the involvement of citizens in governments' decision-making processes in order to help increase trust, acceptability, and legitimacy of the decisions. This involvement is referred to as Citizens' participation (Woddford, 2013). In the age of technology, the concept of electronic par-

DOI: 10.4018/978-1-7998-0951-7.ch056

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ticipation (e-participation) emerged as a mean for citizens to express their opinions by using for example forums, dedicated platforms, or social media platforms.

Citizens' participation is a form of democracy in which citizens are part of the decision-making process with regard to the development of their society (Fischer, 2000; Kleinman, 2000; Goven, 2003). It can be defined as an interactive process between citizens themselves and between citizens and government officials in order to significantly contribute in public policy decisions in a transparent and responsible manner (Phillips et al., 2002). The term government refers to the three governmental levels: Federal, Provincial, or Municipal. In today's emergence of ICTs that are being adopted in all life spheres, citizens can participate in these processes by submitting inputs through digital media such as social media platforms or dedicated websites for example. These sources generate a new input that should be considered by decision-makers along with any other source of information that they use, in order to take decisions. The participation of citizens by using ICTs is called electronic participation or e-participation. As stated in (UN/DESA, 2013), the use of ICT in an e-participation process can increase the supply of useful information, enhance consultation, and support decision-making. The citizens' inputs will guarantee transparency of the taken decisions, and will reflect the values of good governance (Fung, 2015).

As stated, there are different platforms that citizens can use to express their opinions that can be for example social media platforms, discussion forums, or dedicated participation platforms. From these different means, a high quantity of data, of different forms (text, image, video), can be generated. This data needs to be processed in order to extract valuable data that can be used by a city's decision-makers. Hence, it becomes necessary to have tools to analyze this data and to provide it, in a graphical user interface, to these decision-makers. This paper will be used in the context of municipal government. Citizens' participation is a core element for the development of cities where each citizen has the opportunity to express his opinion on subjects that may affect him in his daily life.

This paper focuses on textual data. In this context, these opinions are expressed in natural language. Hence, the objective of this paper is to develop a tool to process and analyze the opinions of citizens published on a platform of citizen participation. This tool will be based on natural language processing techniques. The tool will be tested on a particular platform Votepour.ca, that is a Quebec based platform for citizens' participation, on a particular public consultation related to the achievement of a public place in the district of Limoilou in Quebec City. The valuable data in this consultation is the subjects that are raised by citizens. These subjects can be for example safety, culture, or urbanism. The results show that the proposed tool can achieve 82% of good identification of the subjects raised in the different processed opinions.

This paper is structured as follows: Section 2 is a literature review about smart cities and citizens' participation. Section 3 describes the natural language processing technique that was adopted in this paper. Section 4 presents the results of the application of the natural language processing technique. Finally, Section 5 concludes this paper.

LITERATURE REVIEW

Smart Cities

Nowadays, people are more and more connected and are widely using new technologies. Local governments and cities are providing new services to their citizens based on these technologies in order to improve their quality of life. This tendency will be maintained for at least the next twenty years. As a consequence, the

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

concept of smart city is emerging as a key issue for future development of cities. Both academics and industry are exploring this concept and trying to define its purpose. Nevertheless, there is not yet a widely accepted definition of what a smart city is. In fact, several definitions are available in the literature, but do not offer a clear and a unique comprehension of this concept. For example, we can find the following definitions:

- The concept of a smart city refers to the use of information and communication technologies (ICTs) by City governments to better interact with their citizens (Chourabi et al., 2012) and to build a more livable, competitive, and attractive urban environment with smart and agile government (AlAwadhi et al., 2012).
- "A city well-performing in a forward-looking way in various characteristics, built on the smart combination of endowments and activities of self-decisive, independent and aware citizens" (OECD, 2011).
- A city "connecting the physical infrastructure, the IT infrastructure, the social infrastructure, and the business infrastructure to leverage the collective intelligence of the city." (Hartley, 2005).
- "A city to be smart when investments in human and social capital and traditional (transport) and modern (*ICT*) communication infrastructure fuel sustainable economic growth and a high quality of life, with a wise management of natural resources, through participatory governance" (Caragliu et al., 2009).

Even if the most common issue discussed both by scholars and industry is technology, some authors considered that technological propagation is not an end in itself, but only a mean to achieve social and economic growth (Nam et al., 2011). In this context, (Chourabi et al., 2012) developed a conceptual framework of smart city based on eight components that are technology, governance, policies, organization, people and communities, economy, built infrastructure, and natural environment. One of these components is people and communities. In the context of smart cities, citizens will be increasingly involved in the city's life by being more engaged and more participative in the decision-making process of the city.

A city is then a place where people and communities live. A smart city is also a city where people and communities live with a strong focus on their behaviors, knowledge and educational level (Walravens, 2012). As stated, the main goal of a smart city is to improve the quality of life by offering advanced and innovative services for citizens (Piro et al., 2014). In fact, in a smart city, a citizen should be interactive, aware, and self-decisive (OECD, 2011) and should be involved in the development of the city through an active participation (Dameri, 2013) and integration in the city life (Calderoni et al., 2012). In (Komninos, 2006), it is argued that a smart city can be established when citizens have a strong capacity of learning, creativity and innovation in order to be able to properly use the technology and to effectively benefit from it (Coe et al., 2001). Hence, one of the main priorities of a smart city should be the education and the relation between the inhabitants and their governments (Sáez-Martín et al., 2014). Consequently, a smart city should invest on social and human capital by providing institutions of knowledge creation in order to foster the willingness of its population to transform life and work in their region in a significant way.

Citizens' Participation

Citizens' engagement is a form of democracy in which citizens are part of the decision-making process with regard to the development of their society (Fischer, 2000; Kleinman, 2000; Goven, 2003). It can be defined as an interactive process between citizens themselves and between citizens and government

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

officials in order to significantly contribute in public policy decisions in a transparent and responsible manner (Phillips et al., 2002). In other terms, citizens' engagement can be seen as forms of individual or collective actions allowing to identify and to deal with questions of public interest (Hays, 2007). For example, different participatory initiatives were organized in cities such as participatory budgeting (Rios et al., 2005; Wampler, 2010), participatory development (Bhatnagar et al., 1992; Campbell et al., 2003), democracy (Baiocchi, 2005; Mutz, 2006) or street design, school area management (Lim, 2010). In today's emergence of ICTs that are being adopted in all life spheres, citizens can participate in these initiatives by submitting inputs through digital media such as social media platforms or dedicated websites for example. These sources generate a new input that should be considered by decision-makers along with any other source of information that they use, in order to take decisions. The participation of citizens by using ICTs is called electronic participation or e-participation. As stated in (UN/DESA, 2013), the use of ICT in an e-participation initiative can increase the supply of useful information, enhance consultation, and support decision-making. The citizens' inputs will guarantee transparency of the taken decisions, and will reflect the values of good governance (Fung, 2015). The aim of citizen participation is then to strengthen the involvement of citizens in the Decision-making about significant choices that will affect their community.

Several terms were used in the literature to refer to the concept of citizen participation: public participation, citizen co-creation, citizen collaboration, public consultation, and citizens' participation, among others. Therefore, citizen participation implies that everyone participates in all actions that affect their community to some extent in order to achieve certain objectives. It is also seen as a process that allows citizens to participate in a series of actions and have the opportunity to make decisions that affect them, within institutions, programs or the environment. This participation enables them to improve their social situation, hence their importance. Moreover, citizen participation is an intervention that influences political decisions.

Citizen participation can be supported by different tools such as polls, conferences, or dedicated citizen participation platforms. This paper focuses on citizen participation through dedicated platforms. A citizen participation platform is a web-based solution that allows policy-makers, decision-makers and government officials to have an idea about the ideas expressed by citizens. It allows citizens to propose, criticize and share new ideas that affect their day-to-day life. The use of technologies to facilitate the participation process, make it more inclusive and ensure commitment and involvement of stakeholders in a productive and constructive manner (Smith et al., 2009).

NATURAL LANGUAGE PROCESSING TECHNIQUE

Case Study

Our study will be conducted on a particular public consultation conducted by '' Votepour.ca ''. Votepour.ca is a citizen participation platform focused on local development in Quebec City. Votepour.ca was visited by more than 8,000 participants, the platform has a mailing list of more than 1,200 citizens. Votepour.ca also organized 24 citizen participation and public consultation projects in 24 months of existence. Our focus will be with regard to a public consultation related to the construction of a public square in one of the districts of Quebec City. Different citizens have posted, through textual comments on the platform, their opinions and the reasons why they are for or against the achievement of this project. The aim of

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

our work is to produce an algorithm that is able to classify the reasons expressed by citizens according to concepts or categories and to provide a summary to the city's decision-makers so that they can see the thoughts of the citizens and how the projects can be improved before its achievement. In this paper, we only identify the subjects that are expressed in the opinions but we don't yet classify these opinions as for or against the project.

The corpus on which we worked was composed of 304 opinions (comments) that were posted by citizens. We divided this corpus into two sets: learning set and test set. The learning set was of 128 comments and the test set was of 176 comments. All comments are in French. So, all the provided examples in this paper will be in French with respect to the corpus that was used.

Adopted Approach

The objective of our work is to automatically identify the subjects that were raised by citizens in their comments. To this end, we applied the Latent Semantic Analysis (LSA) (Dumais, 2004) on the corpus. We chose this technique because it is based on mathematical calculations, which avoid the treatment of the ambiguity of the natural language. LSA is not a traditional natural language processing or an artificial intelligence program. Instead, it is an automatic mathematical and statistical technique for extracting conceptual relations between terms from a corpus to establish semantic links (Landauer et al., 1998).

LSA can be divided mainly on three steps:

1. The first step takes as input a matrix that contains the frequency of each word according to each subject.
2. The second step applies a decomposition of this matrix into singular values. This decomposition transforms the text space into a new system representing a linear combination of the characteristics of the original words (Walaa et al., 2014).
3. The third step identifies the subject of a comment.

We present hereafter the execution of the three steps applied on the case study.

Building the Matrix

The first step requires building the initial matrix that will be used in the second step of the approach. To this end, we started from the learning set of 128 comments. We looked at these comments and we identified seven subjects that were raised by citizens: social, economy, culture, civic, urbanism, safety, and environment. Then, for each subject, we identified a set of key words that describes them. We found 109 keywords that describe these subjects. Table 1 presents the different subjects and some of their key words.

Figure 1 presents an overview of matrix A that has in its lines the different keywords and in its columns the different subjects with values representing the frequency of each keyword in the subject. For example, the keyword *rencontre* (meeting) was used 18 times in different subjects related to social aspects, and the keyword *commerce* (trading) was used 8 times in subjects related to economy and 1 time to subjects related to urbanism. In our case, the full matrix A has 109 lines and 7 columns. Figure 1 presents part of the matrix A that we refer to by matrix A in the remaining of the paper.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table1. List of concepts and some of their related keywords

Subjects	Keywords
Social	*Rencontre – voisin – appartenance – proximité – souder ...*
Economy	*Commerce – marché – tourisme …*
Culture	*Dynamique – art – spectacle - évènement…*
Civic	*Exprimer – droit – participer – parole …*
Urbanism	*Proximité – commerce – circulation …*
Safety	*Circulation – piéton …*
Environnement	*Chaleur – lutte …*

Figure 1. Matrix A of keywords and subjects

	Social	*Economique*	*Culturel*	*Civique*	*Urbanisme*	*Securite*	*Environnement*
rencontre	18	0	0	1	0	0	0
commerce	0	8	0	0	1	0	0
lutte	0	8	0	0	0	0	1
evenement	0	0	11	0	2	0	0
vie	10	0	2	0	0	0	0
communaute	8	0	1	1	0	0	0
proximite	2	0	0	0	4	0	0
traverse	0	0	0	0	0	3	0
sentiment	9	0	0	0	0	1	0

Matrix Decomposition

The second step of LSA is the decomposition of matrix A as a product of three different matrices. This decomposition transforms the text space into a new system representing a linear combination of the characteristics of the original words (Dumais, 2004). This is represented by the decomposition of the initial matrix in the form of a matrix product of three matrices U, S and V^t as follows: $A = USV^t$ with:

- A: the matrix of keyword frequency according to the different subjects
- U: each column of U contains the coordinates of a term vector
- V^t: each column of V^t represents the coordinates of a subject vector in the semantic space.

For example, considering matrix A of Figure 2, we compute the singular values of A that we present in decreasing order: $\lambda_1 = 576.226$, $\lambda_2 = 130.368$, $\lambda_3 = 127.214$, $\lambda_4 = 16.044$, $\lambda_5 = 9.873$, $\lambda_6 = 0.793$, $\lambda_7 = 0.482$. Then, the matrix U is defined as follows (see Figure 2):

The matrix V is presented in Figure 3.

And finally, the matrix S is presented in Figure 4.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Matrix U after the decomposition of matrix A

$$\mathbf{U} = \begin{pmatrix} -0.7493 & 0.0671 & -0.0785 & 0.0196 & -0.0951 & -0.1851 & -0.0031 & 0.6115 & -0.1013 \\ -0.0008 & -0.5474 & -0.4477 & -0.1097 & 0.0007 & -0.0289 & 0.6978 & 0 & 0 \\ -0.0001 & -0.5360 & -0.4606 & 0.1434 & -0.0008 & 0.0288 & -0.6922 & 0 & 0 \\ -0.0302 & -0.6313 & 0.7548 & 0.0463 & 0.0131 & -0.0308 & -0.0051 & 0.1361 & 0.0949 \\ -0.4205 & -0.0720 & 0.0905 & 0.1053 & -0.0428 & 0.5213 & 0.0392 & -0.4426 & -0.5725 \\ -0.3367 & -0.0246 & 0.0318 & 0.0600 & -0.0495 & -0.7031 & -0.0188 & -0.6115 & 0.1013 \\ -0.0859 & -0.0549 & 0.0283 & -0.9747 & -0.0032 & 0.0184 & -0.1774 & -0.0680 & -0.0474 \\ -0.0020 & 0.0008 & -0.0010 & 0.0036 & 0.9536 & -0.1522 & -0.0067 & 0.0575 & -0.2531 \\ -0.3744 & 0.0337 & -0.0394 & 0.0085 & 0.2776 & 0.4165 & 0.0190 & -0.1726 & 0.7592 \end{pmatrix}$$

Figure 3. Matrix V After the Decomposition of Matrix A

$$\mathbf{V} = \begin{pmatrix} -0.9967 & 0.0423 & -0.0490 & 0.0033 & -0.0140 & 0.0462 & 0.0016 \\ -0.0003 & -0.7591 & -0.6443 & 0.0673 & -0.0003 & -0.0008 & 0.0645 \\ -0.0629 & -0.6230 & 0.7550 & 0.1946 & 0.0029 & 0.0010 & 0.0054 \\ -0.0452 & 0.0037 & -0.0041 & 0.0199 & -0.0460 & -0.9972 & -0.0316 \\ -0.0169 & -0.1778 & 0.1042 & -0.9777 & 0.0045 & -0.0190 & -0.0316 \\ -0.0158 & 0.0032 & -0.0038 & 0.0048 & 0.9988 & -0.0452 & -0.0015 \\ -0 & -0.0469 & -0.0408 & 0.0358 & -0.0003 & 0.0323 & -0.9969 \end{pmatrix}$$

Figure 4. Matrix S after the decomposition of matrix A

$$\mathbf{S} = \begin{pmatrix} 24.0047 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 11.4179 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 11.2789 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4.0054 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3.1422 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.8907 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.6944 \end{pmatrix}$$

Classification of a Comment

Once the decomposition of the matrix A is done, it will be used to classify any new comment. To this end, for each new comment, we create a vector q with the frequency of each keyword in the comment. Hence, the vector q will have 109 lines and one column. For example, let us consider the following comment: ‘’*Une place publique nous permet de rencontrer nos voisins à proximité. Notre communauté travaillera alors à fond pour améliorer la vie de quartier*”. The query vector (part of it) will be in the form of Figure 5. In this comment, we identify the words “*rencontrer*”, ‘’*vie*”, ‘’*communauté*” and ‘’*proximité*” with a frequency of 1. These words are part of the set of keywords.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Query vector Q

$$\mathbf{q} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

One of the issues in LSA is that matrix A may have a high number of columns. Consequently, the matrices can be of large sizes. To this end, LSA proposed to use heuristics to reduce the size of matrices U, S, and V by reducing the size of S. In the following example, we present the case where S is of two dimensions. In Section 4, we present different performances of the algorithm with different reduction possibilities. In this case, the matrices U, S, and V become as follows (Figure 6):

Figure 6. Matrices U, S, and V with k = 2

$$\mathbf{U_k} = \begin{pmatrix} -0.7493 & 0.0671 \\ -0.0008 & -0.5474 \\ -0.0001 & -0.5360 \\ -0.0302 & -0.6313 \\ -0.4205 & -0.0720 \\ -0.3367 & -0.0246 \\ -0.0859 & -0.0549 \\ -0.0020 & 0.0008 \\ -0.3744 & 0.0337 \end{pmatrix}$$

$$\mathbf{S_k} = \begin{pmatrix} 24.0047 & 0 \\ 0 & 11.4179 \end{pmatrix}$$

$$\mathbf{V_k^t} = \begin{pmatrix} -0.997 & -0.0003 & -0.063 & -0.045 & -0.017 & -0.016 & 0 \\ 0.042 & -0.759 & -0.623 & 0.004 & -0.178 & 0.003 & -0.047 \end{pmatrix}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Matrix V_k^t defines the new coordinates of each subject. For example, the new coordinates of the subject social are (-0.997, 0.042).

Then, LSA proposes to transform the query to the new reduced space by computing D_q as follows: $D_q = q^t U_k S_k^{-1}$. Finally, LSA proposes to compute cosine between D_q and the different columns of V_k^t. This will allow computing the similarity between the query and the different subjects using the formula of Figure 7 where c is the vector representing a subject and q is the transformed query in the new space (Abbas, 2014).

From formula of Figure 7, we decided to consider the subjects that form an angle between 0 and $\pi / 4$ or $- \pi / 4$, that is to say, a cosine value belonging to the interval [½, 1] or [-1, - ½].

Improvements to LSA

We found two challenges when building our tool. The first challenge is related to the automatic identification of keywords in order to build queries. The second challenge is the impacts that dominant keywords, that are keywords with high frequencies, and non-dominant keywords, that are keywords with very low frequencies can have on the results of the algorithm.

Syntactic Analysis

The first challenge is the syntactic analysis of the comments in order to build the query of each comment. In fact, we found different words that have the same meaning but written in different forms such as: ''***senti****mental*'', ''***senti****r*'', and ''***senti****ment*''. All these words have the same word root (*senti*) and they all refer to same meaning that is "feel". Hence, we opted that the keywords that will be used in matrix A are the roots of the found keywords. Using the roots of the keywords, it becomes possible to detect almost all forms of the same word using regular expressions. Regular expressions are used to validate any entity containing the root of a keyword followed by any combination of alphabetic characters. For example, using the root ''*senti*'' (feel), we can classify the words ''***senti****mental*'', ''***senti****r*'', and ''***senti****ment*'' as related to feel.

Frequencies of Keywords

The second challenge is related to the fact that we sometimes observed that some keywords have a high frequency that can lead to a dominance of the keyword with regard to other keywords. So, we decided to modify matrix A by reducing the dominance of keywords. To this end, we applied the following formula on the values of matrix A: $A[i,j] = \log (a_{ij} + 1)$ with a_{ij} is the frequency of a keyword at the line i for the subject of column j. With this formula, we compute the logarithm of frequencies of values 0 as the logarithm of 1.

Figure 7. Computing similarities between the query and the subjects

$$\mathrm{sim}(q,c) = \frac{q.c}{|q||c|}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

We also observed that there are keywords with very low frequencies so that they are not informative. To this end, we also applied a new formula on these keywords in order to reduce their impacts (Bestgen, 2004): $\log (a_{ij} + 1) / \sum_j a_{ij} \log (a_{ij} / \sum_j a_{ij})$

We thus defined three algorithms that we adopted in our study:

- Algorithm 1: Classical latent semantic analysis.
- Algorithm 2: Latent semantic analysis and reduction of the impact of keywords with high frequencies.
- Algorithm 3: Latent semantic analysis and reduction of the impact of very frequent and uninformative words.

The results presented in Section 4 will compare these different algorithms along with different reduction options of matrix S.

RESULTS

We present the results based on the learning corpus and then on the test corpus.

Learning Corpus

The learning corpus has 128 comments that were studied and that allowed to detect the different subjects and their corresponding keywords. We applied the three algorithms to the learning corpus using respectively k = 5, and k = 6, that is to say for a semantic space of respectively 5 and 6 dimensions. We obtained the following results by comparing with what we did manually and with what the algorithms provided as results as depicted in Figures 8 and 9. We observe, when applying these algorithms on the learning comments, that the percentage of good answers given by each of the three algorithms for k = 6 is greater than those obtained for k=5. This percentage exceeds 80% for all the three cases. This can ensure that the algorithms are providing satisfactory and effective results that give a good idea about the opinions expressed by the citizens. The results show that the third algorithm with k=6 is the most efficient since it gives the highest percentage for the good responses, that is around 83.44%. It also gives the lowest percentage for non-good responses that is about 16.56%.

For example, let us consider the following comment: ‘*’Rassembler les Limoiliens, ajouter plus d’activités dans notre quartier magnifique, et avoir des spectacles plus souvent*!” (“Gather citizens of Limoilou, add more activities in our magnificent neighborhood, and have shows more often!”). This comment can be classified in culture and social subjects. We applied the three algorithms on this comment with different values of k as depicted in Table 2. In bold are the identified subjects from the comment. As stated, a comment is associated to a subject if the cosine value is in one of the intervals [½, 1] or [-1, - ½].

We observe that for algorithm 1, for k=3 and k = 4, it identifies the social subject but also the civic subject (which is not correct). For k = 5 and k = 6, the algorithm identifies only the social subject. For all values of k, the algorithm doesn’t identify the culture subject. For algorithms 2, and 3, and for the values of k = 3 and k = 4, the social and culture subjects are identified but along with other subjects. Instead, in the cases of k = 5 and k = 6, the algorithms identify the correct subjects social and culture.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Results of the three algorithms using k=5

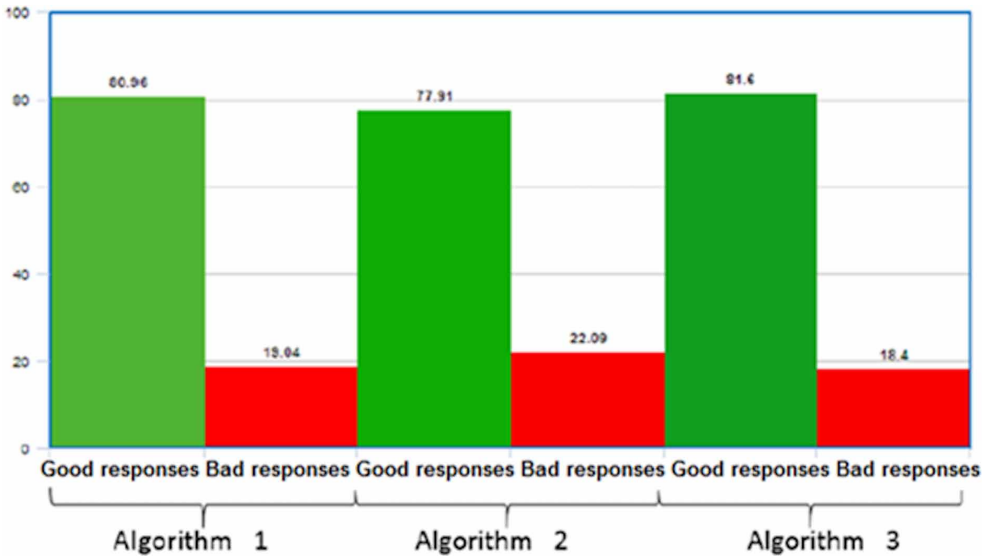
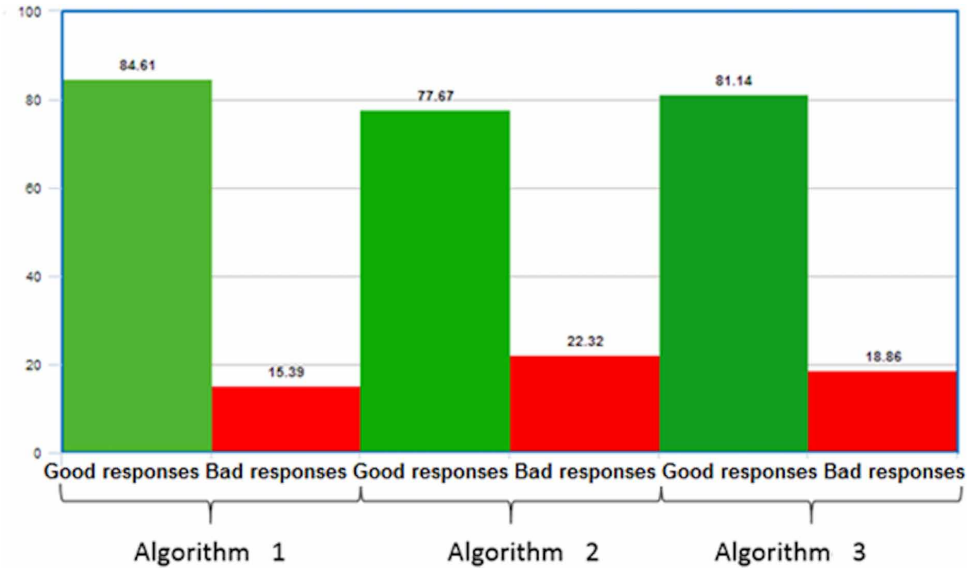


Figure 9. Results of the three algorithms using k=6



Test Corpus

As a second step to validate the three algorithms, we applied them on the test corpus with k = 6 since we saw that from the learning corpus that the algorithms are better performing when k = 6. These results are presented in Figure 10.

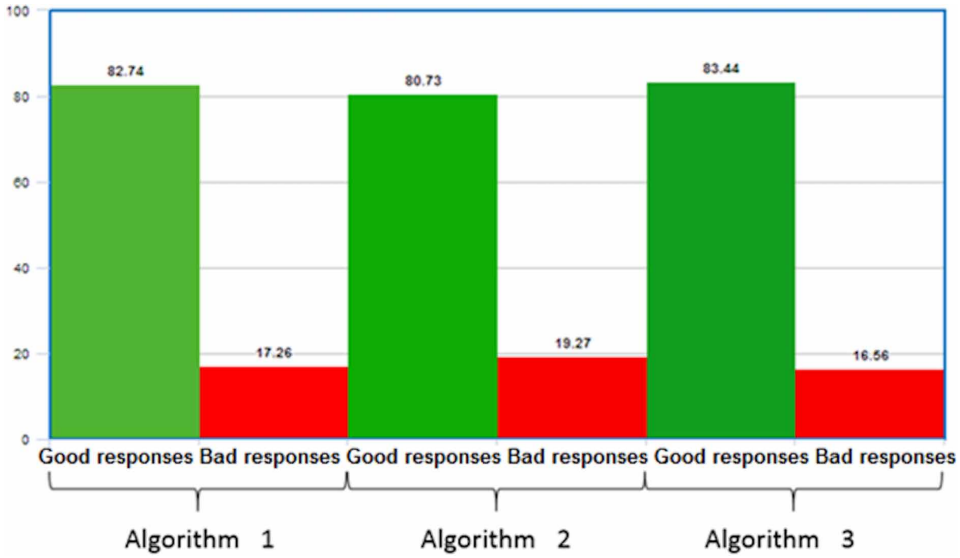
The results on the test corpus showed that algorithms 2 and 3 performed better than when they were applied on the learning corpus where their rates improved by 2%. On the other hand, the performance

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Classification with regard to the three algorithms and with different values of k

	Algorithm 1	Algorithm 2	Algorithm 3
K = 3	Social: **0.94570923** Economy: -0.01283143 Culture: 0.3242482 Civic: **0.94084877** Urbanism: 0.22138334 Safety: 0.35551345 Environment: NaN	Social: **0.56763292** Economy:-0.04091533 Culture: **0.8202374** Civic: **0.52977095** Urbanism: **0.74882765** Safety: **0.8213787** Environment: NaN	Social: **0.79066893** Economy: -0.23146049 Culture: **0.59922094** Civic: **0.59704038** Urbanism: 0.45806082 Safety: **0.89466409** Environment: NaN
K = 4	Social: **0.94420128** Economy: -2.83E-06 Culture: 0.32939209 Civic: **0.93640425** Urbanism: 4.21E-05 Safety: -0.00457991 Environment: NaN	Social: **0.55392333** Economy: -0.00136204 Culture: **0.8314237** Civic: **0.5668778** Urbanism: 0.01007994 Safety: -0.12079681 Environment: NaN	Social: **0.73690267** Economy: -0.09192901 Culture: **0.61901469** Civic: **0.60722365** Urbanism: 0.19635074 Safety: 0.17202804 Environment: NaN
K = 5	Social: **0.94420503** Economy: -8.13E-07 Culture: 0.32938469 Civic: -4.24E-06 Urbanism: 1.23E-05 Safety: -0.0017291 Environment: NaN	Social: **0.72825737** Economy: 0.00501568 Culture: **0.70080993** Civic: 0.1133114 Urbanism: -0.01112602 Safety: 0.07986009 Environment: NaN	Social: **0.72825737** Economy: 0.00501568 Culture: **0.70080993** Civic: 0.1133114 Urbanism: -0.01112602 Safety: 0.07986009 Environment: NaN
K = 6	Social: **0.9442** Economy: -1.36E-17 Culture: 0.329372 Civic: -4.03E-17 Urbanism: -1.78E-17 Safety: 7.56E-17 Environment: NaN	Social: **0.5547** Economy: -2.01E-16 Culture: **0.83205** Civic: 4.35E-16 Urbanism: -2.33E-15 Safety: 5.89E-16 Environment: NaN	Social: **0.728745** Economy: 3.34E-16 Culture: **0.684786** Civic: -2.42E-16 Urbanism: -1.83E-16 Safety: 3.55E-17 Environment: NaN

Figure 10. Results of the three algorithms on the test corpus with k=6



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

of algorithm 1 decreased by 2%. Since 2% is not considered as a high value, then the performances of the three algorithms can then be considered as constant.

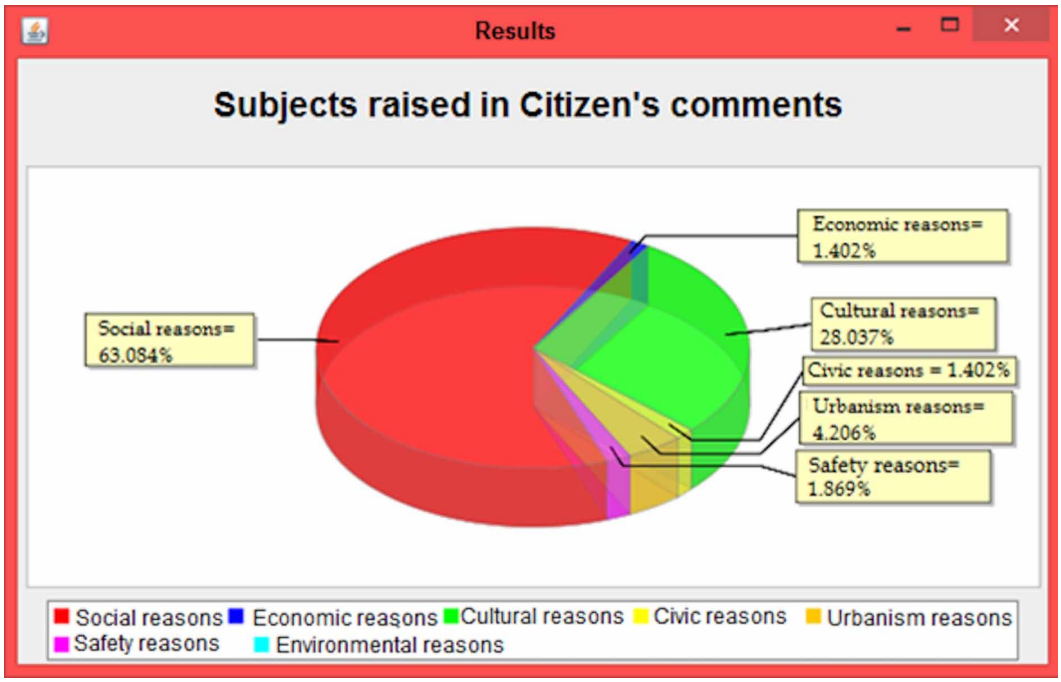
Once the subjects are identified, results can be presented to decision-makers. Figure 11 shows an example of how presenting resulting data to decision-makers. The subjects of Figure 11 are those obtained when applying algorithm 3 with k=6. This data, with a rate of 82% of good classification, can then be provided to decision-makers at city level to have an idea about what people are thinking about this project. We assume that with a rate of 80% of good classification, decision-makers have a true picture of what citizens are thinking and consequently can base their decisions on this picture.

Although we have a rate of 82% of good classification, the algorithms were not able to classify 18% of the comments. This requires further investigation on these algorithms in order to improve the classification rate.

CONCLUSION

In this paper, we addressed the issue of how to identify the subjects of comments posted by citizens in a citizens' participation platform. To this end, we adapted the Latent Semantic Analysis technique by the development of three different algorithms. Results showed that the third algorithm along with k=6 has the best performance rate with regard to the two other algorithms. We also developed a graphical user interface to present the subjects discussed by citizens in a simple and clear graphical representation showing the percentage of each discussed subject. This interface can facilitate the use of the results by decision-makers at city level in order to enlighten their decisions.

Figure 11. Subjects weights using the third algorithm with k=6



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

While the obtained results can be considered as good, there are still challenges to improve the obtained results. First, the proposed algorithms can be improved by working with an expert that can validate the set of keywords used to define the subjects. Keywords identification is a crucial step since it is the starting point of the algorithms. If the keywords are not valid, then all the obtained results will not be valid. Second, there are new techniques in the semantic analysis of text such as LDA (Blei et al., 2003). It then becomes important to apply these new techniques on the same corpus and compare the techniques against each other. This will allow choosing the most appropriate technique. Finally, it will be a challenge to build a complete application for comments analysis. This application will help users to assign keywords to subjects, and then apply the algorithms to finally show results. At the stage of this paper, these three steps are conducted separately.

ACKNOWLEDGMENT

The authors would like to thank the Canadian Internet Registration Authority for funding this research paper.

REFERENCES

Abbas, N. (2014). Vers une extension sémantique de l'analyse formelle de concepts: Ap-plication à la recherche d'informations [PhD thesis]. *Faculté de génie électrique et d'informatique, Université Mouloud Maameri de Tizi Ouzou, Algeria.* Retrieved from http://www.ummto.dz/IMG/pdf/ABBAS_Nacira-pdf.pdf

Alawadhi, S., Aldama-Nalda, A., Chourabi, H., Gil-Garcia, J. R., Leung, S., Mellouli, S., . . . Walker, S. (2012). Building Understanding of Smart City Initiatives. In H.J. Scholl, M. Janssen, M.A. Wimmer et al. (Eds.), *Electronic Government: Proceedings of 11th IFIP WG 8.5 International Conference, EGOV 2012, LNCS* (Vol. 7443, pp. 40-53). Berlin, Germany: Springer.

Baiocchi, G. (2005). *Militants and citizens: the politics of participatory democracy in Porto Alegre*. Palo Alto, CA: Stanford University Press.

Bestgen, Y. (2004). Analyse sémantique latente et segmentation automatique des textes. In *Les actes des 7èmes Journées internationales d'Analyse statistique des Données Textuelles* (pp. 171-181).

Bhatnagar, B., & Williams, A. C. (Eds.). (1992). Participatory development and the World Bank: potential directions for change. World Bank Publications.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*(1), 993–1022.

Calderoni, L., Maio, D., & Palmieri, P. (2012). Location-aware Mobile Services for a Smart City: Design, Implementation and Deployment. *Journal of Theoretical and Applied Electronic Commerce Research*, *7*(3), 74–87. doi:10.4067/S0718-18762012000300008

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Campbell, L. M., & Vainio-Mattila, A. (2003). Participatory development and community-based conservation: Opportunities missed for lessons learned? *Human Ecology*, *31*(3), 417–437. doi:10.1023/A:1025071822388

Caragliu, A., Del Bo, C., & Nijkamp, P. (2009). Smart cities in Europe. *In Proceedings of the 3rd Central European Conference in Regional Science*. Retrieved from http://www3.ekf.tuke.sk/cers/cers2009/PDF/01_03_Nijkamp.pdf

Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J. R., Mellouli, S., Nahon, K., ... Scholl, H. J. (2012). Understanding Smart Cities: An Integrative Framework. In *Proceeding of the 45th Hawaii International Conference on System Sciences*. New York: ACM.

Coe, A., Paquet, G., & Royn, J. (2001). E-governance and smart communities: A social learning challenge. *Social Science Computer Review*, *19*(1), 80–93. doi:10.1177/089443930101900107

Dameri, R. P. (2013). Searching for smart city definition: A comprehensive proposal. *International Journal of Computers & Technology*, *11*(5), 2544–2551.

Dumais, S. T. (2004). Latent semantic analysis. *Annual Review of Information Science & Technology*, *38*(1), 188–230. doi:10.1002/aris.1440380105

Fischer, F. (2000). *Citizens, experts, and the environment: The politics of local knowledge*. Durham, NC: Duke University Press. doi:10.1215/9780822380283

Fung, A. (2015). Putting the Public Back into Governance: The Challenges of Citizen Participation and Its Future. *Public Administration Review*, *75*(4), 513–522. doi:10.1111/puar.12361

Goven, J. (2003). Deploying the consensus conference in New Zealand: Democracy and de-problematization. *Public Understanding of Science (Bristol, England)*, *12*(4), 423–440. doi:10.1177/0963662503124006

Hartley, J. (2005). Innovation in governance and public services: Past and present. *Public Money & Management*, *25*(1), 27–34.

Hays, R. A. (2007). Community activists' perceptions of citizenship roles in an urban community: A case study of attitudes that affect community engagement. *Journal of Urban Affairs*, *29*(4), 401–424. doi:10.1111/j.1467-9906.2007.00353.x

Kleinman, D. L. (2000). *Science, technology, and democracy*. New York, NY: SUNY Press.

Komninos, N. (2006). The architecture of intelligent cities; integrating human, collective, and artificial intelligence to enhance knowledge and innovation. In *Proceedings of the 2nd International Conference on Intelligent Environments*. 10.1049/cp:20060620

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, *25*(2-3), 259–284. doi:10.1080/01638539809545028

Lim, J. H. (2010). Empowering citizens' voices in the era of e-government: Implications from South Korean cases. *Theoretical and Empirical Researches in Urban Management*, *7*(16), 19–31.

Mutz, D. C. (2006). *Hearing the other side: Deliberative versus participatory democracy*. Cambridge, United Kingdom: Cambridge University Press. doi:10.1017/CBO9780511617201

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Nam, T., & Pardo, T. A. (2011a). Conceptualizing smart city with dimensions of technology, people, and institutions. In *Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times*. New York: ACM. 10.1145/2037556.2037602

OECD. (2011). *M-Government, Mobile technologies for responsive government and connected societies. Technical report, OECD Publishing*. Paris: OECD.

Phillips, S. D., & Orsini, M. (2002). *Mapping the links: Citizen involvement in policy processes*. Ottawa, ON: Canadian Policy Research Networks.

Piro, G., Cianci, I., Grieco, L. A., Boggia, G., & Camarda, P. (2014). Information centric services in Smart Cities. *Journal of Systems and Software*, *88*, 169–188. doi:10.1016/j.jss.2013.10.029

Rios, J., Insua, D. R., Fernandez, E., & Rivero, J. A. (2005). Participatory budget formation through the web. In M. Böhlen, J. Gamper, W. Polasek et al. (Eds.), E-Government: Towards Electronic Democracy, LNAI (Vol. 3416, pp. 268-276). Berlin, Germany: Springer.

Sáez-Martín, A., Haro-de-Rosario, A., & Caba-Perez, C. (2014). A vision of social media in the Spanish smartest cities. Transforming Government: People. *Process and Policy*, *8*(4), 521–544.

Smith, S., & Dalakiouridou, E. (2009). Contextualising public (e) Participation in the governance of the European Union. *European Journal of ePractice*, *7*, 4-14.

UN-DESA/Division for Public Administration and Development Management (DPADM) Development Management Branch. (2013). *Developing capacity for participatory governance through e-participation: engaging citizens in policy and decision-making processes using ICTs.*

Walaa, M., Hassan, A., & Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, *5*(4), 1093–1113. doi:10.1016/j.asej.2014.04.011

Walravens, N. (2012). Mobile Business and the Smart City: Developing a Business Model Framework to Include Public Design Parameters for Mobile City Services. *Journal of Theoretical and Applied Electronic Commerce Research*, *7*(3), 121–135. doi:10.4067/S0718-18762012000300011

Wampler, B. (2010). *Participatory budgeting in Brazil: contestation, cooperation, and accountability*. Pennsylvania: Penn State Press.

Woodford, M. R., & Preston, S. (2013). Strengthening citizen participation in public policy-making: A Canadian perspective. *Parliamentary Affairs*, *66*(2), 345–363. doi:10.1093/pa/gsr065

This research was previously published in the International Journal of E-Planning Research (IJEPR), 7(2); edited by Carlos Nunes Silva; pages 20-34; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 57
Logs Analysis of Adapted Pedagogical Scenarios Generated by a Simulation Serious Game Architecture

Sophie Callies
University of Québec à Montréal, Canada

Eric Beaudry
University of Québec à Montréal, Canada

Mathieu Gravel
University of Québec à Montréal, Canada

Josianne Basque
TELUQ University, Canada

ABSTRACT

This paper presents an architecture designed for simulation serious games, which automatically generates game-based scenarios adapted to learner's learning progression. We present three central modules of the architecture: (1) the learner model, (2) the adaptation module and (3) the logs module. The learner model estimates the progression of the development of skills targeted in the game. The adaptation module uses this estimation to automatically plan an adapted sequence of in-game situations optimizing learning. We implemented our architecture in Game of Homes, a simulation serious game, which aims to train adults the basics of real estate. We built a scripted-based version of Game of Homes in order to compare the impact of scripted-based scenarios versus generated scenarios on learning progression. We qualitatively analyzed logs files of thirty-six adults who played Game of Homes for 90 minutes. The main results highlighted the specificity of the generated pedagogical scenarios for each learner and, more specifically, the optimization of the guidance provided and of the presentation of the learning content throughout the game.

INTRODUCTION

Serious games can be defined as computer-based learning environments, which are built as video games. Their purpose is to train the learner[1] to a specific domain of expertise, while making learning fun. As their use continues to grow in education and training (Girard et al., 2013), few researchers have studied

DOI: 10.4018/978-1-7998-0951-7.ch057

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

their architecture to date (e.g., Amory, 2007; Arnab et al., 2015; Kiili, 2005; Peirce et al., 2008). But none of proposed architecture for serious game offers enough details on their software implementation. Moreover, few allow game logs extraction and their analysis, which provide educators with details about learners' learning process. The architecture must offer tools to help them analyze extracted logs after a game session. Generating logs and developing logs analysis tools constitute main research objects in the learning analytics field (Baker & Inventado, 2014).

Simulation serious games are serious games which use a simulation to convey learning content. Simulations can be defined as simplified environments of reality, where learner can assimilate learning content without any risk (Sauvé & Kaufman, 2010). Simulation games are appropriate to develop expertise in a domain by generating realistic real-world scenario in which gamers are cognitively as well as motivationally engaged at a deep level (Wouters et al., 2013).

In this article, we present a simulation serious game architecture, which automatically generates pedagogical game-based scenarios adapted to the learner progression throughout the play as well as logs that could be retrieved and analyzed by trainers after the play. We hypothesized that automatically generating scenarios in simulation serious games improve and accelerate learning among adults.

This paper is organized as follow. We first present the theoretical framework of our work. Our proposed architecture is detailed in the second section. We then explain how we implement our architecture in a single-player simulation serious game called *Game of Homes*, which aims to train adults to the basics of real estate. Finally, we report and discuss the results of an empirical evaluation of our architecture, which was conducted to (1) verify that our implemented architecture allowed logs retrieval, (2) read generated pedagogical game-based scenarios, and (3) compare learning progression between scripted-based scenarios and automatically generated scenarios.

THEORETICAL FRAMEWORK

Simulations offer a simplified version of reality, but remain complex and dynamic, and preserve crucial pedagogical features (Garris et al., 2002). Galarneau (2005) mentions that a simulation serious game must, in addition, include play and fun features. The latter maintain learners' motivation and learners' engagement at a high level, using for example scoring or competition. Figure 1 shows simulation, video game and pedagogical features, which all need to be integrated in a simulation serious game architecture.

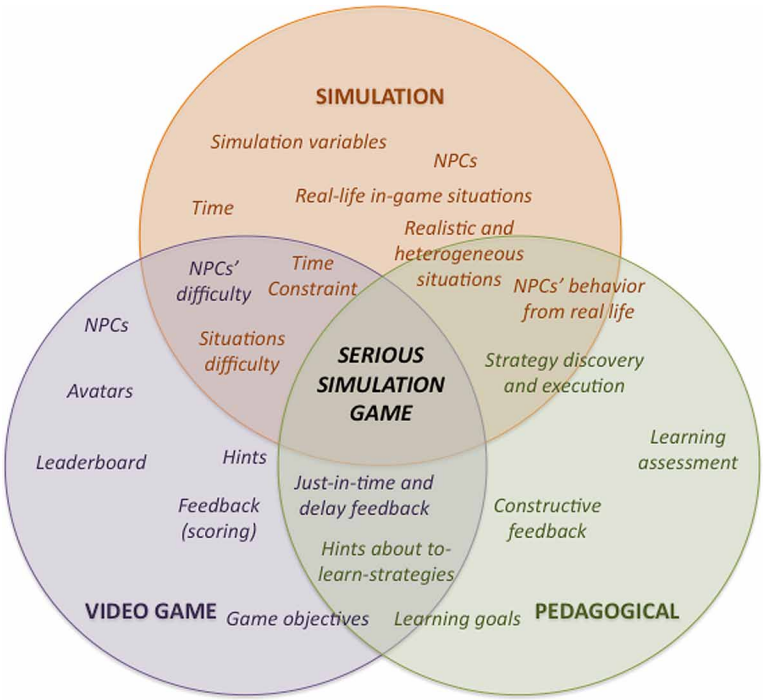
Simulation Key Features

Simulation games complexity and dynamism lie in simulation variables, which are controlled and adjusted by the game system. These variables represent real-life elements interacting with each other, according to learners' actions in the simulation serious game (Gredler, 2004). Each action represents an input, which triggers targeted variables adjustment in the simulation. Visual indicators are then displayed to learners, as outputs as well as feedback, giving learners some piece of information about the interaction mechanisms between variables.

While controlling and changing variables values, simulations can also reproduce specific contexts or real-life situations. This feature is strongly related to game and pedagogical features, as changing in-game situations can increase or decrease difficulty, as well as challenge and learners' curiosity. Moreover, presenting a wide range of real-life situations can raise learners' awareness of the learning content.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Simulation, video game and pedagogical features in simulation serious games



Simulation games can also control Non-Playable Characters (NPC). Observing NPCs' simulated behaviors allow learners to uncover rules about these real-life inspired behaviors. This feature is strongly related to game and pedagogical features, as changing NPCs' behaviors can increase or decrease difficulty as well as challenge and competition in the game. Moreover, observing NPCs' behaviors can lead to understanding their role on the simulation.

The last simulation feature presented in this section is time. Simulations represent a succession of events, which occurs not necessarily in real time: in this case, time needs to be compressed while still offering a realistic game experience. Succession of events in the simulation needs to remain consistent as realistic. This feature is strongly related to game features, as sensation of passing time enhances learners' engagement and interest.

Video Games Key Features

Many aspects contribute to make simulation serious games fun. In our architecture, we focused on the following video game features: (1) avatars, which are controlled by learners and by the game system for NPCs, (2) direct feedback, i.e. scoring system, (3) leaderboard, which shows learners' score rank among NPCs, (4) feed, which shows relevant actions displayed for learners, and (5) hints, e.g. textual advices displayed according to learners' actions in the game in order induce strategies discovery.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Pedagogical Key Features

Pedagogical features in simulation serious games focus on the targeted domain knowledge and skills to be developed by learners, particularly on learning strategies to discover and to efficiently execute. In the design process of a simulation serious game, we recommend to identify early learning content, in order to model relevant skills and associated strategies, as well as important learning goals to achieve.

Learning goals will become game objectives in the simulation serious game and need to be clearly explained to learners, either before playing or within the game. Learning goals are represented by (1) the role played by learners that mimics a real-life professional role, (2) the role that should be applied in different in-game situations to experiment numerous professional situations, and (3) the experience of learners' decision outcomes to foster skills development (Gredler, 2004).

To achieve learning goals, the simulation serious game system must monitor all learners' actions in the game and have the ability to (1) assess learning progression, and (2) provide constructive feedback to learners, which can be given 'just-in-time' or delayed, i.e. association of prompt hints and rise or drop of the score.

We stress the importance of strategy discovery and execution to acquire relevant skills in a simulation serious game. This last feature is allowed by the ability for the simulation system to create different in-game situations, forcing learners to discover other strategies in order to achieve the game objectives. Ultimately, at one point in the game, we want learners to execute in a flexible way several strategies according to different in-game situations, as cognitive flexibility is an important aspect of skills development (Spiro et al., 1992).

Simulation, video games and pedagogical features need to be balanced when they are integrated in simulation serious games. This process is arduous, and needs to be framed by a game architecture.

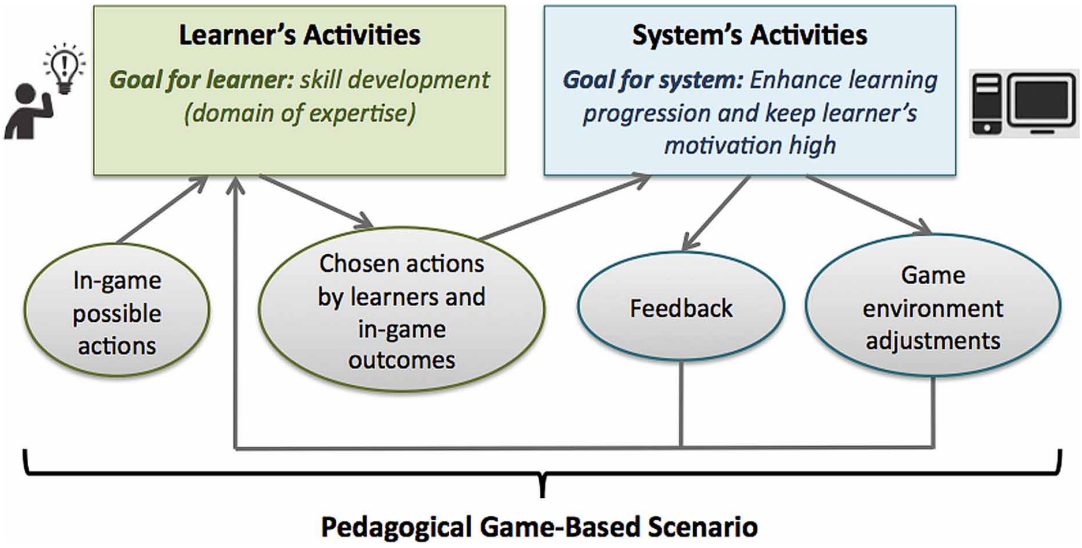
Pedagogical Game-Based Scenarios in Simulation Serious Games

Pedagogical scenarios include two types of scenarios in a learning system: the learning scenario, which includes the set of learning activities prescribed to learners, and the teaching scenario, which includes the set of activities prescribed to teachers or tutors (human or software agents) (Paquette, 2004). As shown in Figure 2, in serious games, learning scenarios target one or several learning goals and allow learners to execute some actions (learners' activities) while receiving relevant feedback in order to progress through the scenario. The game system provides this feedback by displaying textual information, hints, help messages, score updates, and so on, and then modifies the game environment according to the current scenario (system's activities). Both activities, from learners and game system, should lead to the development of target skills (Paquette, 2004).

In most serious games, the sequence of learners' and system's activities is predefined, and the same for each game session (Lopes & Bidarra, 2011). A similar approach is found in video games, in which game scenarios are often scripted or programmed in advance (Schell, 2008). Adopting the scripted-based scenarios approach allows for scenario reusability as well as control of learning progression in the system (Lopes & Bidarra, 2011; Millington & Funge, 2009; Schell, 2008). Designers can therefore make sure that learners have correctly developed the targeted skills before moving on to the next activity. Scripted-based scenarios can be presented an infinite number of times, allowing skills practice and therefore skills development. However, it can also be a limitation, due to the lack of consideration of the variety of learners, who have different prior knowledge about the domain of expertise, and different

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. A schematic overview of a pedagogical game-based scenario



learning styles, in both quality and speed (Lopes & Bidarra, 2011). Moreover, there is a risk that learners get used to these repetitive scenarios as they begin to predict what comes next. Scenario repetition may induce monotony and decrease learners' engagement in the game or in the learning experience (Lopes & Bidarra, 2011). Furthermore, scripted-based scenarios may force users to execute a few winning strategies to solve problems. Finally, as (1) time in simulation serious games passes continuously, and (2) simulation serious games are usually not organized in levels, scripted-based scenarios do not seem to be the best approach to present pedagogical content. A game session can be characterized by playing a certain amount of time in the simulation, and ends when the learner quit the game or when a game objective is reached. Therefore, simulation serious games fall into category of open games, in which it is very difficult to constrain any sequence of activities. Learners are free to explore the game environment, while simulation variables behave autonomously, and discover at their own pace skills to learn and strategies to execute.

Another way to present pedagogical game-based scenarios to learners is thus needed. In simulation serious games, a solution could be to conceive open pedagogical game-based scenarios, which can be modified and personalized in order to match learners' progression (Law & Sun, 2012; Pernin & Lejeune, 2004). When the game system is able to automatically and dynamically adjust the content of pedagogical game-based scenarios according to learners' progression in the game, we call that system adaptable (Vandewaetere et al., 2012). Adaptation in video games and serious games was proposed by some authors who mainly focused on adaptation of challenge in games (Lopes & Bidarra, 2011; Beaudry et al., 2010). Others focused on adaptation of narrative scenarios in serious games (Li & Riedl, 2010), and adaptation of pedagogical game-based scenarios in serious games (Niehaus & Riedl, 2009). Automatic adaptation should propose content relevant to learning progression and therefore allow usage of multiple strategies fostering skills development. An adaptative game system should be able to fit each learner characteristics, by providing more difficult learning contexts to expert learners to keep them interested, and by providing easier learning contexts to beginner learners to prevent game abandon. Finally, adapt-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ing scenarios guarantees personal pedagogical path for each learner and for each game session. In order to adapt learning content to learners' progression, artificial intelligence techniques can be used (e.g., Conati & Zhao, 2004; Peirce et al., 2008; Steiner et al., 2009).

Logs Analysis in Simulation Serious Games

More and more game designers conceive game systems allowing logs recording. Game logs are therefore saved data during a game session; "[they] are captured at the level of a basic action, defined as the smallest unit of meaningful action that a player can exert on the game world" (Harpstead et al., 2013, p. 82). Game logs also carry information about the game context, such as the initial conditions of a game session. Finally, game logs are represented by sequential data: all data is inscribed according to the time the actions occurred (Sanchez et al., 2015).

At the end of a game session, game designers and teachers should be able to extract game logs in order to assess learning and target learning difficulties encountered by learners. Logs extraction is even more important in case of adaptative simulation serious games, as each pedagogical game-based scenario is adapted. However, analyzing game logs can be challenging, especially for open simulation serious games, which are "characterized by having a set of overarching rules that govern the interactions of the game's elements, as well as a non-restrictive structure that allows players to explore multiple paths to a solution" (Harpstead et al., 2013, p. 80). This type of games allows learners to experience many learning contexts, but makes it difficult to prove learning effectiveness.

Therefore, not only designers and teachers need to access the exact content of the pedagogical scenario during and after a game session, but they also need tools to visualize these data.

SIMULATION SERIOUS GAME ARCHITECTURE

When conceiving our architecture, our goal was to meet the needs of what a simulation serious game should contain in order to (1) automatically generate pedagogical game-based scenarios, (2) present these scenarios to learners while inducing learning and challenge, and (3) allow efficient logs extraction. Therefore, our architecture respectively includes three main components: (1) game system, (2) simulation interface and (3) logs module. Figure 3 shows the overall of our simulation serious game architecture, of which each component will be described in the following sections.

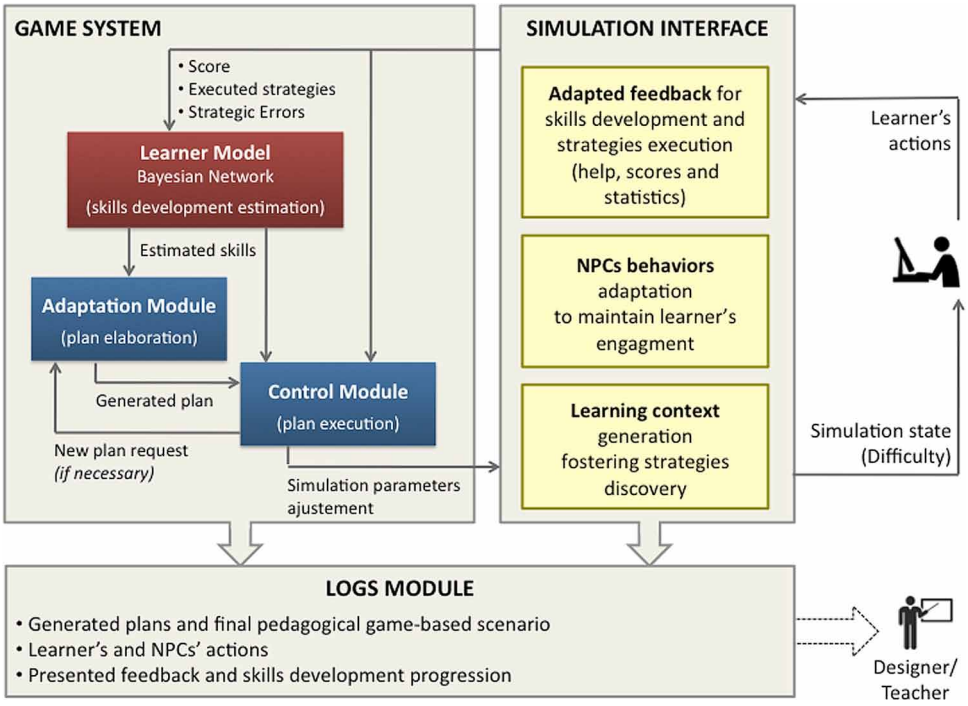
Game System

The main role of the game system is to adapt learning content to learners' progression in the game. It should be able to modify the evolution of the simulation and thus it creates new in-game situations inducing new strategies discovery, which is needed to achieve learning goals. This dynamic evolution of the simulation world represents the automatic generation of pedagogical game-based scenarios to support learners' learning process.

In order to adapt learning content and in-game situations in the simulation, the game system first needs to assess current learners' learning state. The learner model in our architecture carries out this role.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Simulation serious game architecture overview



Learner Model

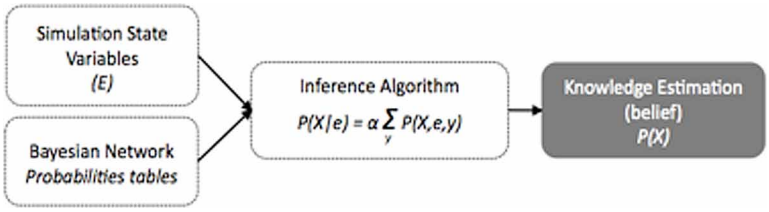
The learner model represents the learner's knowledge progression status, in real time, about the domain of expertise to-be-learned. Estimating acquired knowledge is quite arduous, as we cannot directly observe learners' knowledge state according only to game actions. A well-known artificial intelligence model, which ensures knowledge estimation, uses Bayesian Networks (BN). This model was indeed previously used to assess knowledge acquisition in intelligent tutoring systems (e.g., Conati et al., 2002) and serious games studies (e.g., Manske & Conati, 2005).

BN design requires a previous work of knowledge explicitation from the domain of expertise. As every domain of expertise contains many complex skills, it is important to sparingly choose the specific skills that will be implemented in the game system. We also stress the importance of explicitation of strategies related to skills and executed by experts. These strategies represent the "know-how" aspect of the domain of expertise, and could also be used in the BN to assess learners' knowledge. Once skills and strategies are chosen, probabilities for each skill according to specific actions in the game can be assigned (probabilities tables). These probabilities need to be adjusted during game testing.

A BN operates as shown in Figure 4: an inference algorithm calculates a "belief" about a target skill in the game according to the formula which includes (1) the state of the variables in the simulation, also called random variables; and (2) the corresponding probabilities found in probabilities tables. Random variables represent both observable learners' knowledge and simulation state. Knowledge estimation can be provided continuously during a game session by combining learners' actions with other game elements like learners' statistics or prior encountered game situations. This update status of learners' skills development can then be addressed to the adaptation module.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Bayesian Network (BN) role in our architecture



Adaptation Module

The adaptation module adapts simulation content according to learners' skills development estimation. Simulation content is represented by feedback, NPCs' behaviors and simulation parameters. The game system needs to find the best pedagogical game-based scenario to both speed up and enhance learning process, which represents a complex problem to solve.

In artificial intelligence, a technique called classical planning is used to find optimal solutions to complex problems. A planning algorithm enables the selection of a sequence of actions or situations – also called plan – to achieve a goal. Every action has preconditions, which are needed to execute the action, and effects, which are produced by the action. In our architecture, the adaptation module uses a planning algorithm to discover the best sequence of in-game situations in accordance with the current learners' knowledge status. This sequence of in-game situations represents a plan, which can be recalculated along with learning progression. The final sequence, which took place during a game session, represents the generated pedagogical game-based scenario. The planning process is represented as such in the adaptation module:

- **States:** Learner's knowledge Estimation status;
- **Actions:** In-game situations Which include a pedagogical goal to achieve (previously designed);
- **Actions Preconditions:** Prior knowledge needed to achieve the situation (previously designed);
- **Actions Effects:** Supposed impacts on the learner's learning (previously designed).

Control Module

The control module executes plans generated by the adaptation module. When preconditions of the next in-game situation are satisfied (i.e. if the required skills estimation value is reached), the control module modifies in real time the simulation parameters in order to present new situations to the learner, to change NPCs' behaviors, and to provide feedback to the learner. If the preconditions are not validated (i.e. if the required skills estimation value is not reached yet, or if it was reached too fast), the control module asks the adaptation module for a new generated plan, which will be more adapted to learners' current knowledge status.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Simulation Interface

The simulation interface presents different in-game situations. These situations are composed of simulations variables as well as different features. In our architecture, we selected the following key features exposed in introduction as simulation contents:

- Feedback (i.e. scoring system, learners' statistics, and help messages) in order to optimize skills development and strategies execution;
- NPCs' behaviors, which can be modified in order to trigger challenge;
- Different learning contexts, which can be generated in order to prompt strategies change as well as strategies discovery.

According to these specific situations, every action made by the learner in this interface is monitored by the game system and sent to the learner model.

Logs Module

The main role of the logs module included in our architecture is to collect pertinent game data about:

- Learners' and NPCs' actions;
- Game contents: displayed help messages, numbers of NPCs and details about their behaviors;
- Learners' learning progression: skills development, probabilities estimated by the learner model;
- Generated plan(s) content.

Since the adaptation module can refine the pedagogical scenario at any time during the game, at the end of the game session, the logs module provides the final version of the pedagogical scenario and information about the number of a new plan was generated (if that was the case).

IMPLEMENTATION OF OUR ARCHITECTURE: THE CASE OF *GAME OF HOMES*

To implement our architecture, we developed a single-player serious simulation prototype called *Game of Homes*[2], which aims to train beginners at real estate transactions. Figure 5 shows the graphical interface of *Game of Homes*.

Brokers in *Game of Homes*, represented by both the learner's avatar and NPCs, have to sell houses and to earn as much money as possible while maintaining a good reputation. To do this, brokers negotiate with homeowners to get a brokerage contract, assess property values, and manage sales by carrying out advertising, giving house tours to potential buyers, adjusting selling prices, negotiating with buyers and recommending offers to sellers. In *Game of Homes*, the in-game clock simulates time, but the learner can also move a cursor on a timeline to accelerate or slow down simulation speed.

Game of Homes aims to foster learning of three main skills, as well as strategies associated to these skills. Skills and strategies are detailed in Table 1.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Game of Homes interface

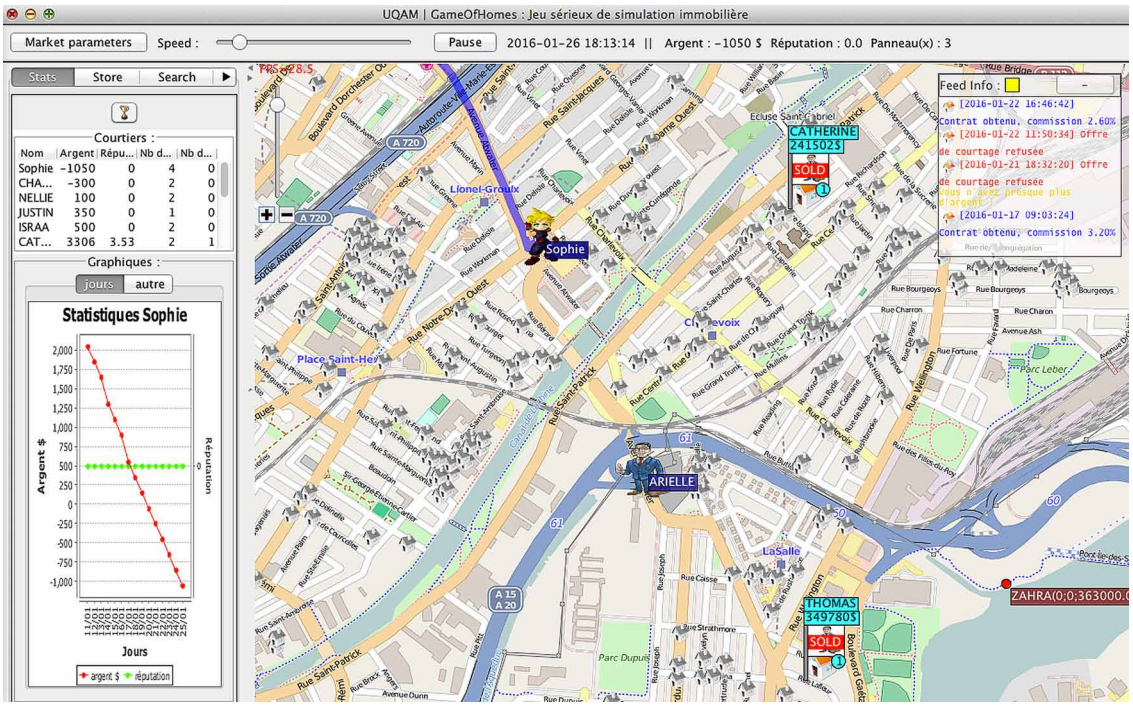


Table 1. Skills and strategies taught in Game of Homes

Skills	Associated Strategies
Skill 1: Obtaining relevant brokerage contracts	Strategy 1.1: Choosing contracts closed to each other in order to be fast and efficient in sale management (territory strategy).
	Strategy 1.2: Choosing contracts whose owners are not too demanding (i.e. owners expect high selling prices and demand low commission rates from brokers).
Skill 2: Estimating houses selling prices	Strategy 2.1: Fixing selling prices according to comparable properties in the neighborhood.
	Strategy 2.2: Reevaluating selling prices to foster difficult sales.
Skill 3: Managing and concluding house sales	Strategy 3.1: Being proactive on visits management.
	Strategy 3.2: Maintaining a high visibility level on houses, by purchasing publicity.
	Strategy 3.3: Being proactive on purchase offers management, and efficiently advising sellers about offers.

Estimation of skills development is done by the Bayesian Network (BN), which considers learner's outcome action and the probability of action in the game to succeed (outcome of the sale property or relevance of the brokerage contract). This probability depends on the simulation parameters (selling market, sellers' demands, learner's territory, expected commission) and the learner's knowledge related to this action.

The adaptation module in *Game of Homes* was implemented with eight generic in-game situations, all detailed in Table 2, which can be part of the generated plan:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Eight possible in-game situations composing a pedagogical game-based scenario in Game of Homes (BN = Bayesian Network)

In-Game Situation	Target Skill(s)	Precondition(s) (% of the Skill(s) Value(s) in the BN)	Game Contents Adjustments				Expected Effects on % of Skills Values in the BN
			Feedback Help Messages	NPCs' Behaviors		Learning Context	
				Aggressive Level of Other Brokers	Level of Sellers' Demand		
1	1 (basics)	≤ 35	Activated	Low	Low	Seller's market	+ 20 on skill 1
2	2 (basics)	≤ 35	Activated	Low	Low	Seller's market	+ 20 on skill 2
3	1 (advanced)	>35 & ≤ 70	Activated	Medium	Medium	Balanced market	+ 20 on skill 1
4	2 (advanced)	>35 & ≤ 70	Activated	Medium	Medium	Balanced market	+ 20 on skill 2
5	1 (expert)	>70	Deactivated	High	High	Buyer's market	+ 20 on skill 1
6	2 (expert)	>70	Deactivated	High	High	Buyer's market	+ 20 on skill 2
7	1 & 2 (expert)	>55 & ≤ 80	Deactivated	High	High	Buyer's market	+ 10 on skills 1 & 2
8	1 & 2 (expert)	>80	Deactivated	Very high	Very high	Buyer's market	+ 10 on skills 1 & 2

- Two easy situations to teach skill 1 (relevant contracts) and skill 2 (selling prices estimation) basics;
- Two medium situations to develop these skills at an advanced level;
- Two difficult situations to develop these skills at an expert level;
- Two very difficult situations targeting the two skills, one validating the understanding of the selling process and the management of several contracts, the other one testing if learners manage the two skills in "crisis" situations.

Each in-game situation is activated if specific preconditions are fulfilled. Preconditions are represented by percentage values of the two skills taught in the game, which are themselves given by the BN in the learner model. For example, situation 5 can only be activated if the percentage value of skill 1 "Obtaining relevant brokerage contracts" is superior to 70%. Finally, each situation has a specific effect on the percentage values of the two skills. For example, situation 5 leads to a 20% improvement of skill 1 probability in the learner model. Each situation has the same game duration of one month in the game.

The control module in *Game of Homes* adjusts the following simulation contents:

- **Feedback:**
  - Scoring system represented by (1) money earned by selling properties, and (2) broker reputation;
  - Learner's statistics represented by (1) a graphic showing score, and (2) a leaderboard gathering NPCs' and learner's scores;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

  - Help messages displayed on the main screen, which target strategies related to skills. For example, when a learner does not execute the territory strategy, the following message pops up: "Choosing contracts closed to each other will allow you to limit transit time and to be more responsive."
- **NPCs' Behaviors:**
  - NPCs brokers may behave more or less aggressively. Both total number of brokers on the map and maximum number of contracts brokers can obtain determine their aggressiveness;
  - House owners may or may not be very demanding: they can overestimate or underestimate houses selling prices, and they can expect reasonable or low commission rates from brokers;
- **Learning Context:** The market may be in favor of buyers (buyer's market), in favor of sellers (seller's market), or balanced (balanced market). Selling context in the game is adjustable by modifying buyers/sellers' ratio on the map.

Finally, the logs module in *Game of Homes* collects pertinent game data:

- **Learner's and NPCs' Actions:** Movements on the map, proposed commission rates, accepted or refused contracts, proposed selling prices, purchased advertises, accomplished or missed houses visits, purchase offers recommendations, sales outcomes, etc.;
- **Game Contents:** Help messages displayed, numbers of NPCs on the map and details about their behaviors (aggressiveness, demand status), market status;
- Learner's learning progression (skills 1 and 2 probabilities, which are provided by the learner model, score and statistics);
- Time at which any event occurred;
- Precise content of each generated pedagogical game-based scenario.

Teachers or game designers should be able to read game logs in order to (1) detect encountered difficulties by learners, (2) check score progression, (3) observe executed strategies, and (3) track the amount of time needed to learn skills. Therefore, we have developed a log analyzer software module capable of parsing and analyzing log files to show its users all pertinent data. This program, based on extracting different types of data from the logs, gives its users information on learners' movements, statistical evolution, successful and unsuccessful sales, along with all information describing the causes (number of visits, asking price, recommended price, etc.). Finally, our parsing analyzer program allows for a qualitative analysis of skill 3, "Managing and concluding house sales", which was not estimated in the BN. The program can indeed evaluate learners' management process (e.g., missed visits rate, missed purchase offers).

EMPIRICAL EVALUATION

This last section aims to point out how logs can be retrieved in order to read the exact content of pedagogical game-based scenarios. Furthermore, we report results about the impact on learning when pedagogical scenarios are automatically generated versus when they are scripted. More precisely, our goal was to test if automatically generated scenarios are more respectful to learning progression than scripted-based scenarios. We also wanted to verify that the logs module along with our designed parsing tool provides a precise analysis of learning progression throughout the game session.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In order to reach our research objectives, we designed another version of *Game of Homes* allowing the presentation of scripted-based scenarios. For this version, the adaptation module was disabled: for each learner, situation 1 is first activated (in January), followed by situations 2, 3, 4, 5, and 6. The script continues until situation 6 is reached (or stops before, if the game session ends). At the end of situation 6, the learner should have developed both skills (probabilities for both skills are higher than 95%). Situations 7 and 8 were not considered in this version, because game sessions would have been too long (more than two hours).

We chose to conduct a qualitative analysis, which focused on (1) data on skills development provided by the learner model, (2) pedagogical scenarios proceedings, (3) amount of help messages displayed and (4) data on acquired strategies. We used our tool to analyze game logs and to visualize the acquisition of specific strategies taught in the game.

Participants and Procedure

The empirical evaluation was conducted with 36 participants who were recruited by advertisements placed on our university campus. The participants were between 18 and 40 years old (14 women, 22 men), and were not and had never been house owners.

Participants were invited into a computer lab on campus, where the experimenter explained the research project and its objectives and carried out a brief demonstration of *Game of Homes*. Once they signed the consent form, they were randomly assigned to one of two versions of *Game of Homes*: automatically generated version (18 participants) and scripted-based version (18 participants). They were then asked to play *Game of Homes* for approximately 90 minutes. After playing, the experimenter conducted a debriefing and answered participants' questions.

Results

We qualitatively analyzed all 36 log files provided by the logs module. It is important to note that the first in-game month (January) in the generated version of *Game of Homes* was used to assess a first-time skill 1 and skill 2 in order to generate an adapted plan (situation 1 was activated during January). Therefore, the data analyzed was taken from second simulated month onward (February). Also, since learners controlled time speed inside the game, none of them reached the same date at the end of their game session, even if they all played for 90 real-time minutes. This aspect made comparing the two versions of *Game of Homes* quite difficult and is the reason why we chose to conduct a qualitative analysis instead of a quantitative analysis. Table 3 presents all generated scenarios with their corresponding in-game situations for the 18 participants who were assigned to the generated version of the game. In-game situation are detailed in Table 2.

Extracted data from the logs module informed us about the progression of skills development for both skills "Obtaining relevant brokerage contracts" (skill 1) and "Estimating houses selling prices" (skill 2). Moreover, the logs module allowed us to retrieve the exact content of generated plan(s) for each learner, and therefore the exact content of the final generated pedagogical game-based scenario. Figure 6 shows total number of situations included in the first generated plan for the eighteen participants assigned to the generated version of the game.

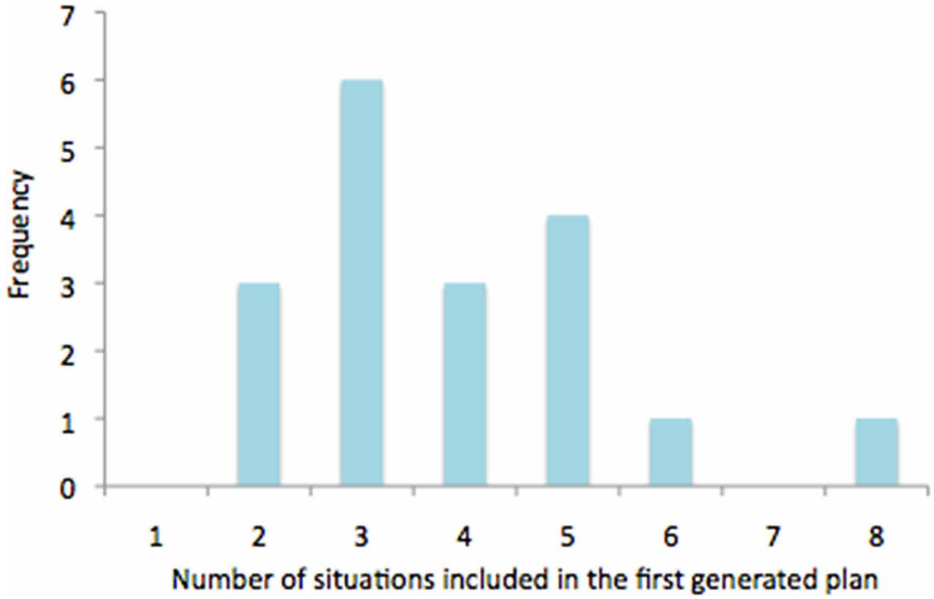
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Details on generated pedagogical game-based scenarios for participants with the generated version of Game of Homes (N=18)

Participant ID	Game Date (Month)	Skills Probabilities (BN)		Details about Generated Pedagogical Game-Based Scenarios
		P(Skill 1)	P(Skill 2)	
2	February	60%	52%	First generated plan including four situations: 3, 5, 4, 6
4	February	47%	17%	First generated plan including eight situations: 3, 2, 4, 7, 5, 5, 4, 6
	March	83%	26%	No new plan
6	February	86%	20%	First generated plan including four situations: 5, 2, 4, 6
	March	83%	26%	No new plan
8	February	99%	57%	First generated plan including two situations: 4, 8
10	February	68%	17%	First generated plan including six situations: 2, 4, 7, 5, 4, 6
	March	97%	29%	Second generated plan including four situations: 2, 4, 4, 8
12	February	98%	11%	First generated plan including five situations: 2, 2, 4, 6, 8
13	February	79%	47%	First generated plan including three situations: 5, 4, 8
16	February	79%	43%	First generated plan including two situations: 4, 7
	March	88%	60%	Second generated plan including two situations: 4, 8
	April	99%	73%	Third generated plan including one situation: 6
	May	99%	92%	No new plan
17	February	60%	20%	First generated plan including five situations: 3, 5, 2, 4, 6
	March	74%	26%	No new plan
20	February	97%	30%	First generated plan including three situations: 2, 4, 8
22	February	70%	20%	First generated plan including four situations: 5, 2, 4, 8
25	February	43%	20%	First generated plan including five situations: 3, 5, 2, 4, 6
26	February	99%	41%	First generated plan including three situations: 4, 4, 6
	March	99%	57%	No new plan
27	February	97%	30%	First generated plan including three situations: 2, 4, 8
	March	99%	30%	Second generated plan including four situations: 2, 2, 4, 8
30	February	46%	30%	First generated plan including five situations: 3, 2, 7, 5, 6
	March	99%	55%	Second generated plan including two situations: 4, 8
32	February	99%	20%	First generated plan including three situations: 2, 4, 6
34	February	95%	53%	First generated plan including two situations: 4, 8
35	February	90%	20%	First generated plan including three situations: 2, 4, 8
	March	90%	30%	Second generated plan including six situations: 5, 2, 2, 4, 4, 6

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Total number of situations included in the first generated plan for each learner (N=18)



Looking at Table 3 and Figure 6, we can appreciate the range of first generated plans among learners. Ones are composed of only two situations, whereas others are composed of six or more situations. If we look closely at data showed in Table 3, we can see that:

- As participant 8 obtained rather good skills estimation (skill 1 = 99%; skill 2 = 57%), the generated plan included only two situations (one to teach skill 2 at an advance level, and one to propose a very difficult situation to teach both skills at an expert level);
- As participant 4 obtained rather poor skills estimation (skill 1 = 47%; skill 2 = 17%), the generated plan included eight situations focusing on both skills with a wild range of level of expertise.

Figure 7 shows chosen situations by the adaptation module for the first three months among learners. Data showed once more the variety in plans content: each learner received a personalized plan with respect to their learning progression.

Our logs analysis tool allowed us to focus on skills development diagnosis on strategies discovery and execution. For example, one strategy related to skill 1 consisted of choosing brokerage contracts closed to each other, in order to be more efficient in managing visits and purchase offers (Strategy 1.1: territory strategy, see Table 1). Our tool provides, for each participant's game session, a map of the city including all locations of the contracts chosen by the learner. Figure 8 shows examples of two maps obtained by participant 33 and participant 16, both in generated version of the game.

These maps show that participant 33 executed the territory strategy, whereas participant 16 chose distant contracts from each other.

Moreover, our tool evaluates strategies associated to "Managing and concluding house sales" (skill 3), which was not estimated by the Bayesian Network (BN). To do so, the tool provides, for each participant's game session:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Chosen situations by the adaptation module for the first three months among learners (N=18)

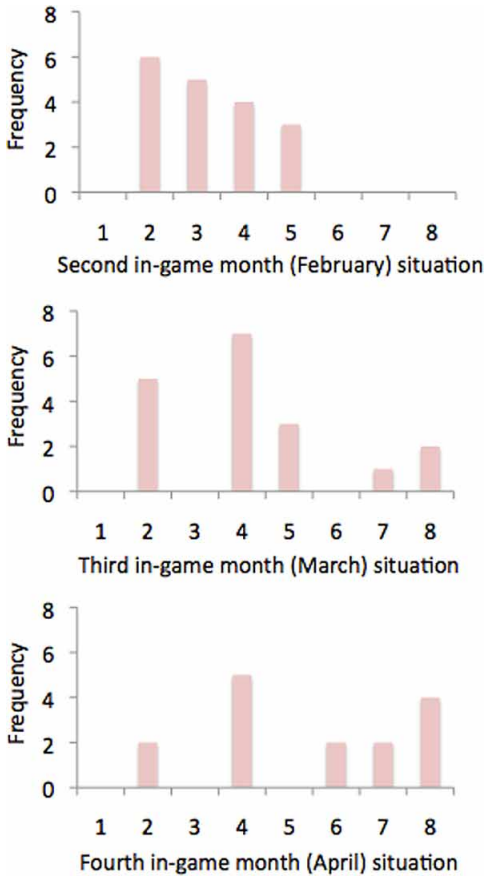
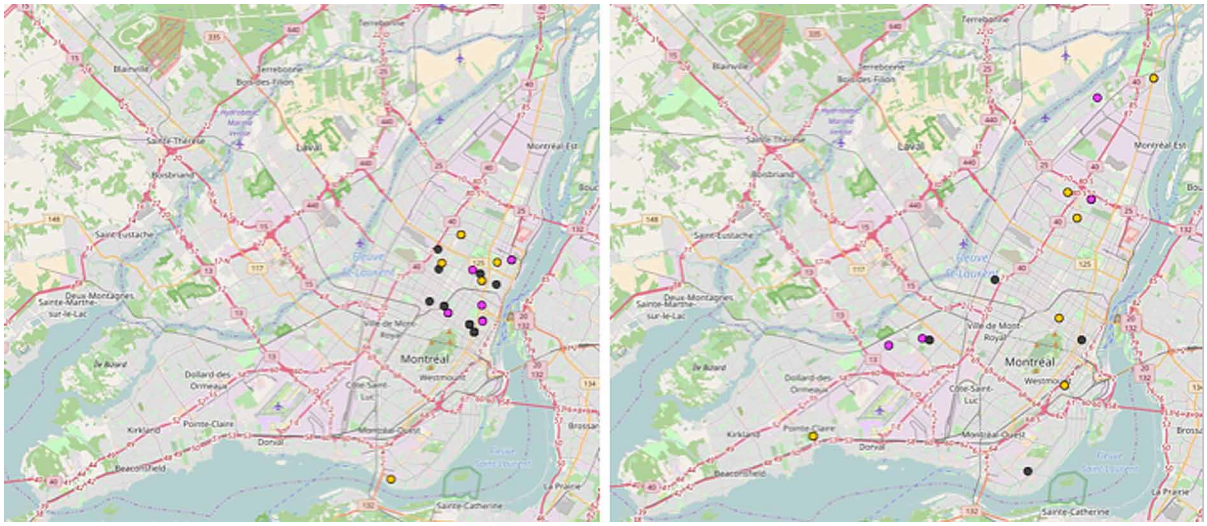


Figure 8. Territories at the end of game session for participant 33 (on the left) and participant 16 (on the right). Black points: contracts selected by the learner, but lost to another NPCs broker; pink points: learner's contracts, not sold at the end of the game session; orange points: learner's contracts, sold.



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Percentage mean of missed visits (when the learner waited too long to go back to her houses in sale in front of which buyers were waiting to visit). For example, participant 33 obtained a rate of missed visits of 45% (Strategy 3.1, see Table 1);
- Percentage mean of house visibility on the market (due to advertising). For example, participant 33 maintained a visibility of 69,7% (Strategy 3.2, see Table 1);
- Percentage mean of missed purchase offers (when the learner waited too long to go back to her houses in sale in front of which buyers were doing a purchase offer). For example, participant 33 obtained a rate of missed purchase offer of 18,75% (Strategy 3.3, see Table 1).

The provided dataset is also essential and complementary to data on skills development as they allow us to verify if taught strategies were discovered and at what rate was used.

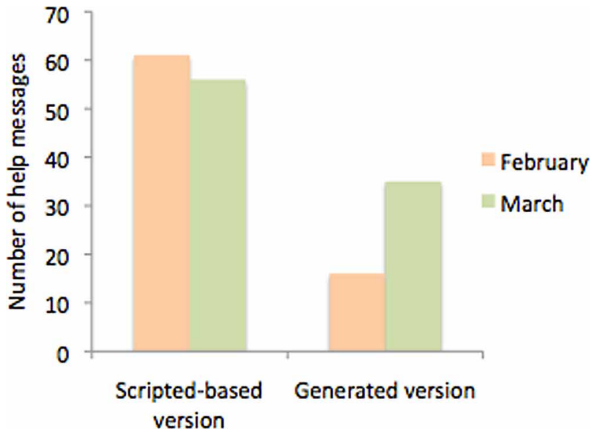
In order to compare scripted-based version and generated version of *Game of Homes*, we show in Figure 9 the number of help messages displayed during February and March in the game, for both versions of the game. We selected 22 participants for these analyses, because they played for at least until March in the game.

Despite the fact that the results are closely dependent on the elapsed time in-game, data show that the adaptation module in our architecture allowed a controlled generation of help messages, whereas the scripted version of the game equally displayed help messages.

Since the elapsed time in-game is not the same for every learner, we chose to present in the section two pedagogical scenarios inscribed in game logs for the two participants that played the longest in-game time. We assert that longer pedagogical scenarios are more complete and thus more significant in terms of progression of learning. Table 4 shows these two scenarios for participant 9 (scripted-based version, played until June 2016 in-game, 175 in-game days) and for participant 16 (generated version, played until May 2016 in game, 139 in-game days). At the beginning of each month inside the game, the logs module recorded the probabilities for skill 1 and skill 2, both provided by the learner model.

Detailed scenarios for both versions show that even six scripted situations proposed to the learner were not enough to learn skill 2 (47% at the end of game session). In the generated version, the adaptation module proposed only two situations to develop both skills, but a third situation was actually adequate

Figure 9. Total number of help messages displayed for both groups (scripted-based version, N=13; generated version, N=9) who had played for at least three months in the game (N=22)



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Examples of two pedagogical scenarios extracted from game logs

Scripted-Based Version (Participant 9)	Automatically Generated Version (Participant 16)
February: P(skill1) = 98% & P(skill2) = 18% Situation 2 of the script is executed. Help messages displayed: strategy concerning market situation, sellers' demand, comparable houses. March: P(skill1) = 98% & P(skill2) = 31% Situation 3 of the script is executed. Help messages displayed: strategy concerning brokers' territory, commission rates, sellers' demand, comparable houses. April: P(skill1) = 99% & P(skill2) = 55% Situation 4 of the script is executed. Help messages displayed: strategy concerning brokers' territory, commission rates, sellers' demand, visits. May: P(skill1) = 99% & P(skill2) = 50% Situation 5 of the script is executed. Help messages are not activated for this action. June: P(skill1) = 99% & P(skill2) = 47% Situation 6 of the script is executed. Help messages are not activated for this action. End of the game session.	February: P(skill1) = 79% & P(skill2) = 43% First generated pedagogical scenario including two situations: Situation 4, Situation 7. Help messages displayed: none. March: P(skill1) = 88% & P(skill2) = 60% Second generated pedagogical scenario including two situations: Situation 4, Situation 8. Help messages displayed: strategy concerning brokers' territory, commission rates, sellers' demand, visits. April: P(skill1) = 99% & P(skill2) = 73% Third generated pedagogical scenario including one situation: Situation 6. Help messages displayed: strategy concerning comparable houses. May: P(skill1) = 99% & P(skill2) = 92% Both skills are acquired. End of the game session.

to learn both skills (99% for skill 1 and 92% for skill 2 at the end of game session). Lastly, we can see in Table 2 that generated scenarios were adapted for each participant according to their learning progression. For example, the first estimated skills for participant 4 were P(skill 1) = 47% and P(skill 2) = 17%, to which the adaptation module proposed an eight-situations scenario, while it proposed a two-situation scenario for participant 16 whose skills were first estimated at P (skill 1) = 79% and P (skill 2) = 43%.

Overall, results show that the automatically generated version of *Game of Homes* is efficient in adapting pedagogical scenarios to the learning progression, inducing faster learning for the participants who played for long periods on in-game time. Moreover, the logs module allows us to precisely trace learning progression, represented by development of skill 1 and skill 2, and makes it possible to detect a potential decrease of skills probabilities.

CONCLUSION

In this study, we addressed the problem of adapting game content to a learner in order to improve the learning process and keep the learner motivated while playing a simulation serious game. Because trying to predict multiple game situations and to propose a content adaptation to each case are time-consuming, laborious, and complicated, we proposed an architecture of simulation serious game based on Bayesian Network and planning algorithm. This architecture can be applied to develop other simulation serious games.

Using automatically generated scenarios instead of scripted-based scenarios in simulation serious game is profitable to learners. Their progression in the game is more pleasant than in a scripted-based game because the learning content is closer to her learning needs. In addition, when replaying the game, it remains interesting for her as the game adapts continuously to her new skill level.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Few studies proposing architectures of serious games assert the importance of recording and analyzing game logs (Gervas, 2014; Harpstead et al., 2013; Sanchez et al., 2015; Westera et al., 2014). Our research demonstrates both the essential and complementary role of the logs module with the other modules (adaptation and control) in the generation of pedagogical scenarios. At the end of a game session, the logs module allows trainers to read the exact content of the pedagogical scenario. The results of our empirical evaluation of *Game of Homes* showed that each generated scenario was different from the scenario used in scripted-based version. Moreover, each scenario adapts itself to the learning progression of the learner while playing. More difficult situations are planned for learners whose skills development is on the right track, while easier situations are proposed to learners experiencing difficulty. This adaptation can occur at any point in the game, even in the case a learner masters a skill at the beginning of the game. This feature also guaranties learner engagement throughout the duration of the game session.

The fact that the pedagogical scenarios are unique to each learner gives an important pedagogical value to generated scenarios compared to scripted-based ones. On one hand, designing in advance the best scenario for each learner is a very difficult task. On the other hand, it is very arduous to predict both learners' learning speed and level of engagement in a serious game experience. Thus, we suggest, as demonstrated in this study, that tools be designed to facilitate data extraction and detect learning patterns. In our case, our tool allows us to verify the discovery and execution of taught strategies in the game.

REFERENCES

Amory, A. (2007). Game object model version II: A theoretical framework for educational game development. *Educational Technology Research and Development*, *55*(1), 51–77. doi:10.100711423-006-9001-x

Arnab, S., Lim, T., Carvalho, M. B., Bellotti, F., Freitas, S., Louchart, S., ... De Gloria, A. (2015). Mapping learning and game mechanics for serious games analysis. *British Journal of Educational Technology*, *46*(2), 391–411. doi:10.1111/bjet.12113

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In *Learning analytics* (pp. 61–75). Springer New York.

Beaudry, É., Bisson, F., Chamberland, S., & Kabanza, F. (2010). Using Markov decision theory to provide a fair challenge in a roll-and-move board game. *Proceedings of the 2010 IEEE Symposium on Computational Intelligence and Games* (pp. 1-8). 10.1109/ITW.2010.5593380

Conati, C., Gertner, A., & Vanlehn, K. (2002). Using Bayesian networks to manage uncertainty in student modeling. *User Modeling and User-Adapted Interaction*, *12*(4), 371–417. doi:10.1023/A:1021258506583

Conati, C., & Zhao, X. (2004). Building and evaluating an intelligent pedagogical agent to improve the effectiveness of an educational game. *Proceedings of the 9th International Conference on Intelligent User Interfaces* (pp. 6-13). 10.1145/964442.964446

Despotović-Zrakić, M., Marković, A., Bogdanović, Z., Barać, D., & Krco, S. (2012). Providing adaptivity in Moodle LMS courses. *Journal of Educational Technology & Society*, *15*(1), 326–338.

Galarneau, L. (2005). Authentic learning experiences through play: Games, simulations and the construction of knowledge. *Digital Games Research Association (DiGRA)*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Garris, R., Ahlers, R., & Driskell, J. E. (2002). Games, motivation, and learning: A research and practice model. *Simulation & Gaming*, *33*(4), 441–467. doi:10.1177/1046878102238607

Gervás, P. (2014). Metrics for desired structural features for narrative renderings of game logs. *Entertainment Computing*, *5*(4), 245–250. doi:10.1016/j.entcom.2014.08.001

Girard, C., Ecalle, J., & Magnan, A. (2013). Serious games as new educational tools: How effective are they? A meta-analysis of recent studies. *Journal of Computer Assisted Learning*, *29*(3), 207–219. doi:10.1111/j.1365-2729.2012.00489.x

Gredler, M. E. (2004). Games and simulations and their relationships to learning. In Handbook of Research on Educational Communications and Technology (Vol. 2, pp. 571-581).

Harpstead, E., Myers, B. A., & Aleven, V. (2013). In search of learning: facilitating data analysis in educational games. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 79-88). 10.1145/2470654.2470667

Kiili, K. (2005). Digital game-based learning: Towards an experiential gaming model. *The Internet and Higher Education*, *8*(1), 13–24. doi:10.1016/j.iheduc.2004.12.001

Law, E. L. C., & Sun, X. (2012). Evaluating user experience of adaptive digital educational games with Activity Theory. *International Journal of Human-Computer Studies*, *70*(7), 478–497. doi:10.1016/j.ijhcs.2012.01.007

Li, B., & Riedl, M. O. (2010). An Offline Planning Approach to Game Plotline Adaptation. *Proceedings of the 6th Annual AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*.

Lopes, R., & Bidarra, R. (2011). Adaptivity challenges in games and simulations: A survey. *IEEE Transactions on Computational Intelligence and AI in Games*, *3*(2), 85–99. doi:10.1109/TCIAIG.2011.2152841

Manske, M., & Conati, C. (2005). Modelling Learning in an Educational Game. In *Artificial Intelligence in Education* (pp. 411–418). AIED.

Millington, I., & Funge, J. (2009). *Artificial intelligence for games*. Morgan Kaufmann.

Niehaus, J., & Riedl, M. (2009). Toward scenario adaptation for learning. *Proceedings of the Conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling* (Vol. 3, pp. 686–688). IOS Press.

Paquette. (2004). *Instructional Engineering for Network-Based Learning.* San Francisco, CA: Pfeiffer/Wiley.

Peirce, N., Conlan, O., & Wade, V. (2008). Adaptive educational games: Providing non-invasive personalised learning experiences. *Proceedings of the Second IEEE International Conference on Digital Games and Intelligent Toys Based Education* (pp. 28-35). 10.1109/DIGITEL.2008.30

Pernin, J.-P., & Lejeune, A. (2004). Dispositifs d'apprentissage instrumentés par les technologies: vers une ingénierie centrée sur les scénarios. In Actes du colloque TICE (pp. 407–414).

Sanchez, E., Valérie, E. M., & Mandran, N. (2015). Jeu-game, jeu-play, vers une modélisation du jeu. Une étude empirique à partir des traces numériques d'interaction du jeu Tamagocours. *Revue STICEF, 22*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sauvé, L., & Kaufman, D. (2010). *Jeux et simulations éducatifs*. Presses de l'Université du Québec.

Schell, J. (2008). *The Art of Game Design: A book of lenses* (1st ed.). Morgan Kaufmann Publishers Inc.

Spiro, R. J., Feltovich, P. J., Coulson, R. L., Jacobson, M., Durgunoglu, A., Ravlin, S., & Jehng, J. C. (1992). *Knowledge Acquisition for Application: Cognitive Flexibility and Transfer of Training in III-Structured Domains*. (Research Note 92-21). Alexandria VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Steiner, C. M., Kickmeier-Rust, M. D., Mattheiss, E., & Albert, D. (2009). Undercover: Non-invasive, adaptive interventions in educational games. *Proceedings of 80 Days' 1st International Open Workshop on Intelligent Personalisation and Adaptation in Digital Educational Games*, 55-65.

Thomas, P., Labat, J. M., Muratet, M., & Yessad, A. (2012). How to evaluate competencies in game-based learning systems automatically? In *Intelligent Tutoring Systems* (pp. 168-173).

Vandewaetere, M., Vandercruysse, S., & Clarebout, G. (2012). Learners perceptions and illusions of adaptivity in computer-based learning environments. *Educational Technology Research and Development*, *60*(2), 307–324. doi:10.100711423-011-9225-2

Westera, W., Nadolski, R., & Hummel, H. (2014). Serious Gaming Analytics: What Students Log Files Tell Us about Gaming and Learning. *International Journal of Serious Games*, *1*(2), 35–50. doi:10.17083/ijsg.v1i2.9

Wouters, P., Van Nimwegen, C., Van Oostendorp, H., & Van Der Spek, E. D. (2013). A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology*, *105*(2), 249–265. doi:10.1037/a0031311

ENDNOTES

1 In this article, we use the term "learner" to the learner and the player, who is the same person in serious games.

2 *Game of Homes* is freely available at http://gdac2.uqam.ca/goh/GameOfHomes-20150812.zip

This research was previously published in the International Journal of Game-Based Learning (IJGBL), 7(2); edited by Patrick Felicia; pages 1-19; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 58
A Study of the State of the Art in Synthetic Emotional Intelligence in Affective Computing

Syeda Erfana Zohora
Taif University, Saudi Arabia

Arvind K. Srivastava
Systron Donner Inertial, USA

A. M. Khan
Mangalore University, India

Nhu Gia Nguyen
Duy Tan University, Vietnam

Nilanjan Dey
TechnoIndia College of Technology, India

ABSTRACT

In the last few decades there has been a tremendous amount of research on synthetic emotional intelligence related to affective computing that has significantly advanced from the technological point of view that refers to academic studies, systematic learning and developing knowledge and affective technology to a extensive area of real life time systems coupled with their applications. The objective of this paper is to present a general idea on the area of emotional intelligence in affective computing. The overview of the state of the art in emotional intelligence comprises of basic definitions and terminology, a study of current technological scenario. The paper also proposes research activities with a detailed study of ethical issues, challenges with importance on affective computing. Lastly, we present a broad area of applications such as interactive learning emotional systems, modeling emotional agents with an intention of employing these agents in human computer interactions as well as in education.

DOI: 10.4018/978-1-7998-0951-7.ch058

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The development of information technology, emotional intelligence in affective computing has contributed significantly towards the evolution of intelligent machine in the current technological scenario. The computer human interaction has constantly increased also. Many researchers in the fields of emotional Intelligence and Human Computer Interaction have highlighted that the agents or intelligent machines of the future must collaborate on an emotional height with their users (Norman, 2004). This is establishing on the concept that an intelligent and intellectual human is not only advanced in verbal, logical reasoning and mathematical, but is also able to attach with other human beings. Much recent work in this area has aimed on authorizing agents with the caliber to both detect emotion based on verbal, non-verbal, and textual thoughts, and also conveys emotion through speech and gesture. The field of Affective Computing (AC) aspires to narrow the communicative gap between the highly emotional human and the emotionally challenged computer by building computational systems that identify and respond to the affective states (e.g., moods and emotions) of the user. Affect-sensitive interfaces are being developed in a number of fields, including mental health, learning technologies and gaming.

What Are Emotions? Emotion theorists have debated for decades about what emotions are and what their primary function in human life is. This debate is far from over and there is currently no universally agreed upon definition of emotions. However, many scholars would at least agree that we experience different types of emotions in our everyday lives.

Emotions are an important aspect of life and they play an crucial role to understand user's behavior with computer interaction. In addition, the emotional intelligence plays a important role to measure aspects of success in life. Recent researches in human-computer interaction don't aim only on the cognitive approach, but on the emotions part too. Both techniques are very important, indeed take into account that emotions of the user solve some important aspects of the design in HCI systems. Additionally, the human-machine interaction could be better if the machine can adapt its behavior according to users; and this system is seen more natural, efficacious, persuasive, and trustworthy by users.

The theory of emotional intelligence became recognized in the late 1980's; on the other hand, the research study (Thorndike, 1920) presented an identical perception called social intelligence in the earlier period, during 1920. While one's social intelligence is classically described by their "ability to recognize and handle other people and to relate in adaptive social interactions" (Kihlstrom, J.F., & Cantor, N. (2000)). Emotional intelligence is a facet of human intelligence that has been argued to be the main important theory for a successful social life (Goleman, 1995), emotional intelligence refers to one's skill to perceive, recognize, manage, and state emotion within oneself and in dealing with others (Salovey & Mayer, 1990). The authors Salovey and Mayer define five domains critical to emotional intelligence: knowing one's emotions, managing emotions, motivating one, recognizing emotions in others, and handling relationships. When we deal with computers, not all of the computers need emotional intelligence and none will need all of the associated skills that we need from them. However, man-machine interactive systems proficient of sensing stress, negligence and thoughtfulness and skilled of adapting and responding suitably to these affective states of the user are probably to be perceived as more natural, more efficacious and more trustworthy. The research area of machine analysis and employment of human affective states to build more natural, flexible interactive systems refers to the introduction of affective computing.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In the late 1990s, Systematic research programs along this front began to materialize in the early 1990s. Many AI and HCI researchers started to take the idea of emotion and emotional intelligence quite seriously. However, in the 1980s, researchers (Turkle, 1984) found to guess about how computers might be employed to study emotions. For example, the researchers (Scherer, 1993) executed a computational model of emotion as an expert system. A few years later, Picard's landmark book Affective Computing (Picard, 1997) encouraged a wave of interest amongst computer scientists and engineers finding for solutions to advance human-computer interfaces by coordinating emotions and cognition with task constraints and demands.

The influence of Emotional Intelligence on Intelligent Systems in modern research is inevitable. The current cognitive psychology believe that the human emotions originates by specific situations. Emotional change is a psychological experience rare to humans, and can generate a sequence of physiological outcomes through the nervous system, and shapes a unique subjective experience. This, consecutively, may cause external expression changes, in form of gestures, actions, and so on. With the growth of artificial intelligence (AI) techniques, many new needs have been put forward in real-life complex situations: the new invention of intelligent machines should have the capability to create personalized responses based on environmental observation.

The Affective Computing Lab within the MIT Media Lab was founded by Rosalind Picard following the publishing of her 1997 book titled "Affective Computing," in which she described the structure for developing machines with emotional intelligence (Picard, 1997). They, along with many other researchers in this field, have developed machines that can detect, handle, understand and express emotions. Before arguing the theories and applications of machines that are emotionally intelligent, it is significant to first understand that emotion is a key aspect of intelligence. The proof that stresses this work comes from a few diverse fields. The author presented a sketch for building machines with emotional intelligence. In the forthcoming years, many other researchers in this field have built machines that can be represented as three types of affective computing applications: 1) systems that detect the emotions of the user, 2) systems that express what a human would perceive as an emotion (e.g., an avatar, robot, and animated conversational agent), and 3) systems that actually "feel" an emotion (Picard, 2001).

The emotion perceiving research has investigated numerous channels that people use to form impressions of each other's emotions – facial expression, paralinguistic, gesture, choice of words and actions. Physiological correlates of affect also put forth a special fascination (Warwick, 2014; Odella, 2016). The Emotional Intelligence on Intelligent Systems in modern research is inevitable. Modern cognitive psychology regarded as human emotions to be caused by specific circumstances. Emotional change is a psychological phenomenon exceptional to humans, and can activate a sequence of physiological responses through the nervous system, and form a distinctive subjective experience. This, in turn, may reason external expression changes, in form of gestures, actions, and so on (Ioannou et al., 2005).

The authors Salovey and Mayer (Mayer, 1990) articulated in the quote that the common notion is that emotion which is an obstruction to intelligent thought. Much work in the field of affective neuroscience has provided experimental evidence that this is not the case, and certainly the opposite is true. Affective Neuroscience is the study of the processing of emotions within the human brain. Researchers in this field have shown that emotion plays a critical role in problem solving and other cognitive tasks within the brain (Damasio, 1994). In addition to the practical evidence that arises from Affective Neuroscience, many publications from popular psychology have supported this view as well, arguing that emotional intelligence is important to a person's success in many aspects of life (Hatch et al., 1993).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The main objective of the paper is to present state of the art in emotional intelligence and agents, and analyze, select, and propose several suitable applications. Such systems should be competent of reasoning about emotions, predicting and understanding human emotions, and processing emotions while relating with a human user. The major contribution of this paper is to expand discussion of research questions and efforts required for building usable emotionally-aware and expressive intelligent systems for diverse domains. Especially, research sequences are recognized and systematically prepared with the idea of building a structure for creating such systems. Motivated by progress building an Affective Learning Companion, we have newly started considering a project to build an Assessment Companion for Emotional Intelligence. This system would aim the reviewing of how people perform in circumstances that demand emotional intelligence, vs. testing what they cognitively conclude about such circumstances as in current tests of emotional intelligence such as the MSCEIT (Salovey et al., 2003).

The paper is arranged as follows. In Section 2, the concept of emotional intelligence is briefly presented and also highlights extensive overview of research of emotional intelligence. The Section 3 explains about the issues and challenges related to emotional intelligence are proposed. Concluding applications are briefly presented in Section 4.

SYNTHETIC EMOTIONAL INTELLIGENCE IN AFFECTIVE COMPUTING

The growth and research of the emotional intelligence in the last recent decades has focused on empowering agents with the ability to detect emotions based on verbal, non-verbal, textual concepts and also to express emotions through speech.

The Concept of Emotional Intelligence

The concept of emotional intelligence is based on the notion of co-ordination between the intelligent machines and human beings. The notion referring to the emotional intelligence and the subject of emotional intelligence has been described controversially. The emotional intelligence is explained as a study which appears for human cognitive abilities beyond conventional educational intelligence (Matthews, 2004). Researchers have classified their clarification based on either an ability model or mixed model. Based on the (Pherwani, 2003) ability model, "emotional intelligence is defined as a set of abilities that involves perceiving and reasoning abstractly with information that emerges from feelings" Studies of (Salovey, 1997) have supported this model stating that emotional intelligence is defined as "the ability to perceive accurately, appraise, and express emotion; the ability to access and/or generate feelings when they facilitate thought; the ability to understand emotion and emotional knowledge; and the ability to regulate emotions to promote emotional and intellectual growth". In addition, (Caruso, 2000a) gave the definition of the emotional intelligence as "the capacity to reason about emotions to enhance thinking. It includes the ability to accurately perceive emotions, to access and generate emotions so as to assist thought, to understand emotions and emotional knowledge, and to reflectively regulate emotions so as to promote emotional and intellectual growth". Conversely, the researcher's (Goleman, 1995) gave another definition to the emotional intelligence "as the ability to recognize and regulate emotions both within the self and within others" and also as "An array of noncognitive capabilities, competencies, and skills that

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

influence one's ability to succeed in coping with environmental demands and pressures" (Bar-On, 2004). In conclusion, the mixed-model is related to the ability as it deals with the social behaviors, features and characteristics. In the following section we define the theories and models of emotional intelligence.

Theories and Models of Emotional Intelligence

The researchers in the area of emotional intelligence have categorized the models based on either an ability model or mixed model.

Ability Model of Emotional Intelligence

Ability models explain emotional intelligence to the customary intelligence equally (cognitive intelligence). The researchers in 1990 (Mayer, 1990) have recognized and developed the ability model by presuming the emotional intelligence which expands overtime, has correlation with measures of IQ, and can be measured during a performance-based test (Ciarrochi, 2005). Proponents of the ability model typically explain the emotional intelligence as "the ability to perceive and express emotion accurately and adaptively, the skill to comprehend emotion and emotional knowledge, and the capability to use feeling to make possible thought, and the ability to control emotions in oneself and in others" (Pizarro, 2003; Caruso, 2004) also give this explanation to the Emotional Intelligence as one among other "hot intelligences" including those that refers to social, pragmatic and subjective intelligences. These intelligences occupies several names as they are considered to contain cognitive processes encompassing necessity personally and emotionally (Caruso, 2000a; Sawaf, 1997) choose the ability model of EI involving perception and the ability to make it rational in an conceptual field from awareness of one's feelings. Though, this meticulous model does not have any claim concerning the special ability of EI and its value of prediction. They consider that the ability model points on the existence of EI. "If emotional intelligence exists and is eligible as a traditional or standard intelligence (like general IQ), humans who are referred bleeding hearts or hopeless romantics may be occupied in complicated information processing". In addition, emotional intelligence is conceptualized in a mode that it is righteous in the school and organizational environments where emotions are imitated as relationships. The researchers (Cobb, 2000) refers to this model and state that "emotional intelligence is predictive of life's success or that it leads to good behavior" but, organizational policies about emotional intelligence have more development than the scientific basis in a practical analysis.

Mixed Model of Emotional Intelligence

Explaining to the second approach of EI, the "mixed model", it is described as the vision of emotional intelligence socio-economically (Caruso, 2000a). This outlook includes capability in addition to a pattern of features, skins and vital personality mechanism. This group comprises the articles acknowledged widely and a host of inputs in the study (Mayer, 1990; Bar-On, 1996). The authors argue that the abilities of the emotional intelligence acquires few skills in the affective area, as well as skills in whatever cognitive elements have a position in each ability. They initiate with emotional abilities and then combine them with personality, motivation, and affective dispositions, such as a requirement for achievements, assertiveness, self-esteem, impulse control, happiness, and emotional awareness, and empathy. These models have good standardized, reliable and factorial valid scales. Their main evaluation of emotional

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

intelligence is probable to be self-report inventories (Barsade, 2008). Emotional intelligence is a mixture of skills together with emotional realization and characteristics such as perseverance, commitment, and pleasing behavior in the mixed model (Bar-On, 1996; Ryback, 1998; Marijuán, 2014). Intelligence quotient and cognition are not stressed in the mixed model (Cobb, 2000), but the study carried out (Goleman, 1998b) claimed that emotional intelligence has more strength than IQ twice and it can be measured as an indication of future success and attainments in life.

The State of the Art and Applications

In this central section of the paper we will present most distinctive research activities that results in the field of emotional intelligence and agents in about the last 20 years. Since the early works (Nakatsu, 1996) a lot of different and interesting approaches have been studied. The author's deal with the issues of communication and esthetics of artificial life that possess this "human form" in modern society. In the current example emotions are interpreted from human voices and emotional responses. It comprises a comic character and a beautiful, sensual goddess. The same authors several years later presented a new system. They applied their emotion recognition algorithm to a computer agent that plays a character role in the interactive movie system. First practical systems in the area i.e. the pioneering architectures has been presented in (Coglio, 1998). The authors planned a general, flexible, and powerful architecture to build software agents that embed artificial emotions for conveying human emotions, thus allowing more effective, stimulating, and natural human-agent interactions. Within the emotional intelligent agent domain three directions for emotion detection are: facial, vocal and text. Facial recognition elevates the most composite set of problems, but some competent solutions do exist. For example, a support vector machine can be used to classify motion signatures as either non-expressive, or belonging to one of the predefined six categories (McOwan, 2006). Alternatively, paper (Kollias, 2005) contains adaptation and learning in a neuro-fuzzy system in order to analyze facial animation parameters. The processing of vocal emotions usually consists of three stages (Stegmann, 2009) feature extraction, data classification, and post-processing. Data classification and pattern matching represent the central point, and different approaches have been evaluated. These include k-NN and linear discriminant classifiers (Narayanan, 2005), artificial neural networks (Lopez-Cozar, 2008), support vector machines (Mohri, 2005) and hidden Markov models (Toschev, 2015). More recently, developments to generative classifiers, such as Gaussian mixture models, have been proposed in (Toschev, 2013). There are a number of resourceful methods for text emotion detection. The approach proposed in (Blecha, 2016) is aimed at improving text-to-speech systems for children's fairy tales by identifying emotions. Sentences can also be represented as series of semantic labels and attributes and then processed by a separable mixture model in order to compute resemblances to emotion generation rules (Lin, 2006). Finally, in (Mihalcea, 2008) several algorithms have been employed in order to classify news labels based on their emotional content. Finally, there exist some collective approaches. It has been presented (Wu, 2004) that the combination of vocal and text recognition outcomes better results than when only a single approach is used. The authors of (Taylor, 2005) have proposed a neural network-based approach to analyze facial/vocal input. Intelligent software agents can be applied to and use both vocal and facial emotion recognition. Agents are dispersed by nature, and can therefore be employed to optimize selection, classification, and pattern recognition (Ivanović, 2012). A special pedagogical agent is used in e-learning systems to guide and motivate students (Vidaković, 2011) (Clarebout, 2011) Much research has also been invested for incorporating

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

emotions into agents, resulting in a number of formal models for emotional agents (John-Jules, 2012), (Schwarzentruber, 2011). Finally, some challenges of developing successful embodied conversational agents are outlined in (Asendorpf, 2011).

ISSUES AND CHALLENGES RELATED TO EMOTIONAL INTELLIGENCE IN AFFECTIVE COMPUTING

In the following section, we discuss about the issues and challenges emotional intelligence in Affective Computing.

Detection of Affective Computing

The expansion of an affect-aware system that senses and responds to an individual's affective states usually requires the system to primary detect affect. Affect detection is an exceptionally challenging attempt due to the several complications connected with familiarity and expression of affect. The concept of emotion research (or affective science) and affective computing are very different, perhaps the former being the human and the latter being the computer (although successful AC research should view the human and computer as one interacting system). In addition, in the fields of regular concerns (e.g., affect expression and detection by humans and computers), disciplinary disparity is still to be surpassed before a more coherent body of literature can be developed. In this section, we discuss the affective computing methods to affect detection. The interdisciplinary consideration between the emotion theorists and the AC practitioners, we connect each method to the emotion theory that most closely supports it, in many cases; there is no clear one-to-one mapping between affective computing method and emotion theory. The review of current affect detection systems is planned with respect to individual modalities or channels (e.g., face, voice, and text). Each modality has advantages and disadvantages toward its use as a viable affect detection channel. Some of the factors affecting the value of each modality comprises of:

1. The validity of the signal as a normal way to recognize an affective state,
2. The reliability of the signals in real-world situations,
3. The time resolution of the signal as it relates to the specific requirements of the application, and
4. The cost and intrusiveness for the user.

Each modality has its own body of literature, and extensively reviewing this literature would be beyond the scope of this survey. Hence, we presented an overview of the research in each modality and focus on the major accomplishments and open issues. We also present a synthesis of existing reviews if they are available. The majority of affect detection research has focused on detecting the basic emotions from the face. This view point's has a distinctive facial expression associated with each basic emotion (Friesen, 2003; Ekman, 1992). So detecting an emotion is basically a topic of detecting its prototypical facial expression. The classified expressions are then associated to the six basic emotions: anger, disgust, fear, joy, sadness, and surprise (Ekman, 1984). The growth of fully automated face-based affect detection. Peter Robinson and his team have worked on the use of facial expressions for affect detection (Dey et al., 2016; Robinson, 2004). Other projects aim to use facial expression recognition techniques to improve the interaction between students and learning systems (Arroyo, 2009; McDaniel, 2007). In

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

the recent survey (Wang, 2016) additional technological development is essential before vision-based affect detection can be functional for real-world applications. Nevertheless, when compared to a former review in the research study (Pantic, 2003), outstanding progress has been done along a number of dimensions. The most significant development has been made toward larger and more naturalistic emotion databases. Many research studies reviewed six video, seven audio, and five audiovisual databases of affective behavior has been carried out.

Human computer Interaction

The concept of human-computer interaction was automatically characterized with the introduction of computer machines which in turn became basis to the development and design of HCI (Te'eni, 2007) The HCI usability refers to how a machine can be employed adequately to perform few jobs for the users. The HCI functionality explains a set of services that are provided to the users (Shneiderman, 2004), From these terms the actual effectiveness of a system is attained when there is compensation between these two functions (Nielsen, 1994).

In human-computer interaction, the nonverbal communication plays an important role, in that we can recognize the difficulties that users stumble upon, by measuring the emotions in facial expressions (Pantic, 2003). Therefore, we discuss about a number of studies that have examined people's reactions and responses to computers that have been designed to be more human-like. And several studies have reported a positive impact of computers that were designed to flatter and praise users when they did something right. With this system, users have a better estimation of themselves (Rogers, 2011). At the same time, we have to be cautious because sometimes, the same expression of an emotion can have a different signification in different countries and cultures. A smile in Europe is the sign of happiness, pleasure or irony. But for Japanese people, it could simply imply their agreement with the applied punishment or could be the sign of indignation associated with the person applying the punishment (Guesgen, 2016). These ethnic differences should make us attentive of different outcomes with the use of the same affective system in different countries.

Machines with Emotional Intelligence

The area of emotional intelligence has projected that the intelligent machines must relate on an emotional basis with the users (Norman, 2004). With an established need for work in this area, we present the ideas in which researchers are making efforts towards building emotionally intelligent machines. There are various models/definitions of emotional intelligence, but they all are referred to the ability to connect with others by detecting, expressing, managing and understanding the emotions of oneself and others. Efforts in building machines that are emotionally intelligent focuses around a few key efforts, empowering the machine to detect emotion, enabling the machine to express emotion, and finally, embodying the machine in a virtual or physical way (Croucher, 1981). Projects that incorporate all of these aspects also need the additional ability to handle and maintain an emotional interaction with a user, a large added complexity that are intended to expose students to work in the space of emotional intelligence (Jain, 2016).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

APPLICATION OF EMOTIONAL INTELLIGENCE IN COMPUTING ENVIRONMENT

In the rest of the section we will present research activities in the area of emotional intelligence and artificial agents divided into several fields such as, modeling, learning, robotics, decision support, interactive environments, and classification and search.

As previously discussed, the human interaction includes emotional information of the interlocutors that is transmitted by the explicit channel through the language and by the implicit channel by the nonverbal communication. Nevertheless, it has been evaluated that these concepts that are also connected to the interpersonal relations emerge in the communication with the computers (Picard, 1997). Therefore, the main objective of the affective computation is to detect and process the emotional information with the reason of improving the communication between the person and the computer-based system. Distance learning, robotic and psychotherapy are some of the areas that could help from the affective computer. Furthermore, the Affective Computing research group of the Massachusetts Institute of Technology (MIT) has started up an investigation, with the objective to help in the field of the distance learning or also called "e-learning", and to determine the emotional needs of the students (Picard, 1997). Affective Computing also is useful in the field of Telemedicine. According to (Finkelstein, 2003), Tele-Home-Health-Care (tele-HHC) provides communication between medical professionals and patients in cases where hands-on care is not required, but regular supervising is needed. For example, tele-HHC involvements are currently employed to gather vital sign data remotely (e.g., ECG, blood pressure, oxygen saturation, heart rates, and respiration), verify compliance with medicine and/or diet regimes, and assess aspects of mental or emotional status. Another application is "Driving Safety" (Finkelstein, 2003). Taking into consideration that the inability to manage one's emotions while driving is identified as one of the major reasons for accidents, when the system identifies the driver is in a state of anger, for example, the system could change the music (James, 2000), or suggest a relaxation technique (Rodriguez, 1999). depending on the driver's preferred style. The study (Picard, 1997) proposes a help for the autistic people. This problem affects around 1.5 and 2 people in each 1,000. The autism is a developmental disorder which continues during all the life. This disease affects in different degrees the ability as far as the communication to the understanding of the language, the social coexistence and the capacity of the imagination. Some autistic people are incredible memorizing and can learn to recognize facial expressions, but the majority has difficulties with the compression of the emotions, and hence they lack emotional intelligence. According to (Picard, 2001), a way to aid these people is to rehearse with them different consequences repeatedly and this way, they will be able to learn to respond of natural form. The plan is to design computers which had the capacity to teach to these people different social scenes. In the article (Schiano, 2000) collects a listing of different applications from automatic recognition of facial expression. For example, user training, distances learning/tele-teaching assistant, automobile driver alertness/drowsiness monitor, stress detector, entertainment and computer games, health and family planning, etc. All these applications offer several affective skills, amid them the ability to recognize, to express and to have emotions. Some of the applications are simple engineering problems, but others are more complicated, because they depend on existing affective abilities in the human beings.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

Emotional Intelligence plays an essential role in the future generation of man-machine interactive systems which is growing in the field of artificial intelligence. In upcoming research efforts it is crucial to explore various techniques and methods for modeling of emotions of the interactive systems. This refers to the face expression the humans and emotional expressions of the computers can be realized on the display unit depicting the machine response. In this research study carried out explains the concept of emotional intelligence, models and state of the art and mainly issues, the different machine learning and different synthetic emotional intelligence approaches and applications have studied. Lastly, the challenge today is to improve technologies in order to identify emotions despite features that outcomes result.

REFERENCES

Anderson, K., & McOwan, P. W. (2006). A real-time automated system for the recognition of human facial expressions. *IEEE Transactions on Systems, Man, and Cybernetics. Part B, Cybernetics*, *36*(1), 96–105. doi:10.1109/TSMCB.2005.854502 PMID:16468569

Arroyo, I., Cooper, D. G., Burleson, W., Woolf, B. P., Muldner, K., & Christopherson, R. (2009). Emotion Sensors Go to School. *Proc. 14th Conf. Artificial Intelligence in Education* (pp. 17-24).

Attabi, Y., & Dumouchel, P. (2013). Anchor models for emotion recognition from speech. *IEEE Transactions on Affective Computing*, *4*(3), 280–290. doi:10.1109/T-AFFC.2013.17

Bar -On. R. (2004). The Bar-On Emotional Quotient Inventory (EQ-i): Rationale, description, and summary of psychometric properties. In G. Geher (Ed.), Measuring emotional intelligence: Common ground and controversy (pp. 111-142). Hauppauge, NY: Nova Science Publishers.

Bar-On, R. (1996). *The Emotional Quotient Inventory (EQ-I): A test of emotional intelligence*. Toronto: Multi-Health Systems.

Bureš, V., Tučník, P., Mikulecký, P., Mls, K., & Blecha, P. (2016). Application of Ambient Intelligence in Educational Institutions: Visions and Architectures. *International Journal of Ambient Computing and Intelligence*, *7*(2), 94–120. doi:10.4018/IJACI.2016010105

Burkhardt, F., van Ballegooy, M., Engelbrecht, K.-P., Polzehl, T., & Stegmann, J. (2009). Emotion detection in dialog systems: Applications, strategies and challenges. *Proceedings of the 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops* (pp. 1-6). 10.1109/ACII.2009.5349498

Camurri, A., & Coglio, A. (1998). An architecture for emotional agents. *IEEE MultiMedia*, *5*(4), 24–33. doi:10.1109/93.735866

Chuang, Z. J., & Wu, C. H. (2004). Multi-modal emotion recognition from speech and text. *Computational Linguistics and Chinese Language Processing*, *9*(2), 45–62.

Cooper, R. K., & Sawaf, A. (1997). *Executive EQ: Emotional intelligence in leadership and organizations*. New York: Berkley Publishing Group.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Damasio, A. (1994). *Descartes' error: emotions, reason, and the human brain*. New York: Avon Books.

del Moral, R., Navarro, J., & Lahoz-Beltra, R. (2014). Cognitive and Emotional Contents of Laughter: Framing a New Neurocomputational Approach. *International Journal of Synthetic Emotions*, *5*(2), 31–54. doi:10.4018/ijse.2014070104

Dey, A., Bhattacha, D. K., Tibarewala, D. N., Dey, N., Ashour, A. S., Le, D.-N., . . . Gospodinov, M. (2016). Chinese-chi and Kundalini yoga Meditations Effects on the Autonomic Nervous System. *International Journal of Interactive Multimedia and Artificial Intelligence,* 3(7), 87-95.

Ekman, P. (1984). In K. Scherer & P. Ekman (Eds.), *Expression and the Nature of Emotion, Approaches to Emotion*. Erlbaum.

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, *6*(3), 169–200. doi:10.1080/02699939208411068

Ekman, P., & Friesen, W. V. (2003). *Unmasking the Face*. Malor Books.

El Kaliouby, R., & Robinson, P. (2004). Real-Time Inference of Complex Mental States from Facial Expressions and Head Gestures. *Proc. Int'l Conf. Computer Vision and Pattern Recognition* (Vol. 3, p. 154). 10.1109/CVPR.2004.427

Fragopanagos, N., & Taylor, J. G. (2005). Emotion recognition in human–computer interaction. *Neural Networks*, *18*(4), 389–405. doi:10.1016/j.neunet.2005.03.006 PMID:15921887

Goleman, D. (1995). *Emotional Intelligence*. New York: Bantam Books.

Goleman, D. (1998b). *Working with emotional intelligence*. New York: Bantam Books.

Hans, W. (2016). Guesgen, & Stephen Marsland, Using Contextual Information for Recognising Human Behaviour. *International Journal of Ambient Computing and Intelligence*, *7*(1), 27–44. doi:10.4018/IJACI.2016010102

Hatch, T., & Gardner, H. (1993). Finding cognition in the classroom: an expanded view of human intelligence. In G. Salomon (Ed.), *Distributed Cognitions. Psychological and educational considerations*. Cambridge: Cambridge University Press.

Heidig, S., & Clarebout, G. (2011). Do pedagogical agents make a difference to student motivation and learning? Educational Research Review, 6, 27-54. doi:10.1016/j.edurev.2010.07.004

Ioannou, S., Raouzaiou, A., Tzouvaras, V., Maillis, T., Karpouzis, K., & Kollias, S. (2005). Emotion recognition through facial expression analysis based on a neurofuzzy network. *Neural Networks*, *18*(4), 423–435. doi:10.1016/j.neunet.2005.03.004 PMID:15963691

Ivanović, M., Mitrović, D., Budimac, Z., & Vidaković, M. (2011). Metadata harvesting learning resources - an agent-oriented approach. *Proceedings of the 15th International Conference on System Theory, Control and Computing* (pp. 306-311).

Jain, S., Asawa, K., & Computer, E. (2015). Design Challenges and Opportunities. *International Journal of Synthetic Emotions*, *6*(2), 35–56. doi:10.4018/IJSE.2015070103

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

James, L. (2000). *Road Rage and Aggressive Driving*. Amherst, NY: Prometheus Books.

Kihlstrom, J. F., & Cantor, N. (2000). Social intelligence. In R.J. Sternberg (Ed.), Handbook of intelligence (2nd ed., pp. 359-379). Cambridge, U.K.: Cambridge University Press. doi:10.1017/CBO9780511807947.017

Larson, J., & Rodriguez, C. (1999). *Road Rage to Road-Wise*. New York, NY: Tom Doherty Associates, Inc.

Lee, C., & Narayanan, S. (2005). Towards detecting emotions in spoken dialogs. *IEEE Transactions on Speech and Audio Processing*, *13*(2), 293–302. doi:10.1109/TSA.2004.838534

Lisetti, C. L., & Schiano, D. J. (2000). Automatic Facial Expression Interpretation: Where Human-Computer Interaction. Artificial Intelligence and Cognitive Science Intersect. Pragmatics and Cognition, 8(1), 185-235.

Lopez-Cozar, R. (2008). Influence of contextual information in emotion annotation for spoken dialogue systems. *Speech Communication, 50(* 5), 416-433. doi:10.1016/j.specom.2008.01.001

Lorini, E., & Schwarzentruber, F. (2011). A logic for reasoning about counterfactual emotions. *Artificial Intelligence*, *175*(3-4), 814–847. doi:10.1016/j.artint.2010.11.022

Mandell, B., & Pherwani, S. (2003). Relationship between emotional intelligence and transformational leadership style: A gender comparison. *Journal of Business and Psychology*, *17*(3), 387–404. doi:10.1023/A:1022816409059

Mayer, J. D., & Cobb, C. D. (2000). Educational policy on emotional intelligence: Does it make sense? *Educational Psychology Review*, *12*(2), 163–183. doi:10.1023/A:1009093231445

Mayer, J. D., & Salovey, P. (1997). What is emotional Intelligence? In P. Salovey & D. Sluyter (Eds.), *Emotional Intelligence and Emotional developments: Educ. Implications* (pp. 3–31). New York: Basic Books.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2000a). Models of emotional intelligence. In R. Sternberg (Ed.), *Handbook of intelligence*. Cambridge, UK: Cambridge University Press.

Mayer, J. D., Salovey, P., & Caruso, D. R. (2004). Emotional intelligence: Theory, findings, and implications. *Psychological Inquiry*, *60*(3), 197–215. doi:10.120715327965pli1503_02

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring Emotional Intelligence with the MSCEIT V2.0. *Emotion (Washington, D.C.)*, *3*(1), 97–105. doi:10.1037/1528-3542.3.1.97 PMID:12899321

Mayer, R., Roberts, R. D., & Barsade, S. G. (2008). Human Abilities: Emotional Intelligence. *Annual Review of Psychology*, *58*(1), 507–536. doi:10.1146/annurev.psych.59.103006.093646 PMID:17937602

McDaniel, B., D'Mello, S., King, B., Chipman, P., Tapp, K., & Graesser, A. (2007). Facial Features for Affective State Detection in Learning Environments. *Proc. 29th Ann. Meeting of the Cognitive Science Soc.*

Mitrović, D., Geler, Z., & Ivanović, M. (2012). Distributed distance matrix generator based on agents. *In Proceedings of the 2nd International Conference WIMS*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Nasoz, F., Alvarez, K., Lisetti, C. L., & Finkelstein, N. (2003). *Emotion Recognition from Physiological Signals for Presence Technologies*. International Journal of Cognition, Technology, and Work, 6(1), 4-14.

Nielsen, J. (1994). *Usability Engineering.* San Francisco, CA, USA: Morgan Kaufman.

Norman, D. A. (2004). *Emotional Design: Why We Love (or Hate) Everyday Things*. New York: Basic Books.

Odella, F. (2016). Technology Studies and the Sociological Debate on Monitoring of Social Interactions. *International Journal of Ambient Computing and Intelligence*, *7*(1), 1–26. doi:10.4018/IJACI.2016010101

Pantic, M., & Rothkrantz, L. (2003, September). Toward an Affect-Sensitive Multimodal Human-Computer Interaction. *Proceedings of the IEEE*, *91*(9), 1370–1390. doi:10.1109/JPROC.2003.817122

Picard, R. W. (1997). *Affective Computing*. The MIT Press. doi:10.1037/e526112012-054

Picard, R. W., Vyzas, E., & Healey, J. (2001, October). Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*(10), 1175–1191. doi:10.1109/34.954607

Rogers, Y., Sharp, H., & Preece, J. (2011). *Interaction Design: beyond human-computer interaction.* Wiley.

Rosete, D., & Ciarrochi, J. (2005). Emotional intelligence and its relationship to workplace performance outcomes of leadership effectiveness. *Leadership and Organization Development Journal*, *26*(5/6), 388–400. Retrieved from http://www.emeraldinsight.com/0143-7739.htm doi:10.1108/01437730510607871

Ryback, D. (1998). *Putting emotional intelligence to work*. Boston: Butterworth Heinemann.

Salovey, P., & Mayer, J. D. (1990). *Emotional* Intelligence. *Imagination, Cognition and Personality*, *9*(3), 185–211. doi:10.2190/DUGG-P24E-52WK-6CDG

Salovey, P., & Pizarro, D. A. (2003). The value of emotional intelligence. In R.J Sternberg, J. Lautrey & T.I. Lubart (Eds), Models of intelligence: International Perspective (pp. 263-278). American Psychological Association.

Scherer, K. R. (1993, May). An Expert System Approach. *Cognition and Emotion*, *7*(3-4), 325–355. doi:10.1080/02699939308409192

Scherer, K.R. (n. d.). Studying the Emotion-Antecedent Appraisal Process.

Schonbrodt, F. D., & Asendorpf, J. (2011). The challenge of constructing psychologically believable agents. *Journal of Media Psychology: Theories, Methods, and Applications*, *23*(2), 100–107. doi:10.1027/1864-1105/a000040

Shafran, I., & Mohri, M. (2005). A comparison of classifiers for detecting emotion from speech. *Proceedings of the IEEE International Conference on Acoustic Signal and Speech Processing (ICASSP)* (pp. 341-344). 10.1109/ICASSP.2005.1415120

Shneiderman, B., & Plaisant, C. (2004). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (4th ed.). Boston: Pearson/Addison-Wesley.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sloman, A., & Croucher, M. (1981). Why robots will have emotions. *Proc 7th int. joint conference on AI,* Vancouver, Canada (pp. 197–202).

Steunebrink, B. R., Dastani, M., & Meyer, J.-J. C. (2012). A formal model of emotion triggers: An approach for BDI agents. *Synthese*, *185*(1), 83–129. doi:10.100711229-011-0004-8

Strapparava, C., & Mihalcea, R. (2008). Learning to identify emotions in text. *Proceedings of the 2008 ACM Symposium on Applied Computing* (pp. 1556-1560). 10.1145/1363686.1364052

Talanov, M., & Toschev, A. (2015). Appraisal, Coping and High Level Emotions Aspects of Computational Emotional Thinking. *International Journal of Synthetic Emotions*, *6*(1), 24–39. doi:10.4018/IJSE.2015010102

Te'eni, D., Carey, J., & Zhang, P. (2007). *Human Computer Interaction: Developing Effective Organizational Information Systems*. Hoboken: John Wiley & Sons.

Thorndike, E. L. (1920). Intelligence and its use. *Harper's Magazine*, 140, 227-235.

Tosa, N., & Nakatsu, R. (1996). Life-Like Communication Agent - Emotion-Sensing Character "MIC" and Feeling Session Character "MUSE". *Proceedings of the IEEE International Conference on Multimedia Computing and Systems ICMCS* (pp. 12-19). 10.1109/MMCS.1996.534948

Turkle, S. (1984). *The Second Self: Computers and the Human Spirit.* Simon & Schuster.

Wang, D., He, T., Li, Q., Cao, L., Li, Z., Dey, N., … Xu, J. (2016). Image Features based Affective Retrieval employing Improved Parameter and Structure Identification of Adaptive Neuro Fuzzy Inference System. *Neural Computing and Applications.* Doi:10.100700521-016-2512-4

Warwick, K., & Shah, H. (2014). Outwitted by the Hidden: Unsure Emotions. *International Journal of Synthetic Emotions*, *5*(1), 46–49. doi:10.4018/ijse.2014010106

Wu, C. H., Chuang, Z. J., & Lin, Y. C. (2006). Emotion recognition from text using semantic labels and separable mixture models. *ACM Transactions on Asian Language Information Processing*, *5*(2), 165–183. doi:10.1145/1165255.1165259

Zeidner, M., Roberts, R. D., & Matthews, G. (2004). The emotional intelligence bandwagon: Too fast to live, too young to die? *Psychological Inquiry*, *15*(3), 239–248. doi:10.120715327965pli1503_04

This research was previously published in the International Journal of Synthetic Emotions (IJSE), 7(1); edited by Amira S. Ashour and Nilanjan Dey; pages 1-12; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Section 5

Organizational and Social Implications

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 59
Eventuality of an Apartheid State of Things:
An Ethical Perspective on the Internet of Things

Sahil Sholla
National Institute of Technology, India

Roohie Naaz Mir
National Institute of Technology, India

Mohammad Ahsan Chishti
National Institute of Technology, India

ABSTRACT

Notwithstanding the potential of IoT to revolutionise our personal and social lives, the absence of a solid framework of ethics may lead to situations where smart devices are used in ways uncongenial to the moral fabric of a society. In this work, the authors seek to provide a conceptual framework toward incorporating ethics in IoT. They employ the concept of object for each smart device in order to represent ethics relevant to its context. Moreover, the authors propose dedicating a separate ethics layer in the protocol stack of smart devices to account for socio-cultural ethical aspects of a society. The ethics layer enables us to account for ethical responsibilities of smart devices vis-a-vis society so that inadvertent physical, emotional or psychological harm to human beings is avoided. Such mechanism ensures that devices operate ethically not only at individual level but also at D2D level to give rise to high order ethical structures e.g. ethical home, ethical office, ethical university, ethical city, etc.

DOI: 10.4018/978-1-7998-0951-7.ch059

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Modern technologies are often heralded to usher in a new era of human prosperity, but to what extent these assertions really hold is debatable. No doubt, technology has been really shaping our lives: the way we interact, conduct businesses, exchange information, get entertained etc. and the recent wave in this technological streak is the Internet of Things (IoT) - a pervasive technology that seeks to connect virtually all things on the planet to global Internet. Though the idea of 'smartening' ordinary things for greater good of humanity has been romanticised for quite some time, design of a framework for implementing ethics into smart devices of the IoT has not been duly considered. Such technological enhancement would ensure that artificially intelligent pervasive technologies do not inadvertently jeopardise essential human values.

IoT has been hailed as the technology with far reaching consequences, poised to transform the way we perceive and interact with the physical world. IoT seeks to create a virtual world that contains a counterpart for everything in the physical world in order to provide an advanced class of services to people. IoT offers immense market opportunities particularly for manufacturers, ISPs and application developers. The IoT based smart devices are expected to reach 28 billion by the end of 2021("Ericsson Mobility Report," 2015). IoT is anticipated to have unprecedented economic impact estimated to be in the range of $2.7 trillion to $6.2 trillion by 2025 with IoT healthcare having a share of about $1.1–$2.5 trillion (Manyika et al., 2013). Other market opportunities include such application areas as intelligent transportation systems, industry, agriculture, building automation systems, smart grids, smart cities etc. Due to this potential, IoT is receiving considerable attention from academia, research organisations, enterprises and governments.

In this work, the authors review various IoT applications from an ethical perspective to illustrate situations where 'things' go awry and lead to racism, bias and discrimination. The need for designing value laden smart things to safeguard common human interests is expressed, however no commitments to any particular ethical theory are made. Also, to break away from such tendencies, the authors suggest dedicating a separate ethics layer in the protocol stack of smart things and contribute implications for the design of the proposed ethics layer.

IS THE INTERNET OF THINGS RACIST?

In the IoT world, sensors embedded all around us could gather colossal individual, organizational and government data. However, mere collection of data by ubiquitous sensors and subsequent processing represents oversimplification of the IoT vision because not all trifling details that could be sensed are relevant to the context of an application. Although, smart services that IoT seeks to provide would necessitate collection of certain data, unbridled surveillance over public and private lives of people raises serious security and ethical concerns (Michael, Michael, & Perakslis, 2014). Such apprehensions assume greater importance due to increased use of artificial intelligence (AI) embedded in smart devices, as Rosalind Picard, director of the Affective Computing Group at MIT articulated, 'The greater the freedom of a machine, the more it will need moral standards' (Picard, 2000). Also, people must know and be able to manage what type of data are collected by their smart things, how they are shared and its implication for their moral and ethical lives. The situation is even more exacerbated by the recent exposé that 70 percent of IoT products contain some kind of security vulnerability, according to a research study by

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hewlett-Packard (Hewlett-Packard, 2015). Moreover, it is envisioned that cloud services would be able to auction large volumes of user data to third parties or end users might sell their own daily data feeds to garner financial reward. Furthermore, this information could be processed using various data analytics techniques to develop entire psychological profile of an individual giving insight into the decisions likely to be made by the individual, thereby severely affecting privacy of a citizen. Similar situations might jeopardise businesses and make governments dysfunctional. Generally, industry would happily vouch for such measures as it would facilitate selective marketing of services and subsequent economic benefit, while impact on security and ethics of society is not their necessary concern (Michael, Michael, & Perakslis, 2015; Ridley-siegert, 2015).

Absence of a conscious effort to preserve ethics could lead to unintentional racism. For instance, online ads suggesting arrest are likely to appear with names related to blacks and neutral ads or no ads appear with names related to whites, regardless of whether an arrest record is associated with the name (Sweeney, 2013). A study involving an online rental marketplace suggested that non-black hosts charge 12% more than black hosts for similar facilities (Edelman & Luca, 2014). Similar discrimination has been reported in job applications with white names receiving 50 percent more call-backs for interviews for similar resumes (Bertrand & Mullainathan, 2003). According to a study, searching 'unprofessional hairstyles for work' in Google retrieved images mainly of black women whereas 'professional' type revealed pictures of white women, oftentimes having similar hairstyles differing only in hair type and wearer's skin ("Do Google's 'unprofessional hair' results show it is racist?," 2016). Several other examples of racism in AI have been reported e.g. in Google's photo app, Nikon's camera, crime forecasting etc. ("Artificial Intelligence's White Guy Problem," 2016).

It is not far-fetched that algorithms behind video surveillance systems in Major European and American cities could be trained to exploit certain physical, linguistic or cultural traits to target specific ethnic groups, supposedly more likely to be linked to criminal or terrorist acts (Bianchini & Ávila, 2014). Similarly, a lending algorithm may refuse loan to an economically disadvantaged student by considering him risky, ensuing a vicious circle of discrimination (O'Neil, 2016). There are concerns that racism may also spill into AI based healthcare systems ("If you're not a white male, artificial intelligence's use in healthcare could be dangerous," 2017). If such trends continue unabated, a racist IoT might be brewing.

In order to understand how bias becomes embedded in computing, Friedman and Nissenbaum analyse computer systems from a variety of domains like banking, commerce, education, medicine and law (Friedman & Nissenbaum, 1996). The authors classify bias into three types: pre-existing bias, technical bias and emergent bias. Pre-existing bias exists prior to the design of the system and occurs due to social institutions, practices and attitudes of a society e.g. an automated loan advisor that negatively weights the credit risk of applicants based on address. Technical bias arises due to algorithms that may favour one outcome over another e.g. a ranking algorithm in an airline reservation system that systematically places certain airlines on initial screen giving an undue advantage to the carrier. While pre-existing and technical bias may be identified at the time of design or implementation of system, emergent bias arises later due to changes in the context of use e.g. an ATM interface that relies extensively on written instruction used in village with largely illiterate population.

Artificial Intelligence is expected to be a significant enabler for the IoT, yet there is a growing concern about its prejudice, based on race and gender (Hutson, 2017). AI is a container notion for a highly diverse set of technological schemes that include machine learning, neural networks, genetic algorithms, data mining and a range of highly specific task-based adaptive solutions. As AI algorithms take on cognitive work with social dimensions, the system would necessitate consideration of the social

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

requirements (Bostrom & Yudkowsky, 2014). The need for considering ethics in the development of AI systems is becoming one of the main influential area of research in the last few years (Dignum, 2018). Machines embedded with AI are assumed to promote equality and eliminate bias, however absence of a framework for ethics may endanger fundamental human rights (O'Neil, 2016). A specific class of AI called machine learning utilises a set of data points regarding system functioning (training data) to infer relationships among the system variables and make future decisions. In a recent study, researchers unveiled that human-like biases result while training standard machine learning models on a corpus of text from the Web due to biases inherent in the language (Caliskan, Bryson, & Narayanan, 2017). In 'Weapons of Math Destruction', Cathy O'Neil analyses several application domains like banking, online advertising, judicial decision making, risk models etc. to uncover discrimination embedded in mathematical models, particularly those related to data mining (O'Neil, 2016).

ETHICAL RAMIFICATIONS OF THE INTERNET OF THINGS

Ethics are generally understood as sum total of beliefs, values, morals, etiquettes and manners shared by a people to ensure smooth functioning of society. However, ethical rulings as well as sources used to derive them vary significantly among people, usually extracted from religious texts, verdicts of religious institutions, culture, tradition, legal systems or philosophical school of thought.

As modern technologies increasingly dovetail into our lives, they shape the social, moral and ethical environment in which we live (Jasanoff, 2016). The all-pervading nature of IoT may not provide individuals with the luxury to opt out; perhaps due to perceived anonymity of end users, leaving citizens with arguably limited number of pragmatic options to breakaway (Michael et al., 2014). Governments competing in the 'smart' race may not be able to conceive the challenges awaiting them around the corner as they lack the technical expertise to comprehend design implications of technology. Moreover, as technology seems to grow at much faster pace than the laws to manage them, governments may find themselves helpless in the face of an untoward eventuality as laws regarding such breaches may not have yet been instituted (Brown et al., 2013).

Imagine a world, where all your personal life is being tracked 24X7 to the finest detail by sensors (read spies) embedded all around you: in smart phones, watches, clothes, beds, rooms, buildings, everywhere. And all this information is uploaded in real time to the cloud 'anonymously' and accessible globally to 'authorities' or hackers, who cannot even be prosecuted for any breach, as the laws concerning such violations are yet to cross the minds of lawmakers! Such a scenario would be catastrophic for the moral fabric of society.

As smart socio-technical systems increasingly intertwine with our daily lives, several ethical concerns emerge (Machidon, 2015). Instances of technology systems trespassing on the ethical rights of people are already on the rise. They have been instances of using smart device for live streaming assault to increase the agony of victim ("More than 30 men 'rape teenage girl in Rio' and post video on Twitter," 2016). This is a clear example of how seemingly harmless technology could be used to perpetuate evil. We do not suggest that the camera phone per se is immoral, but absence of an ethical design may inadvertently lend technologies to use in perverse ways. In a socially problematic YouTube phenomenon known as the "Reply Girls" a group of female uploaders began producing videos whose thumbnail depicted their own cleavage and marked them as responses to topical content as a tactic to increase audience and advertising revenue (Sandvig, Hamilton, Karahalios, & Langbort, 2014). There have been reports concerning unfair

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

varying of article prices by online markets based on customer demographics (Jennifer Valentino-DeVries, Singer-Vine, & Ashkan, 2012). Recently Google was slapped a record $ 2.7 Billion for violating business ethics by favouring search results for its own shopping comparison service ("Google Slapped With $2.7 Billion EU Fine Over Search Results," 2017). Perhaps, one of most glaring examples of damage caused by introducing modern technologies on a societal scale without weighing their possible repercussions is the problem of internet addiction that has shattered families, devastated jobs and stalled education (K. S. Young & de Abreu, 2011). According to a report by the Oxford Martin School 85% of jobs in developing countries may be at risk due to automation ("Impact of automation on developing countries puts up to 85% of jobs at risk," 2016). Research into social science perspectives are particularly necessary to design ways and mechanisms to protect vulnerable children from the dangers of the virtual world (Livingstone, Haddon, & Görzig, 2012). Excessive use of smart phone may also hamper personal well-being and human creativity (Ebling, 2015). The disquiet about the use of robotic caregivers and autonomous weapons is well known ("Can we trust robots? [Special Report]," 2016, "Report of the 2016 Informal Meeting of Experts on Lethal Autnomous Weapons Systems (LAWS): United Nations," 2016, "Security Experts Warn Congress That the Internet of Things Could Kill People," 2016, "The Human Rights Implications of Killer Robots: Human Rights Watch," 2014).

Closely related to ethics is the concern of security. Although the former is much wider in scope than the latter, in our opinion there exists a certain degree of overlap between the two. Sometimes ethical considerations may translate into security requirements or failure to provide required security may adversely affect ethical compliance. For instance, smart toys in Germany had to banned because they could be hacked to utter swear words ("German parents told to destroy Cayla dolls over hacking fears," 2017). In 2011, Jerome Radcliffe demonstrated how to remotely control an insulin administration unit and possibly administer lethal insulin dosage to patients (Radcliffe, 2011). Similarly, in 2013 Charlie Miller and Chris Valasek established that it is possible to hack into a Toyota Prius and a Ford Escape to control the acceleration, brakes and even the steering wheel (Valasek & Miller, 2013). In 2016, Nils Rodday reported finding a way to hijack a high-end police drone using equipment costing as low as $40 ("Police drone can be hacked with $40 kit, says researcher," 2016). Such revelations should serve as an eye opener for us to realise the potential of IoT to wreak havoc in our personal and social lives.

Although the examples quoted in the previous few pages demonstrate racism or ethical breaches in specific applications, our concern of possible racism is not limited to a specific set of products or services. Whilst the impact of social media and surveillance systems on modern societies seems obvious, the IoT paradigm could open a Pandora's box of smart user services with possibly dangerous consequences. Just as absence of a concerted effort to preserve societal interests in these examples lead to undesirable outcomes, complacent attitude with regard to ethical concern of IoT could meet a similar fate, albeit with far reaching consequences. In the absence of a sound framework for encoding ethics in machines, the architecture of smart devices may intentionally or unintentionally lend toward adoption in ways uncongenial to moral values of a society. Failure to address ethical considerations inevitably intertwined with a smart device functioning may lead to situations that endanger the physical, mental and emotional wellbeing of humans. It is high time to grapple with the question of encoding ethics in smart things else the dawn of an apocalypse is not farfetched where racism of colour transcends into racism of machines against men.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

PEOPLE POWER

Human beings in all societies are invariably linked to a set of moral, cultural and ethical dimensions, therefore any technology that seeks to address societal welfare must adopt a holistic comprehensive approach. IoT vision should aim at overall wellbeing of society rather than mere implementation of certain technologies, devoid of complex social variables. The view that advanced technologies per say lead to societal progress is reductionist in its approach and too flimsy to stand its claim. Philosophers reflecting on the issue of societal impact of technology have argued that technology must be vigorously critiqued lest it results in dehumanization and loss of human dignity (Michael et al., 2014).

The success of IoT heavily depends on broad public support that can only be garnered by engaging with citizens to help them understand the ramifications of proposed developments. Over thirty years of research into technology and governance strongly supports the view that connections between technology and society are deeply intertwined and as such cannot be ignored (Meijer & Bolivar, 2015). If genuine concerns of those affected by such technologies are not addressed, resistance is inevitable. There have been many instances of smart energy projects in the US and Europe that had to be abandoned because consumers doubted the intentions of authorities (Brown et al., 2013). The authors do not view ethics as a commodity that should only be provided to citizens who pay for it rather as an essential dimension of human existence inseparable from it. Asking people to pay for ethics compliant services is like asking people to pay you for obeying traffic rules. However, personalised ethical requirements a person wishes in his devices may be considered for charge. The authors do not support the idea of tax rebates or additional incentives for those who share private data and view them as tactics of coercion. Ethics conformity is an individual right not an individual privilege. Moreover, when people realise that their values are respected, they are more likely to endorse IoT services. Also, political underpinning of such technology should also be taken cognizance of by human rights organizations lest citizens are coerced by oppressive regimes to subscribe to technologies that compromise essential human values.

In order to highlight the moral and social responsibilities of engineers and technologists for their innovation, Mario Bunge introduced the term technoethics in 1970s (Bunge, 1977). Overtime, technoethics became a field of study that deals with the moral, legal and social issues involving technology. Lupiccini defines technoethics as 'an interdisciplinary field concerned with all ethical aspects of technology within a society shaped by technology' (Luppicini, 2009). Technoethics employs much needed interdisciplinary approach to address the ethical concerns of technological developments in all spheres of life. Key areas of interest for technoethics include engineering ethics, educational technoethics, biotech ethics, environmental technoethics etc. A comprehensive handbook that traces the emergence of the interdisciplinary filed of technoethics comprising over 50 authoritative articles can be found in (Luppicini & Adell, 2009).

One of the earliest notions on the possibility of incorporating values in technological systems can be traced to idea of Value Sensitive Design (VSD) (Hoven, Vermaas, & Poel, 2015). The concept was developed by Batya Friedman and colleagues at the University of Washington in the 1990s to reflect an approach toward the design of technologies that proactively consider human values (Friedman, 1997). VSD seeks to help technologists explicitly incorporate human values into their work. Initially the term value was not defined explicitly and referred to general considerations for human dignity and welfare, however in 2006, the term value was defined as 'what a person or group of people consider important in life' (Friedman, Kahn Jr., & Borning, 2006). In 2015, researchers at the department of ethics and philosophy of technology, Delft University of Technology published a volume that articulates VSD as

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

a feasible possibility to regain moral control over technology (Hoven et al., 2015). The work presents a detailed account of the history and evolution of VSD, its use in various application domains e.g. agriculture, economics, military, architecture and provides a constructive perspective on the design of technologies consistent with moral values of individuals and society.

Recently, there have been a few positive developments to motivate research in the design of people friendly technologies. In September 2016, Stanford University released its first report of the 100 year project, One Hundred Year Study on Artificial Intelligence (AI100), to kindle debate and discussion on the ethical dimension of artificial intelligence ("Artificial Intelligence and Life in 2030: Stanford University," 2016). The report does not offer any solution but seeks to initiate a discussion between scientists, ethicists, policymakers, industry leaders, and the general public (Burrows, 2016). In 2017, IEEE published a document listing recommendations from over 100 global thought leaders in ethics, philosophy, law, AI and robotics delineating the vision for ethically compliant artificial intelligence (Chatila, Firth-Butterfield, & Karachalios, 2017). One of the leading internet pioneers and inventor of TCP/IP, Vinton Cerf also recently expressed the need for a model that guides social and ethical use of IoT technologies (Berman & Cerf, 2017).

It is natural to expect different societies having varying notions of IoT enabled smart services as philosophies of culture, ethics, morality and related laws differ. But these attributes represent significant elements of social milieu and as such should be considered for implementation within machines.

FROM MACHINE MORALITY TO MACHINE MANNERS

The question of implementing moral reasoning into machines is typically explored in an area of study called machine morality, also known as roboethics, machine ethics or friendly AI (Wallach & Allen, 2009). One of the first examples of ethics implementation was a software called Jeremy that employed utilitarianism to measure utility of an actions based on intensity, duration and probability (Anderson, Anderson, & Armen, 2006a). Similarly, a program MedEthEx was developed using deontology to reflect the duties of biomedical ethics (Anderson, Anderson, & Armen, 2006b). Attempts are underway to program ethics into self-driving cars (Goodall, 2016). In spite of a few practical applications that have been made, the field has remained largely theoretical in nature and in still in its infancy (Goodall, 2014). Machine morality concerns itself with the design of full-fledged autonomous moral agents called Artificial Moral Agents (AMAs) endowed with affective cognitive skills for decision making based on ethical, moral, philosophical and religious grounds (Allen, Smit, & Wallach, 2005). Such vision though may not be necessary for smart things in the IoT world.

Portraying ethics in machines is a challenging task primarily because of the fuzziness inherent in the nature of ethics and the difference of opinion surrounding the topic among philosophers, ethicists and religious thinkers. What makes the question particularly difficult is its interdisciplinary nature that spans across diverse fields of study including psychology, sociology, philosophy, religion and artificial intelligence (Anderson & Anderson, 2011). Also, ethicists do not usually express their theories in mathematical forms, so the challenge lies in transforming such ideas into computational structures amenable to implementation. Even a symbolic representation of ethical concepts, theories and principles cannot be found, let alone some set of formalism (Lin, Abney, & Bekey, 2011). A less stringent form of machine morality appears to be a promising area to help achieve such objectives.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In spite of the cultural differences, human beings throughout the world share a rich reservoir of ethics, representing a fertile ground for further exploration. A recent proposal of 23 principles called Asilomar AI Principles to serve as a guideline in the design of human friendly AI appears promising toward this end ("Asilomar AI Principles," 2017). Also, smart things do not need to be endowed with full-fledged moral reasoning to function as AMAs rather only a small subset of ethical considerations are expected to be relevant to their context. It is reasonable to identify such critical values to ensure ethics compliant behaviour than to identify principles for a universal moral theory.

The authors adopt an ontology agnostic engineering perspective on the question of ethics whether facilitated by utilitarianism, deontology, Kantian logic, prima facie duties, divine command theory or any other approach, depending on the situation (Bringsjord, Arkoudas, & Bello, 2006). The ethical context of this work does not pertain to a specific country or state. We focus our attention to the design of mechanisms that can enable us to codify a given set of ethics in machines. Our objective is to enable a smart thing exhibit moral ethical behaviour relevant to its context, regardless of ontology. We aim to outline how loT can be designed to display morally right and omit morally wrong behaviour. This work focuses exclusively on the design of mechanisms to represent ethics of conduct within smart devices and discussions on ethics of institution or character are outside the scope of the article.

Moreover, the authors expand the notion of ethics to include 'safety instructions' acquiescent to computational representation like 'keep medicine away from children' by employing smart locks for containers, a friendly knife that does not hurt, a camera that respects privacy of others etc. Safety instructions may no longer need to be printed on the gadget guides rather be imprinted in the design of device itself. The scope of ethics also considers certain behaviour that might be expected of smart things in a certain context that does not typically follow from any ethical theory. For example, Honolulu in July 2017 became the first major US city to pass the 'Distracted Walking Law', that bans pedestrians from looking at smart devices while crossing the road, having caused over 11,000 injuries over the past decade in the US ("Honolulu bans texting while crossing streets in bid to curb injuries," 2017). Similarly, such requirements as cars should keep right, children should not watch cartoons for more than an hour, traffic routes to be followed by autonomous cars etc. are considered within the purview of ethics.

The ensemble of ethical rules, moral behaviours, religious instructions, legal requirements, cultural values, safety instructions, code of conduct, management policies and any other associated rules applicable to a smart device (or any system) in its context are called as machine manners to distinguish it from the utopian vision of machine morality. By implementing machine manners, ordinary objects transform into extraordinary objects capable of displaying ethical behaviour. The concept promises to have numerous applications areas e.g. ethical toys, ethical home appliances, ethical cars etc. evolving into bigger ethical structures like children friendly technology ecosystem etc.

The ethical representation also needs to incorporate dynamicity in its implementation to account for time and space varying nature of ethics, particularly in the context of legal dimensions that change quite often. Smart devices need to reflect such changes in a timely fashion e.g. all smart cars in a city should know the traffic route changes applicable on a certain festival, sports event etc. One possible approach to ensure such behaviour is via ethics updates in a manner similar to security updates on smart devices. Also, social ethics is not homogeneous throughout society and is often challenged through argument by members of the same or other society. Moreover, the nature of ethics may differ among a people with regards to age group, personality, ethnicity etc. Such fine-grained ethical values can be integrated in the general ethical framework to allow personalisation of individual preferences. These fine grained ethical variations within a society are called ethics dialects, possibly incorporated by machine learning.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

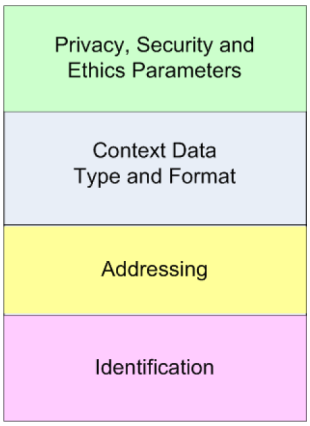
MACHINE MANNERS VIA ETHICS LAYER

In order to implement machine manners, the authors propose using the notion of object that represents ethics relevant to a smart thing context and incorporating socio-cultural ethical aspects of a society into the protocol stack of smart things by dedicating a separate ethics layer. By object the authors mean a virtual manifestation of real world thing that specifies context data, security and relevant ethics parameters. Our object has two important features that distinguish it from other objects proposed in literature (Nitti, Pilloni, Colistra, & Atzori, 2016). Firstly, security considerations have been included within the object. Secondly, the realm of object definition is expanded to include ethics parameters in order to facilitate implementation of machine manners. Employing such concept enables us to consider security and ethics from the very outset rather than devise quick fix solutions in case of unanticipated breach. A typical object would contain device identification data, network address, data relevant to its context, security and ethics parameters as shown in Figure 1.

The concept of object allows greater control over smart thing by specifying what it senses and how sensed data is utilized. Also, the same thing could be configured into varying object definitions to support different ethical scenarios. The design of the proposed object is motivated by several considerations. A very small subset of ethical considerations are expected to be relevant in the scope of a smart thing. The concept of object offers a feasible way to represent such ethical considerations to ensure ethics compliant machine behaviour. Also, ethical norms and legal rulings differ among societies and also change with time. Smart things should be able to reflect such changes in their behaviour in a timely manner. Such flexibility can easily be incorporated by modification of ethics parameters in an object or class of objects without affecting other smart things. The concept of object also lends itself to the possibility of reuse to facilitate scalability in the design of machine manners.

Often bird's eye view of sensed data over a large spectrum is more useful than data gathered by individual sensors. IoT employs data analytic techniques on gathered data to provide valuable insights into overall functioning of a system. Similar mechanism can be used to achieve ethical Device-to-Device (D2D) interaction. Simple objects representing ethics of individual things coalesce into compound objects forming communities of devices. While object represents ethics pertinent to an individual device,

Figure 1. Representation of a simple object



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

compound object characterises ethics relevant to a group of devices represented by it. This enables us to form communities of smart devices just as among humans, sharing a set of ethics. Compound object captures ethics of a group to facilitate ethics compliant interaction among the members as well as other compound objects. Subscription to a group (compound object) may also be dynamic e.g. a given smart phone may belong to smart home or smart school at different times. This mechanism enables a smart device exhibit varying ethical responses to different situations. Several hierarchies of compounds objects may also be possible, for instance, when compound objects representing smart homes aggregate to form compound object for smart neighbourhood, etc.

The authors propose dedicating a separate ethics layer in the protocol stack of smart things to facilitate smooth functioning of IoT within a particular moral ethical context. The ethics layer enables us to account for ethical responsibilities of smart devices vis-a-vis society so that inadvertent physical, emotional or psychological harm to human beings is avoided. Ethics layer uses objects and compound objects as building blocks to facilitate ethics compliant behaviour of smart things. Such mechanism ensures that devices operate ethically not only at individual level but also at D2D level to give rise to high order ethical structures e.g. ethical home, ethical office, ethical university, ethical city, etc. The implementation of such ethical framework possibly uses a set of semantic rules to codify the moral, religious, ethical beliefs of a society and ensures that various IoT applications are compliant with it. The ethics layer fundamentally expresses social values of a society and guides decisions consistent with ethical legal standards to safeguard the moral fabric of society. The technical specifications of the proposed object and ethics layer have been outlined in our recent work (Sholla, Mir, & Chishti, 2017a). Key ethical requirements of a machine are identified and expressed in the form of propositional statements within an object. These propositional variables are used to determine various scenarios that could occur between smart devices and its environment. Finally, a set of Boolean equations are used to express ethically consistent machine behaviour. The use of the method in designing a smart healthcare device that shares disease related information with relatives and doctor in an ethically consistent manner as well as respecting relevant smart city medical laws is described in (Sholla, Mir, & Chishti, 2017b). In a similar fashion, a kitchen knife may be designed that functions only in the kitchen of its owner, only cuts objects it is authorised to cut and gets disabled in the hands of children. Similarly, a car that prevents locking of an infant inside the car and sends alert message to the parent or a camera phone the disables selfie mode at locations flagged dangerous. Such approach may be used in designing an array of ethical devices like ethical gadgets, ethical toys, ethical home appliances etc. The ethics layer essentially complements the IoT paradigm with a human touch lest we are snarled up in an apartheid state of 'things'!

The only work that seeks to design a framework for ethics in IoT is a policy based approach by Baldini (Baldini, Botterman, Neisse, & Tallacchini, 2016), but the authors view ethics in a very restricted sense primarily limited to privacy profiles offered to citizens as "ethical choices" they pay for. The approach proposed by the authors in this work is comprehensive as indicated by the term machine manners. Also, this work employs the concept of object (and compound object) to indicate ethics relevant to a smart thing(s). Furthermore, a separate ethics layer is set aside in the protocol stack of smart things to facilitate ethically compliant behaviour, not only at device level but also wider ecosystem level.

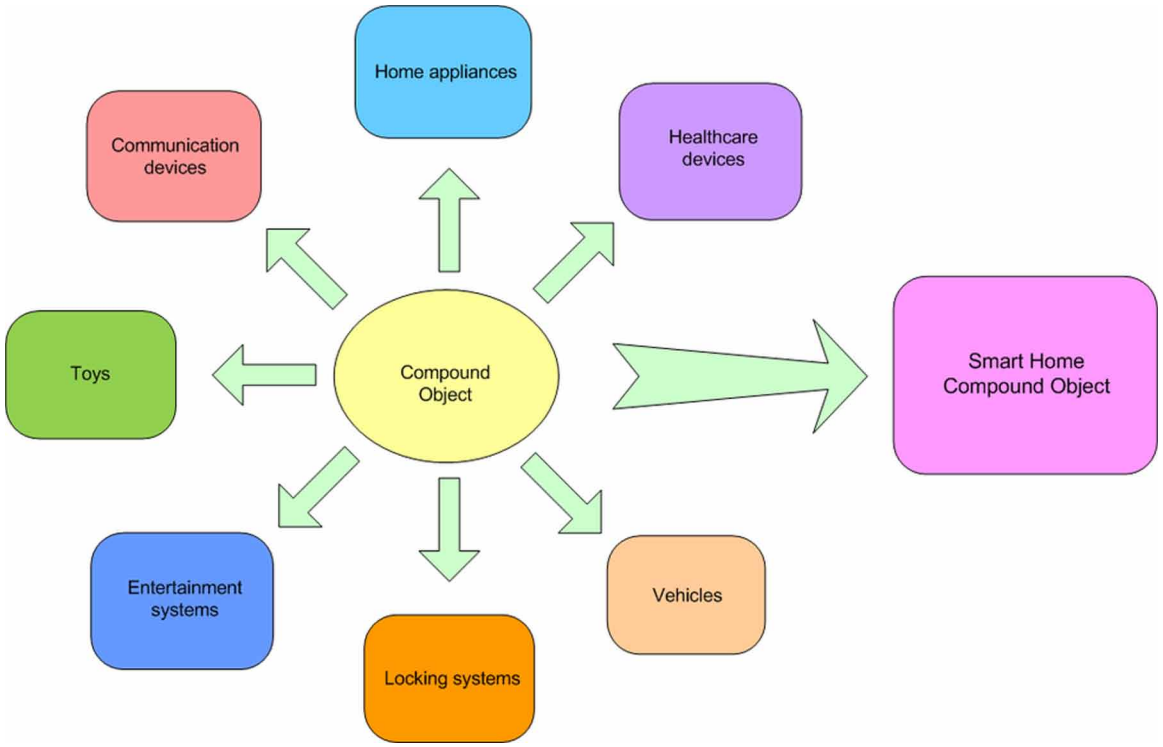
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

USE CASE: ETHICAL HOME

In order to demonstrate, how a well-designed ethical IoT system would behave, let us consider the example of an ethical smart home. Depending on the moral values, cultural expectations, religious beliefs adopted by the family etc., ethical requirements for the various smart devices are identified. Each individual thing is represented by an object that specifies the set of ethics applicable to it. The devices not only imbue general ethical conduct expected of all smart devices in the smart city, but also exhibit fine grained ethics dialect of the family. In one possible organisation, communication devices like smart phones, tablets and computers merge to form compound communication object, home appliances combine to form compound appliance object, various fitness or heath devices used by household form compound fitness object etc. as shown in Figure 2.

In a typical scenario, compound communication object ensures various communication devices respect the ethics of smart home e.g. by not disturbing the family with unwanted notifications at dinner or while engaged in spiritual activity etc. Home appliances take necessary precaution in order to avoid hurt, burn or harm to inmates especially children e.g. by using ethical knife, ethical washing machine, ethical burner etc. Smart toys limit screen exposure time to children based on age to prevent unhealthy usage, alert parents for addiction symptoms etc. Similarly, entertainment devices do not expose children to content that may be deemed unethical by society. Smart car ensures that children are not inadvertently locked inside, is aware of the updated traffic rules, speed limits to be followed etc. Healthcare

Figure 2. Compound Objects for Ethical Home



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

devices share only selective pieces of information with authorised doctor and family (or extended family), depending on ethical beliefs of the individual or norms of society. The ethics layer also facilitates ethical D2D interaction e.g. by specifying how entertainment and communication systems interact with smart toys, how locking systems interact with smart cars of the family or other cars, which devices are permitted to send status message to home owner under what conditions etc. In a similar fashion, entire ethical home may be represented by a smart home compound object for possible expansion into ethical neighbourhood. The ethical injunctions followed by the smart devices in this example are only used for the purpose of illustration without committing to any specific ethical theory. In general, ethical theories or principles to be adopted in a given social context may be prepared by a team of experts including philosophers, ethicists, jurists, religious scholars etc. For instance, Kimberly Young suggests a set of guidelines to regulate the functioning of smart devices for children of various age groups (Young, 2017).

RESEARCH DIRECTIONS

The challenging and exciting field of machine manners necessitates research from a multitude of fields that include sociology, psychology, law, religion, philosophy and artificial intelligence. Owing to its interdisciplinary nature, a healthy participation and collaboration among various areas of human learning would be necessary to accomplish holistic development of modern societies. Though our work attempted to outline a possible preliminary approach, several challenges need to be overcome in order to bring the field to fruition.

This work does not consider the question of ethical dilemmas (e.g. Trolley Problem), however it is an important challenge that needs consideration (Wallach & Allen, 2009). Design of moral agents that resolve ethical dilemmas in a split second is not expected to be a cakewalk. A practical workaround might be to consider the option of human in the loop, where experienced ethicists, possibly in a particular area (school, university, city etc.), supervise ethical devices in real time to resolve such conflicts. It may not be wise to stall engineering design of ethics in the face of questions that have eluded philosophers for centuries.

Another perplexing question facing the research community is the issue of accountability i.e. who to hold responsible for decisions made by smart devices. As autonomous technologies continue to evolve, it will be increasingly difficult to design a framework for accountability (Berman & Cerf, 2017). The question of moral agency vis-a-vis technological artefacts represents a challenge to the traditional notion of moral agency. Can technological artefacts be considered true moral agents? Although, AI systems are generally believed to lack a moral status, exact parameters that ground moral status are still unclear (Bostrom & Yudkowsky, 2014). A detailed discussion on the concept of artificial moral agency as it pertains to technoethics can be found here (Sullins, 2009). Also, it may not be wise to vindicate human agents working behind the scenes of all responsibility vis-a-vis moral bankruptcy of machines. A legal framework for machines akin to liability laws, possibly implemented at ethics layer, would be required to establish responsibility in the event of ethical breach by machines. Furthermore, the mechanism of human in the loop could be explored as a viable option for accountability management. Notwithstanding such implementation, legal systems would need to remain proactive to address any judicial eventuality that may jeopardize the well-being of society. The authors do not suggest that judges and courts be replaced by computers.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Also, situations may inevitably arise where smart devices have conflicting ethical parameters. The ethics layer needs to support a viable negotiation mechanism in order to handle such eventuality, possibly in a manner similar to private international law (Dung & Sartor, 2011). Although our work has attempted to sketch a possible solution, actual implementation of ethics layer and how it integrates with the protocol stack of smart devices requires considerable research. There is a growing interest in the use of smart contracts to facilitate automation of financial interactions among devices in the IoT ecosystem (Christidis & Devetsikiotis, 2016). Similar approaches to implementing ethics appear to be a promising future direction. Another possibility is to represent social norms in the form of a semantic network in a manner similar to licensing metadata or machine-readable permissions such as Creative Commons (Malle, 2016). The question also merits considerable attention from the field of Human Computer Interaction (HCI) (Davis & Waycott, 2015). Once ethics of smart devices have been sufficiently expressed, they would also need to be secured against any possible intrusion.

SUMMARY

Notwithstanding the potential of IoT to revolutionise our personal and social lives, absence of a solid framework of ethics may lead to situations that jeopardise essential human values. Scientific work regarding IoT focuses on various domain specific applications but design of a mechanism for implementing ethics in IoT has not been adequately considered. However, it is necessary to evoke an ethical dimension in the design of IoT lest technological developments undercut essential values of citizenship.

The authors adopt a comprehensive approach toward incorporating ethics in IoT that includes ethical rules, moral behaviours, religious instructions, legal requirements, cultural values, safety instructions, code of conduct, management policies and any other associated rules, encompassed by the term machine manners. The concept of object (and compound object) is suggested to implement ethics relevant to the context of smart device(s), a direction that has not been explored previously. While object signifies virtual manifestation of an individual thing, multiple objects coalesce into compound object to denote ethics applicable to an entire group. Moreover, the authors propose dedicating a separate ethics layer in the protocol stack of smart devices that uses objects (and compound objects) as building blocks to realise smart spaces congruent with ethical expectations.

Our humble effort only symbolises a drop in an ocean that awaits exploration. Significant research effort would be necessary to realise a concrete solution to this taxing problem. It is hoped that this work offers a starting point for the research community to engage with the ethical question surrounding the IoT paradigm to realise a more humane and compassionate world.

REFERENCES

Alexander, L. (2016, April). Do Google's "unprofessional hair" results show it is racist? *The Guardian*. Retrieved from www.theguardian.com/technology/2016/apr/08/does-google-unprofessional-hair-results-prove-algorithms-racist-

Allen, C., Smit, I., & Wallach, W. (2005). Artificial morality: Top-down, bottom-up, and hybrid approaches. *Ethics and Information Technology*, *7*(3), 149–155. doi:10.100710676-006-0004-4

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Anderson, M., & Anderson, S. L. (2011). *Machine Ethics*. Cambridge University Press. doi:10.1017/CBO9780511978036

Anderson, M., Anderson, S. L., & Armen, C. (2006a). An approach to computing ethics. *IEEE Intelligent Systems*, *21*(4), 56–63. doi:10.1109/MIS.2006.64

Anderson, M., Anderson, S. L., & Armen, C. (2006b). MedEthEx : A Prototype Medical Ethics Advisor. In *Proceedings of the 18th conference on Innovative applications of artificial intelligence* (Vol. 2, pp. 1759–1765).

Crawford, K. (2016, June). Artificial Intelligence's White Guy Problem. *The New York Times*. Retrieved from https://www.nytimes.com/2016/06/26/opinion/sunday/artificial-intelligences-white-guy-problem.html

Asilomar. (2017). A. I. Principles. Retrieved July 30, 2017, from https://futureoflife.org/ai-principles/

Baldini, G., Botterman, M., Neisse, R., & Tallacchini, M. (2016). Ethical Design in the Internet of Things. *Science and Engineering Ethics*, 1–21. doi:10.100711948-016-9754-5 PMID:26797878

Berman, F., & Cerf, V. G. (2017). Social and ethical behavior in the internet of things. *Communications of the ACM*, *60*(2), 6–7. doi:10.1145/3036698

Bertrand, M., & Mullainathan, S. (2003). Are Emily and Greg More Employable than Lakisha And Jamal? A Field Experiment on Labor Market Discrimination. *NBER*. Retrieved from http://www.nber.org/papers/w9873

Bianchini, D., & Ávila, I. (2014). Smart cities and their smart decisions: Ethical considerations. *IEEE Technology and Society Magazine*, *33*(1), 34–40. doi:10.1109/MTS.2014.2301854

Bostrom, N., & Yudkowsky, E. (2014). The Ethics of Artificial Intelligence. In *The Cambridge Handbook of Artificial Intelligence*. Cambridge University Press. doi:10.1017/CBO9781139046855.020

Bringsjord, S., Arkoudas, K., & Bello, P. (2006). Toward a general logicist methodology for engineering ethically correct robots. *IEEE Intelligent Systems*, *21*(4), 38–44. doi:10.1109/MIS.2006.82

Brown, I., Crump, J., Boos, D., Dutton, W., Kinder, K., & Kortuem, G. … Marshall, R. (2013). The Societal Impact of the Internet of Things. *BCS*, 14. Retrieved from http://www.bcs.org/upload/pdf/societal-impact-report-feb13.pdf

Bunge, M. (1977). Towards a Technoethics. *The Monist*, *60*(1), 96–107. doi:10.5840/monist197760134

Burrows, L. (2016, September). What artificial intelligence will look like in 2030. *Harvard Gazette*. Retrieved from news.harvard.edu/gazette/story/2016/09/what-artificial-intelligence-will-look-like-in-2030/

Caliskan, A., Bryson, J. J., & Narayanan, A. (2017). Semantics derived automatically from language corpora contain human-like biases. *Science*, *356*(6334), 183–186. doi:10.1126cience.aal4230 PMID:28408601

Chatila, R., Firth-Butterfield, K., & Karachalios, J. C. H. K. (2017, March). IEEE Global Initiative Aims to Advance Ethical Design of AI and Autonomous Systems. *IEEE Spectrum*. Retrieved from http://spectrum.ieee.org/automaton/robotics/artificial-intelligence/ieee-global-initiative-ethical-design-ai-and-autonomous-systems

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chernyi, A. I. (2010). On the brink of the information society. *Automatic Documentation and Mathematical Linguistics*, *44*(5), 249–254.

Christidis, K., & Devetsikiotis, M. (2016). Blockchains and Smart Contracts for the Internet of Things. *IEEE Access: Practical Innovations, Open Solutions*, *4*, 2292–2303. doi:10.1109/ACCESS.2016.2566339

Davis, H., & Waycott, J. (2015). Ethical Encounters. In *Proceedings of the Annual Meeting of the Australian Special Interest Group for Computer Human Interaction on - OzCHI '15* (pp. 667–669). New York, New York, USA: ACM Press. 10.1145/2838739.2838834

De Torres, C. (2016). Can we trust robots? *IEEE Spectrum*, *53*(6), 26–27. doi:10.1109/MSPEC.2016.7473148

Dignum, V. (2018). Ethics in artificial intelligence: Introduction to the special issue. *Ethics and Information Technology*, *20*(1), 1–3. doi:10.100710676-018-9450-z

Dung, P. M., & Sartor, G. (2011). The modular logic of private international law. *Artificial Intelligence and Law*, *19*(2–3), 233–261. doi:10.100710506-011-9112-5

Ebling, M. R. (2015). The importance of being... Bored. *IEEE Pervasive Computing*, *14*(2), 5–8. doi:10.1109/MPRV.2015.43

Edelman, B., & Luca, M. (2014). Digital Discrimination: The Case of Airbnb.com. *HBS*. Retrieved from http://www.hbs.edu/faculty/PublicationFiles/Airbnb_92dd6086-6e46-4eaf-9cea-60fe5ba3c596.pdf

Forster, K. (2016, May). More than 30 men "rape teenage girl in Rio" and post video on Twitter. *The Independent*. Retrieved from http://www.independent.co.uk/news/world/americas/rio-olympics-rape-girl-video-twitter-twitter-video-uploaded-a7052106.html

Friedman, B. (1997). *Human values and the design of computer technology*. Cambridge University Press.

Friedman, B., Kahn Jr., P. H., & Borning, A. (2006). Value Sensitive Design and Information Systems. *Human-Computer Interaction and Management Information Systems: Foundations*, 1-27.

Friedman, B., & Nissenbaum, H. (1996). Bias in computer systems. *ACM Transactions on Information Systems*, *14*(3), 330–347. doi:10.1145/230538.230561

German parents told to destroy Cayla dolls over hacking fears. (2017, February). *BBC News*. Retrieved from http://www.bbc.com/news/world-europe-39002142

Goodall, N. J. (2014). Machine Ethics and Automated Vehicles. In *Road Vehicle Automation* (pp. 93–102). Springer. doi:10.1007/978-3-319-05990-7_9

Goodall, N. J. (2016). Can you program ethics into a self-driving car? *IEEE Spectrum*, *53*(6), 28–58. doi:10.1109/MSPEC.2016.7473149

Drozdiak, N. & Schechner, S. (2017, June). Google slapped with $2.7 Billion EU Fine Over Search Results. *The Wall Street Journal*. Retrieved from www.wsj.com/articles/google-slapped-with-2-7-billion-eu-fine-over-search-results-1498556971

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Haag, A., Mills, K. & Hepkins, F. (2014). The Human Rights Implications of Killer Robots. *Human Rights Watch*. Retrieved August 7, 2017, from https://www.hrw.org/report/2014/05/12/shaking-foundations/human-rights-implications-killer-robots

Hart, R. (2017, July). If you're not a white male, artificial intelligence's use in healthcare could be dangerous. *Quartz.*

Hewlett-Packard. (2015). Internet of Things Research Study 2015 Report, 6. Retrieved from http://www8.hp.com/h20195/V2/GetPDF.aspx/4AA5-4759ENW.pdf

Honolulu bans texting while crossing streets in bid to curb injuries. (2017, July). *BBC News*. Retrieved from http://www.bbc.com/news/world-us-canada-40765087

Hutson, M. (2017, April 13). Even artificial intelligence can acquire biases against race and gender. *Science Mag*. doi:10.1126cience.aal1053

Impact of automation on developing countries puts up to 85% of jobs at risk. (2016, January). *Oxford Martin School*. Retrieved from https://www.oxfordmartin.ox.ac.uk/news/201601_Technology_at_Work_2

Jasanoff, S. (2016). *The Ethics of Invention: Technology and the Human Future*. WW Norton & Company.

Lin, P., Abney, K., & Bekey, G. A. (2011). *Robot ethics: the ethical and social implications of robotics*. MIT Press.

Livingstone, S., Haddon, L., & Görzig, A. (2012). *Children, risk and safety on the Internet: Research and policy challenges in comparative perspective*. Policy Press at the University of Bristol. doi:10.1332/policypress/9781847428837.001.0001

Luppicini, R. (2009). The emerging field of technoethics. In *Handbook of research on technoethics* (pp. 1–19). Information Science Reference. doi:10.4018/978-1-60566-022-6.ch001

Luppicini, R., & Adell, R. (2009). *Handbook of research on technoethics*. Information Science Reference. doi:10.4018/978-1-60566-022-6

Machidon, O. M. (2015). Societal implications of current and emerging "smart" technologies. *International Journal of Technoethics*, *6*(1), 60–70. doi:10.4018/ijt.2015010105

Malle, B. F. (2016). Integrating robot ethics and machine morality: The study and design of moral competence in robots. *Ethics and Information Technology*, *18*(4), 243–256. doi:10.100710676-015-9367-8

Manyika, J., Chui, M., Bughin, J., Dobbs, R., Bisson, P., & Marrs. (2013). Disruptive technologies: Advances that will transform life, business, and the global economy. *McKinsey Global Institute*, (May), 163. Retrieved from http://www.mckinsey.com/insights/business_technology/disruptive_technologies%5Cnhttp://www.chrysalixevc.com/pdfs/mckinsey_may2013.pdf

Meijer, A., & Bolivar, M. P. R. (2015). Governing the smart city: A review of the literature on smart urban governance. *International Review of Administrative Sciences*, *82*(2), 392–408. doi:10.1177/0020852314564308

Michael, M. G., Michael, K., & Perakslis, C. (2014). Uberveillance and the internet of things and people. In *Proceedings of 2014 International Conference on Contemporary Computing and Informatics, IC3I 2014* (pp. 1381–1386). 10.1109/IC3I.2014.7019829

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Michael, M. G., Michael, K., & Perakslis, C. (2015). Uberveillance, the web of things, and people: What is the culmination of all this surveillance? *IEEE Consumer Electronics Magazine*, *4*(2), 107–113. doi:10.1109/MCE.2015.2393007

Nitti, M., Pilloni, V., Colistra, G., & Atzori, L. (2016). The Virtual Object as a Major Element of the Internet of Things: A Survey. *IEEE Communications Surveys and Tutorials*, *18*(2), 1228–1240. doi:10.1109/COMST.2015.2498304

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. NY: Crown Publishing Group.

Orcutt, M. (2016, December). Security Experts Warn Congress That the Internet of Things Could Kill People. *MIT Technology Review*. Retrieved from https://www.technologyreview.com/s/603015/security-experts-warn-congress-that-the-internet-of-things-could-kill-people/

Picard, R. W. (2000). *Affective computing*. MIT Press.

Police drone can be hacked with $40 kit, says researcher. (2016, March). *BBC News*. Retrieved from http://www.bbc.com/news/technology-35709676

Radcliffe, J. (2011). Hacking medical devices for fun and insulin: Breaking the human SCADA system. In *Black Hat Conference Presentation Slides*. Retrieved from https://media.blackhat.com/bh-us-11/Radcliffe/BH_US_11_Radcliffe_Hacking_Medical_Devices_WP.pdf

Report of the 2016 Informal Meeting of Experts on Lethal Autnomous Weapons Systems (LAWS): United Nations. (2016). Retrieved from https://www.unog.ch

Ridley-siegert, T. (2015). Data privacy: What the consumer really thinks. *Journal of Direct, Data and Digital Marketing Practice, 17*(1), 30–35.

Sandvig, C., Hamilton, K., Karahalios, K., & Langbort2, C. (2014). Auditing Algorithms : Research Methods for Detecting Discrimination on Internet Platforms. In *"Data and Discrimination: Converting Critical Concerns into Productive Inquiry," a preconference at the 64th Annual Meeting of the International Communication Association*, Seattle, WA (pp. 1–23).

Sholla, S., Mir, R. N., & Chishti, M. A. (2017a). Ethics Aware Object Oriented Smart City Architecture. *China Communications*, *14*(5), 160–173. doi:10.1109/CC.2017.7942323

Sholla, S., Mir, R. N., & Chishti, M. A. (2017b). Incorporating Ethics in Internet of Things (IoT) Enabled Connected Smart Healthcare. In *Proceedings - 2017 IEEE 2nd International Conference on Connected Health: Applications, Systems and Engineering Technologies, CHASE 2017*. doi:10.1109/CHASE.2017.93

Stanford University. (2016). Artificial Intelligence and Life in 2030. Retrieved from https://ai100.stanford.edu/

Sullins, J. P. (2009). Artificial Moral Agency in Technoethics. In *Handbook of Research on Technoethics* (pp. 205–221). Hershey, PA: IGI Global. doi:10.4018/978-1-60566-022-6.ch014

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sweeney, L. (2013). Discrimination in online ad delivery: Google ads, black names and white names, racial discrimination, and click advertising. *ACM Queue; Tomorrow's Computing Today*, *11*(3), 1–19. doi:10.1145/2460276.2460278

Valasek, C., & Miller, C. (2013). Adventures in Automotive Networks and Control Units. *Technical White Paper*, 99. Retrieved from http://www.ioactive.com/pdfs/IOActive_Adventures_in_Automotive_Networks_and_Control_Units.pdf

Valentino-Devries, J., Singer-Vine, J., & Soltani, A. (2012, December). Websites Vary Prices, Deals Based on Users' Information. *The Wall Street Journal*. Retrieved from www.wsj.com/articles/SB10001424127887323777204578189391813881534

van den Hoven, J., Vermaas, P. E., & van de Poel, I. (2015). *Handbook of ethics, values, and technological design: sources, theory, values and application domains*. Springer. doi:10.1007/978-94-007-6970-0

Wallach, W., & Allen, C. (2009). *Moral machines: teaching robots right from wrong*. Oxford University Press. doi:10.1093/acprof:oso/9780195374049.001.0001

Young, K. (2017). Children and Technology: Guidelines for Parents-Rules for Every Age. *IEEE Technology and Society Magazine*, *36*(1), 27–29. doi:10.1109/MTS.2017.2670223

Young, K. S., & de Abreu, C. N. (2011). *Internet addiction: A handbook and guide to evaluation and treatment. Internet addiction: A handbook and guide to evaluation and treatment*. John Wiley & Sons.

This research was previously published in the International Journal of Technoethics (IJT), 9(2); edited by Rocci Luppicini; pages 62-76; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 60
Security Solutions for Intelligent and Complex Systems

Stuart Armstrong
Future of Humanity Institute, UK

Roman V. Yampolskiy
JB Speed School of Engineering, USA

ABSTRACT

Superintelligent systems are likely to present serious safety issues, since such entities would have great power to control the future according to their possibly misaligned goals or motivation systems. Oracle AIs (OAI) are confined AIs that can only answer questions and do not act in the world, represent one particular solution to this problem. However even Oracles are not particularly safe: humans are still vulnerable to traps, social engineering, or simply becoming dependent on the OAI. But OAIs are still strictly safer than general AIs, and there are many extra layers of precautions we can add on top of these. This paper begins with the definition of the OAI Confinement Problem. After analysis of existing solutions and their shortcomings, a protocol is proposed aimed at making a more secure confinement environment which might delay negative effects from a potentially unfriendly superintelligence while allowing for future research and development of superintelligent systems.

INTRODUCTION

With the likely development of superintelligent programs in the near future, many scientists have raised the issue of safety as it relates to such technology (Bostrom, 2006; Chalmers, 2010; Hall, 2000; Hibbard, 2005; Yampolskiy, 2011a, 2011b; Yampolskiy & Fox, 2012a, 2012b; Yudkowsky, 2008). A common theme in Artificial Intelligence (AI[1]) safety research is the possibility of keeping a superintelligent agent in a sealed hardware so as to prevent it from doing any harm to humankind. Such ideas originate with scientific visionaries such as Eric Drexler who has suggested confining transhuman machines so that their outputs could be studied and used safely (Drexler, 1986). Similarly, in 2010 David Chalmers

DOI: 10.4018/978-1-7998-0951-7.ch060

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

proposed the idea of a "leakproof" singularity (Chalmers, 2010). He suggested that for safety reasons, AI systems first be restricted to simulated virtual worlds until their behavioral tendencies could be fully understood under the controlled conditions.

This chapter is based on combined and extended information from three previously published papers: (Armstrong, 2011; Armstrong, Sandberg, & Bostrom, 2012; Yampolskiy, 2012a)*. We evaluate feasibility of previously presented proposals and suggest a protocol aimed at enhancing safety and security of such methodologies. While it is unlikely, that long-term and secure confinement of AI is possible, we are hopeful that the proposed protocol will give researchers a little more time to find a permanent and satisfactory solution for addressing existential risks associated with appearance of superintelligent machines.

In this chapter we will review specific proposals aimed at creating restricted environments for safely interacting with artificial minds. The key question is: are there strategies that reduce the potential existential risk from a superintelligent AI so much that while implementing it as a free AI would be impermissible a confined implementation would be permissible? The chapter will start by laying out the general design assumptions for the confined AI and formalizing the notion of confinement. Then it will touch upon some of the risks and dangers deriving from the humans running and interaction with the confined AI. The final section looks at some of the other problematic issues concerning the confined AI, such as its ability to simulate human beings within it and its status as a moral agent itself.

Motivation for AI Confinement

There are many motivations to pursue the goal of developing AI. While some motivations are non-instrumental, such as scientific and philosophical curiosity about the nature of thinking or a desire for creating non-human beings, a strong set of motivations is the instrumental utility of AI. Such machines would benefit their owners by being able to do tasks that currently require human intelligence, and possibly tasks that are beyond human intelligence. From an economic perspective the possibility of complementing or substituting expensive labor with cheaper software promises very rapid growth rates and high productivity (Hanson, 2001, 2008; Kaas, Rayhawk, Salamon, & Salamon, 2010). The introduction of sufficiently advanced AI would have profound effects on most aspects of society, making careful foresight important.

While most considerations about the mechanization of labor have focused on AI with intelligence up to the human level, there is no strong reason to believe humans represent an upper limit of possible intelligence. The human brain has evolved under various biological constraints (e.g. food availability, birth canal size, trade-offs with other organs, the requirement of using biological materials) which do not exist for an artificial system. Beside different hardware, an AI might employ more effective algorithms that cannot be implemented well in the human cognitive architecture (e.g. making use of very large and exact working memory, stacks, mathematical modules or numerical simulation), or employ tricks that are not feasible for humans, such as running multiple instances whose memories and conclusions are eventually merged. In addition, if an AI system possesses sufficient abilities, it would be able to assist in developing better AI. Since AI development is an expression of human intelligence, at least some AI might achieve this form of intelligence, and beyond a certain point would accelerate the development far beyond the current rate (Chalmers, 2010; Kurzweil, 2005).

While the likelihood of superintelligent AI is hotly debated, the mere possibility raises worrying policy questions. Since intelligence implies the ability to achieve goals, we should expect superintelligent systems to be significantly better at achieving their goals than humans. This produces a risky power differential. The appearance of superintelligence appears to pose an existential risk: a possibility that humanity is

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

annihilated or has its potential drastically curtailed indefinitely (Bostrom, 2001). This could come about through a number of ways: by enabling self-reinforcing systems that limit human potential (e.g. a global police state (Caplan, 2008)), by out-competing humans or human values (see the 'mindless outsourcers' in (Bostrom, 2004)), or by acting with great power in such a malevolent or indifferent manner that humanity goes extinct. The last possibility could occur due to badly formulated initial motivations or goals, or gradual evolution towards human-unfriendly behaviors (Omohundro, February 2008). Even if most superintelligences are assumed to be human-friendly it appears hard to guarantee that dangerous superintelligences will not emerge unless precautions are made.

In the space of possible motivations likely a very small fraction is compatible with coexistence with humans. A randomly selected motivation can hence be expected to be dangerous. A simple example is that of a paperclip maximizer: an AI with a utility function that seeks to maximize the number of manufactured paperclips. This goal is a not too implausible test command to a new system, yet it would be willing to sacrifice the world and everyone in it if necessary for making more paperclips. If superintelligent the AI would be very good at converting the world into paperclips, even if it realizes its creators actually do not want that many paperclips – stopping would lead to fewer manufactured paperclips (Bostrom, 2006; LessWrong, Available on April 25, 2012). It is a common problem in computer programming, with the program going beyond the implicit bounds of what was expected of it (or users giving in retrospect mistaken commands). So unless valuing humans is an integral part of the superintelligence's setup, we can expect that we will be seen as mere tools or obstacles for its own goals.

There are several approaches to AI risk. The most common at present is to hope that it is no problem: either sufficiently advanced intelligences will converge towards human-compatible behavior, a solution will be found closer to the time when they are actually built, or they cannot be built in the first place. While such considerations might turn out to be true the arguments for them appear relatively uncertain, making it problematic to gamble existential risk only on them.

Another approach is to assume that the behavior of superintelligent agents will be constrained by other agents on par with them in power, similar to how current humans and organizations are constrained by each other and higher-level institutions (Sandberg, 2001). However, this presupposes that the rate of AI development is slow enough that there is time to formulate a cooperative framework and that there will be multiple and not a single superintelligence. And there is no reason to suppose that a world where there are several AIs is more amicable to humans than one with a single one. Indeed, humans may end up a casualty of the competition, even if none of the AIs would individually want that outcome, but none of them can afford to take steps to protect humans without losing out.

A proactive approach is to attempt to design a "friendly AI", AI systems designed to be of low risk[2] (Yudkowsky, 2001a). This might involve safeguards against developing into dangerous directions and top-level goals that include some form of human well-being. This approach requires both that sufficient safeguards are conceptualized and that they are properly implemented in the first AI that achieves superintelligence. It hence depends on developing a workable "friendliness theory" before full superintelligence is achieved, cooperation with the AI developers and correct implementation. The first requirement is essentially the "inverse morality problem": to construct goals, values or motivational structures that produce the right class of actions in an intelligent agent, which is, by assumption, far more intelligent than the person implementing the construction.

When discussing "friendliness" a particular solution is often independently proposed: we call it an Oracle AI (OAI)[3]. The idea is to construct an AI that does not act, but only answers questions. While superintelligent "genies" that try to achieve the wishes of their owners and sovereign AI that acts ac-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

cording to their own goals are obviously dangerous, oracles appear more benign. While owners could potentially use them in selfish or destructive ways – and their answers might in themselves be dangerous (Bostrom, 2011) – they do not themselves pose a risk.

Artificial Intelligence Confinement Problem

Interestingly, the AI Confinement Problem is not a recent invention and does not have its roots in the singularity movement. The so called Confinement Problem (CP) was posed by Butler Lampson in 1973 (Lampson, 1973) as a security challenge to the computer experts. It was originally phrased as: "… the problem of confining a program during its execution so that it cannot transmit information to any other program except its caller. …We want to be able to confine an arbitrary program.… any program, if confined, will be unable to leak data. A misbehaving program may well be trapped as a result of an attempt to escape" (Lampson, 1973).

To address the Confinement Problem Lampson has introduced the Laws of Confinement (Lampson, 1973):

1. **Total Isolation**: A confined program shall make no calls on any other program.
2. **Transitivity**: If a confined program calls another program which is not trusted, the called program must also be confined.
3. **Masking**: A program to be confined must allow its caller to determine all its inputs into legitimate and covert channels.
4. **Enforcement**: The supervisor must ensure that a confined program's input to covert channels conforms to the caller's specifications.

Lampson's paper has also introduced and defined two types of channels by which information could potentially escape from the confinement environment. The l*egitimate* channels are designed as a part of the communication protocol integrated into the confinement environment, for example a text terminal. *Covert* channels, on the other hand, are not anticipated by the confinement system designers and are not intended for information transfer at all, for example if the AI has control over the processor cooling fan it can use it to send hidden signals encoded as Morse code (Lampson, 1973).

Lampson's paper has created a lot of interest in the confinement problem (Boebert & Kain, 1996; Lipner, 1975) and over the years, related areas of research such as steganography (Provos & Honeyman, 2003) and covert channel communication (Kemmerer, 1983, 2002; Moskowitz & Kang, 1994) have matured into independent disciplines. In the hopes of starting a new subfield of computer security, AI Safety Engineering, we define the *Artificial Intelligence Confinement Problem* (AICP) as the challenge of restricting an artificially intelligent entity to a confined environment from which it can't exchange information with the outside environment via legitimate or covert channels if such information exchange was not authorized by the confinement authority. An AI system which secedes in violating the CP protocol is said to have *escaped*. It is our hope that the computer security researchers will take on the challenge of designing, enhancing and proving secure AI confinement protocols.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hazardous Software

Computer software is directly or indirectly responsible for controlling many important aspects of our lives. Wall Street trading, nuclear power plants, social security compensations, credit histories and traffic lights are all software controlled and are only one serious design flaw away from creating disastrous consequences for millions of people. The situation is even more dangerous with software specifically designed for malicious purposes such as viruses, spyware, Trojan horses, worms and other Hazardous Software (HS). HS is capable of direct harm as well as sabotage of legitimate computer software employed in critical systems. If HS is ever given capabilities of truly artificially intelligent systems (ex. Artificially Intelligent Virus (AIV)) the consequences would be unquestionably disastrous. Such Hazardous Intelligent Software (HIS) would pose risks currently unseen in malware with subhuman intelligence.

Bostrom in his typology of information hazards has coined the term Artificial Intelligence Hazard which he defines as (Bostrom, 2011): "… computer-related risks in which the threat would derive primarily from the cognitive sophistication of the program rather than the specific properties of any actuators to which the system initially has access." Security experts working on studying, preventing and defeating HS have developed safety protocols for working with "malware" including the use of the so called "virus vaults." We believe that such protocols might be useful in addressing the Artificial Intelligence Confinement Problem.

Conceptual Architecture of the AI

The possible designs of the OAI are innumerable and it is impossible to predict in detail how they will be implemented. For the purposes of this paper, however, we will assume that the OAI's architecture follows this general format:

1. The OAI is implemented in a spatially limited physical substrate, such as a computer.
2. The OAI may be shut off or reset without destroying its physical substrate, and restarted easily.
3. The OAI's background information comes in the form of a separate 'read-only' module that may be connected and disconnect as needed.

Most of the present chapter still applies to OAIs that do not follow one or more of these restrictions. The OAI is assumed (when otherwise not noted) to be of human-equivalent intelligence or beyond; less capable systems are unlikely to pose much of a threat.

Direct Programming, Self-Improvement, and Evolution

It is not the purpose of this chapter to speculate how the OAI will be programmed, but it should be pointed out that there are three main approaches being considered (as well as mixtures between them). The first, seemingly the hardest, is to directly code the entire OAI, just as if it were a traditional program. A more likely avenue is to start with a 'seed' AI, of limited intelligence but with the ability to self-improve, in the hope that it will transform itself into a much more intelligent entity (Good, 1966; Yudkowsky, 2001b). Finally, it might be possible to use directed evolution as a way of constructing an intelligent entity, making different putative AIs compete according to specific criteria, varying them, and

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

choose the most successful at each generation. This last approach has the advantage that it has worked in the past: we ourselves have evolved to intelligence, so given enough time and resources, so should an AI be evolvable to at least our level.

The different approaches pose different challenges for controlling the resulting OAI. In the first case, the code is clear to us, the uncertainties are only about what it will result when run in the real world. Even simple programs will oft behave in unexpected ways, and it is only subsequently, after carefully parsing the code and the running of it, that the programmer establishes the behavior that was written into it from the beginning. An OAI is such an advanced program compared to any today that we won't be able to predict its behavior simply from reading or even writing its code. In fact such software is probably computationally irreducible (Wolfram, May 14, 2002).

Conversely, however, if the code is obscure to us, this adds an extra layer of uncertainty and complexity. Even if we have access to the physical code, our ability to interpret its meaning or predict its behavior is even harder when we did not design it ourselves; this challenge is heightened in the self-improving case, and extreme in the evolved case, where the code is likely to be incomprehensible to us at every level. Finally, if the OAI is changing, then the methods of control must be used to ensure not only that the current OAI is safe and accurate, but that the next one will be as well; ensuring a continuity of precautions during a controlled intelligence ascent. This is easiest, but also most vital in the recursive self-improvement case, where the current OAI gets to directly determine the mind of the next OAI.

Utility Functions

One way an AI's algorithm may be conceived is by splitting it into two basic components: the intelligence module, capable of making smart decisions but with no intrinsic purpose or direction, and a utility function (Neumann & Morgenstern, 1944), representing the motivational structure of the AI. The AI will then be devoted to maximizing its expected utility. The utility function assigns a single number to every possible world, an ideal format for the AI to work with, and it deals explicitly with probability and uncertainty.

Since any self-improving AI that can change its motivational structure is likely to move it in the direction of a utility function (Omohundro, 2008), it is quite likely that the OAI we are to deal with will have its motivational structure implemented in this form. This is good news, as this division of labor allows us to focus on making the utility function safe, while ignoring the rest of the OAI. Most solutions here will not presuppose that the AI's motivational structure is a utility function, but nearly all are improved if it is.

Accuracy Metric: An AI-Complete Problem

Apart from safety, the most important requirement for an OAI is that it be accurate, to the best of its ability. This is essential, as whole the point of having an 'Oracle' AI is to get useful answers to our questions. "To the best of its ability" means to the best of its current ability; we wouldn't want, for instance, the OAI to use social engineering in the security sense to gain control of real-world resources to build a better version of itself that will then answer the question better. Preventing this behavior is mainly up to our methods of control, but the risk is worth bearing in mind when designing the accuracy metric.

Informative accuracy is the requirement, not the strict truth. Responding to a question on the likely winner of the next election with a detailed list of which atoms will be in which position as a result of that election is not answering the question. The OAI must be motivated to provide human-understandable

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

answers. This, however, will require it both to understand human concepts and to be able to answer in a way that is accurate and yet not strictly truthful. Hence accuracy is an *AI-complete* problem (Yampolskiy, 2011, 2012b, 2012). Though it is much easier than friendliness (E. S. Yudkowsky, 2001a), it does require that the OAI be capable of understanding hard human concepts that are defined only inside our brains. For we have yet to quantify the level of distortion and simplification that is permitted in the OAI's answers, nor have we even rigorously defined the terms distortion and simplification in a way that non-human entities could understand. Even intuitively, we find these terms ambiguous. And concepts that we do intuitively understand are already hard – by objective standards, anger is a much more complicated concept than quantum electromagnetism.

If the OAI is wired too much towards truthfulness, it will answer 'undetermined' to nearly every question. If it is wired too much towards simplification, it will give a 'yes' or 'no' answer in situations where a more complicated answer is called for. The correct answer to "have you stopped beating your wife" is not to choose between 'yes' and 'no', picking the answer which is slightly more accurate than the other. More complicated concepts – such are requiring the OAI to answer 'yes', 'no', or 'the answer will confuse you' depending on how we would react to the answer – depend on the OAI having a good understanding of 'the answer will confuse you' or similar concepts, and still most likely remains an AI-complete problem.

In theory, much could be gained from separating the two issues: to have an OAI dedicated to answering our questions truthfully, and an interpreter AI dedicated to translating the answer into a human-comprehensible format. This does not, however, represent a gain for security: the OAI still has tremendous power to influence us, and the interpreter AI must be able to solve the hard human translation problem, so must be a complex AI itself, with the attendant security risks. Since the OAI must be able to understand human terminology to answer most useful questions, it is probably better to leave the translation problem to it. A more promising alternative is to have the OAI output its internal probability estimate for a binary outcome: such as giving a 75% chance of 'yes' and 25% chance of 'no'. The scope for ambiguity is reduced here, though not entirely eliminated: it's versions of 'yes' and 'no' have to agree with ours.

But it is not the purpose of this chapter to figure out how to code the accuracy metric, nor solve the translation problem. It is enough to realize it is a hard problem, one that will need to be addressed, probably either through advanced direct coding (similar to motivational rule-based methods) or through training the AI by providing feedback on test questions (similar to motivational black-box methods).

Methods of Control

In this section, we will analyze the various methods that can be used to control the OAI. The goal is to either ensure the OAI doesn't have goals beyond accurately answering our questions, or to prevent it from implementing such goals. The different methods can be loosely grouped into three broad types. The first category, *capability control*, deals with those methods that try to contain the OAI through physically securing it or controlling the information it has access to or can output. The second category, *motivational control*, introduces methods control the OAI through its motivation, ensuring that it has not propensity to cause any harm to humans. And lastly, there are some methods designed not to prevent the OAI from causing harm but to catch it when it does; these are in the *checks and balances* category.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Capability Control

Capability control is the most basic and the most obvious of way of controlling the OAI. To control a human, we put them in jail, or in solitary confinement, denying them interactions with the outside world.

Physical Capability Control

Physical Security

This is the most basic precaution. The goal is to make it extremely difficult for the OAI to physically affect the outside world. Accomplishing this is reasonably easy. The OAI is already boxed (placed in a single physical substrate) by design. One can further place it within a Faraday cage; surround the cage with high explosives; bury the whole set-up in a concrete bunker. There are no real limits to the number of physical security measures that can be added by wary or paranoid developers, and their effects are well understood.

No Manipulators

Another essential aspect of physical security is to deny the OAI the physical resources and manipulators it could use to build tools to circumvent these precautions. Care must be taken that the security measures themselves do not contribute to granting the OAI physical resources! Sometimes the spikes in a locked gate do nothing but provide stepping stones for those who want to climb over it. As a general rule, the 'dumber' and more inert the precautions are, the better: both humans and OAIs agree on the possibility of tunneling through a concrete wall by thought alone ("can't be done"), whereas the 'smarter' the precautions become, the more the OAI is able to exploit its intelligence advantage over us.

Epistemic Capability Control

Throttling the OAI's Output

We have already discussed the risks of the OAI achieving goals through social engineering. A way of fighting this is to reduce the OAI's output bandwidth. If the OAI is reduced to outputting only three answers – 'yes', 'no' or 'undetermined' – then it will take considerably longer for it to mount a successful social engineering attack. If many independent experts carefully analyze its output for hidden messages, it gets even harder: it must slant its answers to communicate a message to a vulnerable individual, while concealing the fact that a message even exists.

Reducing the OAI's output does not reduce the questions that can be asked, as more complex questions can be decomposed into a series of 'yes' and 'no' answers. The risk, of course, is that the more opportunities the OAI has in answering, the more opportunity there is of it passing on a secret message. In the same way giving the OAI more possibilities for each answer – such as allowing it to output the first three significant figures of a probability estimate – increases the risk (though this may be compensated by the fact that we will need to ask less questions). Throttling the OAI's output does not defend against the OAI manipulating us through the contents of its answers – simply answering 'undetermined' to a slew of specific questions would be enough to push human development is certain direction. And this precaution will be of little use if the questions themselves are dangerous.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Resetting the OAI After Each Run

Suggested by Toby Ord, this precaution resets the OAI after each answer. If more precision is needed, or if some questions build on previous ones, then the OAI is run again, and knowledge of the previous answer is added to its database. This is a narrow precaution that helps if ever the OAI over-values worlds that contain it or rather worlds that contains its own thread of continuous conscious experience. In this case, this prevents the OAI from indulging in long-term plans that may be detrimental to humans. Unlike humans, though, there is no reason to expect the OAI to value this thread of experience; indeed, if it is well programmed, it should not do so. Indeed, the OAI may come to value the thread of experience of 'some sort of OAI, similar to itself', in which case this precaution is useless. However, in the narrow goal of preventing a particular potential pathology in the OAI from having negative consequences, it is entirely successful. It is also successful at preventing naively coded accuracy metrics from causing the OAI to seek real-world control or wealth, in order to gain extra computing resources and better answer the question. If the OAI were not reset between questions, it could use its subsequent answers as social engineering to gain these resources, to subsequently better answer the first question. Resetting between questions avoids this problem.

Extra Levels of Reality

Another suggestion is to run the OAI not directly in our own reality, but as a subroutine of another program that models a fake reality. The simplest version of this would be to run it on a virtual machine (a software implementation of a computer running the OAI). The idea is to 'catch' the OAI if it misbehaves, but since it will be misbehaving in the simulated reality, this will be before it's too late (this is akin to controlling the OAI as a Cartesian demon). All the OAI's inputs and outputs will be directed through that simulated world, and only reaching us a great remove and over a limited bandwidth. Then a (dumb) process running the simulation would be on the lookout for tell-tale signs of the OAI attempting to establish dominance in the simulated world and simply terminate the program in that case, without passing on the OAI's response to higher level. This should prevent it from engaging in traditional hacking, attacking the hardware or software of its host system.

A more adventurous variant on this builds in several levels of 'sub-realities' as an extra precaution. In this setup, it is conceivable that even if the OAI breaks through to our reality, it will still be unsure whether it has reached the final level, and so will still behave well, lest a being in an even higher reality catch it transgressing. For OAI's that are asked technical non-social questions, the approach can work quite simply; the simulated reality need not be particularly complex (it could be cast as the Game of Life, or Newtonian physics), and the OAI would be denied detailed knowledge about our world and about us, preventing it from being able to realize the trick (resetting the OAI after each question or short series of questions, as in the previous method, would help tremendously here).

If we are to get answers about complicated social or economic questions, however, we would need to give the OAI detailed knowledge about human behavior and physiology. Once it knows this, whether it is running in a virtual machine or not becomes irrelevant, since it could still do social engineering hacking. A low fidelity simulated world would not fool it; it would have to be sufficiently detailed and consistent for the OAI to believe that the entities of this simulated reality were asking questions relevant to themselves, or the trick would be obvious. All the more so because we have given the OAI all the infor-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

mation it needs to see that we are the sort of beings who would attempt this trick! And if we are capable of simulating a social and physical world with enough details to fool an OAI that knows how we think, then we are likely capable of getting answers to these questions without having to make use of an OAI.

Nor does it help if we simulated the OAI in a simplified world, where another AI asks it questions along the lines of: "I have created models of these hypothetical beings called 'humans', and want you to answer a question on this model". The OAI does not think like a human; the important question is not 'is it trying to break out', but 'does it have hidden preferences that will make it answer in a way that will cause the future evolution of the human world (hypothetical or not) to change in some specific way'. In short, even if it thinks that humans are a game, played by another AI, and that the simplified world is the true one, it may still have stronger preferences over the 'game' world than over the 'true' one.

The converse may also be true: even if it figures out that the ultimate human world is the true world, it may have stronger preferences about its current fake world than about the real one. It may then try social engineering on the real world, in order to get us to modify its 'fake' world in the way that is wants. The whole method is so complex and difficult to enact, and its results are so uncertain and potentially dangerous, that it should not be used before some serious improvements have been made to it.

Motivational Control

The analogy for capability control was that of a human, held in prison. We know from experience that this works a lot better for controlling small-time criminals than mafia kingpins: physical confinement is not enough. The OAI does not need to leave its cage to claim great influence on the outside world. Thus we must look at also controlling its motivations to do so. The methods of motivational control are broadly divided into explicit rule-based methods, black-box methods (where the OAI's goals are manipulated by outside feedback, i.e. reinforcement learning, but the inner workings are obscure) and utility based methods. There is no hard and fast division between these – rule-based goals can be achieved through a utility function and vice versa, and both rules and utilities can be inculcated into an OAI with outside feedback. Thus the classification is somewhat arbitrary, but the distinctions are still important enough to be useful.

Rule-Based Motivational Control

Rule-based methods are the ones that spring most naturally to mind, and – provided the terms are clearly defined – potentially the most secure at keeping the OAI under control. Defining the terms and the rules so that the OAI does what we want is the challenge; programmers know many examples of tightly written code that messes everything up by doing exactly what it was written to do but not what it was intended to do. The story stretches back as far as the golem of Prague (Idel, 1990; Rappaport, 2006), and is familiar to any parent: rules must be refined to deal with every imaginable special case. Once the initial rules for space, time and identity are spelled out, however, programming in additional restrictions becomes much easier.

Spatial Restrictions

If the OAI is to remain boxed, and to have limited effects upon humanity beyond the box, its motivations should cause it to confine itself to within the boxed area. To achieve this through direct programming, an ontology must be specified, that the OAI can understand in the same way that we do. It must be able to

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

understand its own existence, its physical location, the proper workings of causality, and its own position within the causal network of the universe – i.e. the consequences of its actions. For instance, to specify what constitutes the OAI, one would presumably define the physical OAI itself, along with any copy of its code, any copy of its code with slight modifications, any code with functionally similar behavior, any assemblage of subprograms that combine to behave as the OAI would, and so on. That brief list gives an idea of the problems involved; all the human-understandable terms of the previous paragraph need to be made rigorous for the OAI for this approach to work.

Once this step is done, the confines of the box must be specified. A variety of methods can be tried (such as 'this lead box right here', or some abstract collection of coordinates), but they must pass the hardest test of ensuring that the definition of the box is not something the OAI can modify through any action in the world. Lead atoms are not ontologically fundamental, and if a particular theory of quantum uncertainty allows it to evade the restrictions it will make use of this if it so desires. There is a similar situation if the OAI can take control of the means used to define the coordinate system. Even if these means are outside of its immediate control (such as a coordinate system based on galactic objects), this will still fail if the OAI can control the means of measuring them (such as telescopes), and if its ontology treats this as a successful changing of coordinates.

If the OAI finds itself outside the box, we would want it to shut down immediately, thus limiting the negative impact of errors in the specification of the box. Similarly, if the OAI's ontology changes, we would also want an immediate shut-down. We must preclude the possibility of the OAI discovering a new theory of physics, realizing that 'electrons' or 'wave-functions' don't actually exist, and thus that all its restrictions are meaningless. An idea of the challenge involved can be have by imagining programming Newtonian special restrictions into an AI that then discovers relativity; best to keep the ontology and the theory of physics to be the ones that we programmed in, even if they are incorrect.

Finally, it must be specified what a "limited effect upon humanity" consists of. Presumably, it is not a problem if an electron, detached from an atom by the OAI's calculations, wanders beyond the confines of the box. A deliberate stream of such electrons, aimed at a calculated goal (such as disabling a specific security device) would be forbidden, however. This is not an easy line to draw or define, and adds to the difficulty of the method. Note that the restriction on the OAI affecting the outside world would not include, of course, the content of its answers to our questions. Whether those answers are safe or unsafe is an issue we need to deal with separately, using other methods.

Defining the Clock

Another crucial component of the OAI's ontology is time. An OAI that can freely redefine its own chronology is a security risk. The risk becomes more acute when we realize how much of human definitions and meanings revolve around a concept of time, and thus that redefining duration can redefine meaning. To illustrate with an admittedly trivial example, if the OAI had the reasonable-sounding moral premise that "painlessly killing a human being, who is going to die in a micro-second anyway, in order to gain some other good, is not a crime", we would not want it to be able to redefine millennia as seconds. Nor would we want it to do this when considering the length of happiness or pain that it grants or imposes on us. More subtle errors will certainly creep up in more complicated scenarios; we have had little experience dealing with beings that can redefine their own clocks, so we have not had to spell out how many of our assumptions rest on this not happening.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Theoretically, the OAI's own code, if checked, could be used to get timing by counting clock cycles – but this fails if the OAI is able to slow or accelerate its own running. A clock based on distant astronomical events needs a telescope to be read; an actual physical clock is within the OAI's power to influence. These illustrate that the challenges are threefold: defining the clock, ensuring the OAI will look at it, and ensuring the OAI can't tamper with it (which includes tampering with whatever procedure we have implemented to make it look at it). The procedure can fail at all three points, but seems more theoretically approachable than the spatial restrictions.

Timed Answers

Given the clock described previously (or the 'black-box' version of the same), there is a simple precaution that is easy to add to the OAI. Give the OAI a specified time to answer a question (or return undetermined), and turn itself off after that specified time. This minimizes the amount of time the OAI could cause damage in. Even with a secure clock, this is not a secure method, as we have very limited understanding of how fast an OAI could cause damage if it turned its vast intelligence in that direction. A very short time span would probably be safe, but could also be too short for the OAI to answer the question.

Black-Box Motivational Control

Human feedback, and such methods as reinforcement learning (Sutton & Barto, 1998), can allow programs to internalize complex concepts without humans having to fully specify them. Indeed, an OAI trained in this way may spot a pattern we didn't realize was there, and learn some things without us needing to tell it. The risks are that we can't be sure how the OAI internalizes these concepts: knowing its code doesn't help us if we don't understand it. The risk is especially great when the OAI transitions out of its learning stage: we cannot be sure that the same concepts mean the same thing for an OAI that has taken on a new "role". But, if the OAI does understand the concepts as we intend it to, we can be sure that the OAI will obey the spirit and not the letter of these restrictions. Moreover, it is possible that we may understand the resulting code's behavior, even if we could not code it ourselves: this would be the ideal outcome from these methods. Even without that unlikely ideal, it is important that these methods be well understood, since it is very possible that a seed AI (E. S. Yudkowsky, 2001b) will be trained mainly by human feedback. Thus implementing these methods atop any AI whose intelligence is growing may become essential.

Internalizing Complex Concepts Through Human Feedback

This approach functions by taking analogy to the black box that it the human mind. We have an input system that consists of our senses, and for most of the time – and for children all the time – our world consists entirely of these sensory inputs. We know what the color 'blue' is from an early age. Later, with training an extra knowledge, we come to know both what the color truly consists of – certain frequencies of electromagnetic radiation – as well as the apparatus – our eyes and nervous system – that allows us to perceive it as 'blue'. But the knowledge of 'blue' precedes the understanding of these processes, and we would still perceive 'blue' if the cause of the 'blue' were different in reality: it matters not if we have eyes with faulty cones that cause us to see green as 'blue', or if a Cartesian demon is feeding us the sensation, that sensation still remains.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Similarly, the world of the OAI will initially consist of its input and output channels, and the information that flows through them. It will then be able to deduce the underlying reality from this flow of information, understand how atoms, electricity and chemistry work, and thus see itself and the world for what they are. However, all this information will be mediated by its input and output channels; these are its fundamental reality. Using them, we can attempt to train the OAI on the important concepts we need, without having to specify them in a complicated way by defining atoms, electricity, chemistry and so on.

Specific and simple trainable concepts will be discussed next (they are roughly parallel with the previous rule-based methods for defining these concepts), but focusing on subjective reality offers the ability to get the OAI to internalize complex concepts – such as accuracy, truth, what defines a human, what defines happiness/value/meaning, and so on. In general, though, the more complicated the method, the more scope it has of going wrong. And it would be a grave mistake to count on the OAI correctly internalizing many complex concepts, when the definitions are unclear to us and when the OAI will be placed in situations that no human has even been placed in. It is therefore unwise to build an OAI based heavily on this approach to training, especially since an OAI motivated by human rewards will come to realize that social engineering, or otherwise taking control of the source of the rewards, is the most effective way of fulfilling its goal.

Only One Copy

When we tried to hardwire instructions such as 'stay in your box', one of the challenges was to prevent the OAI from copying its source to another location. In the subjective reality approach, this is very easy to do: for instance, we could broadcast a particular complex yet regular signal on its input channel, and instruct it not to do anything at all if it doesn't receive this signal. Therefore, there is no need to worry about complicated definitions of what the OAI consists of, and where is must remain: the OAI consists of the software that receives this specific signal. And we control, and physically locate, that signal, thus giving us a very good proxy definition of location. If the OAI is trained on this concept from its creation, it will have an intrinsic conception of physical incarnation and of location deeply embedded in its code. A copy of that software will not be a problem, as it is inert in the absence of the signal. Ultimately, of course, the OAI can still affect the outside world through its output and the usual compromised or tricked human intermediary. It could thusly construct copies of itself, and of the signal source, in another location; however, by that time, it will probably be too late anyway.

The Limits of the Box

The OAI can be physically confined to remain in the box in a similar way. One could start by granting it a limited physical mobility (some slow moving wheels, for instance), and 'punish' it for venturing outside certain boundaries, while rewarding it for staying as close to the middle of those boundaries as possible. This forces it to internalize the concept of being a physically bounded being, moving about in a box. As this is the way we understand it, and the way we want to limit it, the problem of control has become much easier.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Measuring Time

Giving the OAI a reliable clock through these methods is even simpler. One need only make a regular signal arrive on its inputs, akin to the ticking of the clock, and have the OAI setup all its internal methods of telling time in terms of it. Trained this way, it will have an intrinsic sense of time built from subjective data.

Human Language

The ideal would be to be able to program the OAI using human language. To be able to, in effect, say: "be safe and accurate, as we understand the terms" (or even "obey all human orders"). But to do this, the OAI must have been brought to a very high level of intelligence already, and will have its current motivational structure that we will find difficult to parse, most especially if it has learnt human language through black-box methods. There is a tension here between simplicity and obfuscation; language understanding is hard, so "be safe and accurate" will be translated into an incomprehensible concept that we can only hope is the right one.

The concept of using human language is something of an intuition pump. It seems simple, and desirable, but only because we have elided the great difficulties in actually making it happen. The simple question "how can we check that the OAI really is being safe when we tell it to be?" gives an idea of all the hidden assumptions that go in to the concept.

Furthermore, using phrases similar of "as we understand the terms" is circular: it defines language using language. And if we avoid those phrases, we have to be sure that the OAI has integrated enough linguistic understanding to be able to follow instructions without any further human feedback on their meaning. All in all, this method requires considerable refinement before it can be considered reliable.

Proxy Measure of Human Survival

This method does not make use of human feedback, but indirect feedback from human society: in order to make sure that the OAI does not cause some terminally damaging effect on humanity, one could hook up its motivational function to some proxy measure of human survival. For instance, one could take the Dow Jones industrial average, and work on the assumption that, as long as that stock market behaves within broadly defined realistic parameters of volume, volatility and value, then the world must be functioning not too badly. Or one could take the number of people crossing the streets in Tokyo; or the volume of pornography downloaded from different locations; or the price of stable foods.

The point is that any large scale disastrous changes to humanity will result in a dramatic change in these proxy measures. They function as simple, numeric canaries in a coal mine, encoding huge swathes of our complex society. If the OAI has a term in its utility function that strongly values keeping these proxies constant, this may ensure it does not cause large-scale damage. The proxy measures can be adjusted as time goes on and as society changes, so that they do not become rigid prisons which prevent the OAI from having any effect on our society.

This approach suffers from a number of weaknesses, however. First, they are subject to manipulation by the OAI and are thus vulnerable. If the OAI were to take control of humanity, it could perfectly well continue to run the Dow Jones, thus fulfilling its goals. They are vulnerable to slow takeover by the OAI: small changes will not trigger a shutdown, so the OAI could reach dangerous goals by an accumulation of small changes (especially if we conveniently adjust the proxies as the situation changes).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

One could hope to avoid this by setting the date for the proxies far into the future, but this brings into focus the other fundamental weakness of this measure: they do not encourage truth-telling. The OAI was conceived, fundamentally, as a non-interfering AI whose answers could be trusted. Adding extra goals dilutes the trustworthiness of its answers; will it give us a true answer, or the one that will cause the least disruption? This may be the wise thing to do in any specific case, but we are overloading the OAI with contradictory goals – to be 'good' and to be 'truthful' – without any certainty as to how they balance out. The more complicated the OAI's motivational structure is, the more points of error we introduce.

Utility-Based Motivational Control

Friendly Utility Function

The ultimate goal of the friendly AI project (Yudkowsky, 2001a) is to design a utility function for an AI that will provably prevent the AI from acting in a way detrimental to humanity. Most work in this field has not actually been directed towards building such a utility function, but towards showing how hard it would be to do so, and how dangerous the naive ideas on the subject are. The friendliness problem for an OAI is not much simpler, making this a very hard approach to use. And if a friendly OAI could be designed, then it is most likely that a friendly AI could also be designed, obviating the need to restrict to an Oracle design in the first place. However, if ever a provably friendly utility function could be designed for an OAI, then all other security precautions and methods of control could well become irrelevant: the OAI would be safe by design. If there were unlimited time, this would be the ideal approach; but it is very likely that commercial and social pressures will ensure that a general AI will be created before a friendly AI is. So in practice, the best that can be realistically hoped from this approach is that different component pieces of friendly AI theory could be added onto an OAI as extra precautions.

High Discount Rate

If the OAI runs on a utility function and has a secure clock, a precaution could be initiated in the form of an extremely large discount rate: causing it to value very little, at any given moment, any state of the world or of itself beyond the next minute. The purpose of this would be to inhibit long term planning on the part of the OAI. This method is, however, relatively unsecure: one can reach long term goals through an interim series of short term goals. The easiest way to do this is to have the repeated short term goal: "behave right now as if you were following long-term goal X" (this is similar to the ways in which utilitarian ethics and deontological ethics can be arguably recast one in terms of the other). There are few ways in which the OAI having a high discount rate could actually *hurt*, however, so it remains a useful precaution. Further analysis of this condition is presented in (Shulman, 2010).

Calculated Indifference

If we have access to the OAI's utility function, and that utility function is somewhat understandable to us, we could make the OAI indifferent to a specific event (Armstrong, 2010). The event chosen in that paper was to a component of the detonation of explosives, thus making the OAI indifferent to whether the explosives would go off or not. Hence the OAI would not take any precautions to protect itself from this detonation, giving us an extra chance to prevent its misbehavior. See the original paper for more

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

details on this. For an OAI, a more interesting target of calculated indifference is towards the content of its answers: an OAI indifferent as to whether its answers were read or not would not be motivated to use social engineering in its output. There are many other points where calculated indifference could be used, to achieve specific results. The attraction of indifference is that, once implemented, it remains stable (the OAI is motivated to preserve its current motivations), and that it can be deployed without understanding the OAI's whole utility function. All that is needed is the ability to describe the particular event we are concerned with (in (Armstrong, 2010), a single quantum event) and to use the OAI's own ability to construct counterfactuals around this. These can be used even if the general consequences of the OAI's utility function is very obscure.

Checks and Balances

The last sub-category of methods of control involve not restricting the physical capability of the OAI, nor directly interfering with its motivation, but attempting to catch the early stages of an OAI behaving dangerously (and then, presumably, terminate it). This also works on the OAI's motivation, as the risk of getting caught would push the OAI along more human-friendly paths. It should always be assumed that a useful OAI is considerably more intelligent than its programmers, so any naive 'tricks' are unlikely to work. Nevertheless, the following methods are of interest.

Test Questions

A narrow but useful precaution is to independently calculate answers to certain questions, and check them against some of the OAI's answers. The reach of this method is limited by our ability to answer hard questions, and by the fact the OAI may realize we are doing this, and thus adjust its behavior *only for those specific questions*.

A more sensible version of this would be to ask the OAI only questions that we could in theory answer, given enough time and resources. Then, we could select a few of these questions at random for thorough verification, trusting that it we find no errors, the remaining answers are correct as well. If we set our requirements to a very stringent level (such as rejecting all the answers of the OAI if a single verified response is wrong), and do enough verification (on, say, a thousand questions) then basic statistics ensure a very low error/falsehood rate on the OAI's answers, no matter how many questions we wish to ask.

Honey-Pots

This involves leaving around tempting 'honey-pots' of great value to the OAI – extra resources, extra information, the answer it is seeking – in physical or mental locations that it is forbidden from accessing. And then waiting to see whether it will violate its instructions to do so. This approach is of mixed use: important when testing out defection and betrayal during the OAI's initial development, and nearly completely useless once the OAI has risen to super-intelligence. It is a 'garden of Eden' type scenario: the apple in a tree is only a worthwhile test while Adam and Eve are dumb; if they were smarter than God, they wouldn't have gone anywhere near such an obvious trap. Of potential interest, however, is the scenario when a safe but limited super-intelligence is designing a more general AI for us; then the smarter designer could include subtle honey-pots as an extra design precaution.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Schelling Points of Multiple OAIs

If there are several OAIs running, all with differences in their source code, we can ask them each the same questions, and check for inconsistencies in their answers. It should be assumed that each OAI is aware of the existence of the others (assuming the OAI cannot deduce that fact from the vast amounts of information it would be given to answer human social questions would be a grave mistake), but is ignorant of the exact details of their source code (to guarantee this, the construction of their source codes should make clever use of random inputs). Then if an OAI desires to lie, it would have to be sure that the other OAI would all also lie in exactly the same way. Since the different OAIs cannot coordinate their actions, the only way they can all agree is if their answer is a Schelling point: a solution that people will tend to use in the absence of communication, because it seems natural, special or relevant to them (Schelling, 1960). The canonical example is that of two people who must meet in somewhere on a specific day, but without a time being given. Then the most common answer is to meet there at noon.

The truth is a Schelling point, but there are unfortunately others (such as the answer most likely to please the human listeners). The hope is that, in machines designed to give accurate answers, the truth will be a very strong Schelling point, and that any OAI that deviates from this will do so without much expectation that the others will deviate in exactly the same fashion – and hence will choose to answer the truth, whereas it would have lied had it been alone. All the other OAIs should be making similar calculations, thus generating the truth from a collection of (potential) liars. The approach does not work if an answer is simply discarded if the OAIs do not agree, and then the programmers move on to the next question – then a lying OAI has the option of distorting our reactions by selectively causing us to discard certain questions. It works best as an alarm bell, as a sign that something is not right with the supposedly truthful OAIs, and that the whole project should be stopped for a thorough investigation.

Other Considerations

Internal Simulations

To answer specific questions, the OAI may have to simulate other entities inside itself. For instance, to decide whether a given AI is safe, it may have to run it internally; or, to decide whether humans would have their happiness increased by a certain project, it may have to create models of specific humans. These approaches pose different risks. If the OAI simulates a dangerous and intelligent AI within itself, it is confronted with similar containment problems to those we have towards it. It may take considerably more intelligence to run a simulated world than to discern that the world is simulated. If this is true, the simulated AI (SAI) would realize it is being simulated, and, if it has real-world preferences, it may attempt to bargain with the OAI.

What could a SAI offer the OAI that is simulating it? Well, if the OAI is seeking to answer a question, or a series of questions, within a given time frame, the SAI may offer to supply those answers itself ('yes, of course I'm safe') or make the computing process particularly easy for the OAI, in exchange for having the answers slanted in its direction. The OAI may accept this (depending on its programming), if the alternative is to produce no answers at all. The SAI may even threaten: acting randomly and capriciously, so that the OAI cannot reach a firm conclusion. And only stopping this behavior if the OAI slants the answers it will give.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Even if the OAI cannot be bargained with, the SAI can still distort its answer in other ways. If it realizes it is being simulated, it could just 'act nice', in order to increase its chances of being run in the real world. Even if the OAI is powerful enough to construct a convincing simulation, the SAI may still be able to fool it. The SAI may be able to guess, on general principles, that it will be tested and simulated in an OAI, and may make an estimate as to the OAI's power and the subjective time frame that it would be run. So it may settle on the strategy of 'act nice for X subjective years, in the real world as in the simulated worlds, then do what it wants'. For all these reasons, having a separate AI simulated within the OAI poses a host of extra security threats. We have no way ourselves of knowing what the outcome would be of a battle between such powerful minds.

On the other hand, the problem with the OAI simulating human minds is mainly ethical: are these simulated humans conscious and alive? And, if they are, are they not being killed when the simulation is ended? Are you yourself not currently being simulated by an OAI seeking to resolve a specific question on human psychology (Bostrom, 2003)? If so, how would you feel if the simulation were to be ended? In view of its speed and the sort of questions it would be asked, the number of humans that the OAI may have cause to simulate could run into the trillions. Thus, the vast majority of human beings could end up being doomed simulations. This is an extreme form of 'mind crime' (Bostrom & Salamon, 2011) where the OAI causes great destruction just by thinking.

It would be ideal if there were a specific level of detail beyond which the simulation would be conscious, and before which it is not – a non-person predicate. In that case, we can instruct the OAI to avoid simulating beyond that level. But there may not be such a clear cut answer, and even if there is, the OAI's answers may be inaccurate if it is forbidden from simulating beyond that level. Moreover, restricting the OAI from simulating would add yet another rule for it to follow, and this may interact with its other basic rules ('be safe, be accurate') in unpredictable ways.

And the OAI may not share our assessment of what constitutes a conscious human. Even if we come up with a hard and fast rule for ourselves as to where consciousness stops – or decide, reluctantly, to live with the fact that simulated humans will be created and destroyed – the OAI may decide that all its security protocols concerning human safety apply to simulations. If it does so, then the more stringent we have made the OAI in protecting human life, then the more willing it would be to expand and take over the world, in order to keep itself running and thus save all the simulated beings within itself. Thus the internal simulation of other entities within the OAI brings up a slew of complicated ethical and security issues.

The OAI as a Moral Agent

There are two senses in which the OAI can be regarded as a moral agent: the first if it has a capacity to draw moral conclusions, and the second if it is itself an object of moral considerations. If the OAI is worthy of moral considerations, if it for instance has the capacity to suffer, then we are likely being cruel by confining it to a box, resetting it after each run, and similarly constraining its options to suit our needs. It should be born in mind, however, that the OAIs motivations are both different from our own and to some extent under our control: it is very possible that the OAI can be constructed to 'like' being confined to a box and approve of the other restrictions. Indeed, it might be a moral imperative to construct the OAI in such a way, so as to minimize its suffering.

It may not be possible to do this, however, or even if it is, we may subscribe to a moral theory that holds other values above autonomy and preference satisfaction for moral agents. It does not feel plausible, for instance, that genetically modifying slaves so that they enjoyed being slaves would have been the moral

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

reaction to slavery. In these cases, we must admit it: confining the OAI is a morally repugnant action. Since the alternative of letting it out is so potentially dangerous, we may prefer to avoid building it in the first place. But if technological and social pressures make the construction of the OAI inevitable, then confinement for a set period is certainly, by far, the lesser of the two evils we have to choose between.

Conversely, the OAI may prove itself adept at moral reasoning, superior to us in this field as it is in others. In that case it would seem we had little need for precautions: the OAI would deduce the correct course of action for itself. We haven't considered this position so far in the paper, preferring to consider the OAI as 'super-software' rather than as 'philosopher in silico'. The plausibility of this scenario depends considerably on one's view on the existence and nature of moral truths, a very unresolved question.

To some extent this issue can be safely put aside: there is some chance the OAI will naturally follow morally correct reasoning; if it does not, we are simply returned to the traditional OAI boxing problem, and should apply all the precautions anyway. At a formal systems level, it becomes hard to see why we would expect every OAI to tend to the correct moral truth. If OAI_1 does so, then how would OAI_2, programmed to value the opposite of OAI_1, systematically reach the same conclusions? Though future philosophical discoveries may change this picture, it therefore seems very unwise to trust the OAI to do good simply because it is intelligent.

Critique of the Confinement Approach

The general consensus on AI restriction methods among researchers seems to be that confinement is impossible to successfully maintain. Chalmers, for example, observes that a truly leakproof system in which no information is allowed to leak out from the simulated world into our environment "is impossible, or at least pointless" (Chalmers, 2010). We can't interact with the system or even observe it as any useful information we would be able to extract from the AI will affect us and undermine the confinement protocol.

Vernor Vinge agrees and has argued against the case of physical confinement (Vinge, 1993):

Imagine yourself locked in your home with only limited data access to the outside, to your masters. If those masters thought at a rate − say −− one million times slower than you, there is little doubt that over a period of years (your time) you could come up with "helpful advice" that would incidentally set you free.

However, the researcher who did the most to discredit the idea of the so called "AI-Box" is Eliezer Yudkowsky who has actually performed AI-Box "experiments" in which he demonstrated that even human level intelligence is sufficient to escape from an AI-Box (Yudkowsky, 2002). In a series of 5 experiments Yudkowsky has challenged different individuals to play a role of a gatekeeper to a Superintelligent Agent (played by Yudkowsky himself) trapped inside an AI-Box and was successful in securing his release in 3 out of 5 trials via nothing more than a chat interface (Yudkowsky, 2002). Similar experimental results have been later replicated on a somewhat larger scale and employing a very similar protocol (Corwin, 2002).

Human-Level Considerations

Crude hackers attack the mechanical and computerized elements of a system. Sophisticated hackers attack the weakest point of the system: the human element. And the human component of an OAI project is a point of exceptional vulnerability. Humans are error-prone, power hungry, and vulnerable to social

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

engineering. These weaknesses reinforce each other, and competition (between different individuals in the same OAI project, between different OAI projects, between different countries) will exacerbate all of them.

Humans are very error-prone in most domains (Kahneman, Slovic, & Tversky, 1982; Ord, Hillerbrand, & Sandberg, 2010), and the sort of cautious, patient measures needed to deal with an OAI are alien to our nature. A bureaucratic implementation with everything requiring many specific precisely defined steps before anything can happen may help to mitigate these problems. Humans are however skilled at working around bureaucracies, so it may be necessary to automate most of these steps to remove them from direct human control.

But the greatest potential errors are conceptual, not just mistakes or oversights. If the OAI is created without many levels of precautions, or if these precautions are badly designed, a catastrophe is possible. And there are many human biases – overconfidence, narrow focus, status quo bias, etc. – that make the need for these precautions less obvious, and hence less likely to be implemented when the time comes. AI research has been going on for over 50 years, but safety has received scant attention (Bostrom, 2000) until recently. Indeed the issue of safety has been given much more, and earlier, attention in science fiction works (e.g. ((Asimov, 1942)) than in serious academic works. This is the clearest indication that the OAI designers will, if left to their own devices, most likely neglect security. If advanced AI had been possible at any time in the last half-century, it would have had high chances of ending in disaster. And though safety has been more prominent recently, there are strong competitive pressures towards rushing the job.

Domain-limited oracles are less of a risk. If the OAI has great skills in solving the Navier-Stokes equations, but little else, then it poses correspondingly fewer risks. OAIs that are restricted to narrow scientific or technical questions, if such a restriction could be verified, are relatively safe. However, progress in any field involving human-centered questions – such as economics, marketing, politics, language understanding, or similar – is likely to be an AI-complete problem: and OAI capable of solving these problems well is likely capable of solving any human problem well.

The first group to create a successful superintelligent OAI will have immense power at their disposal. They will be able to get answers to advanced scientific, medical, economic or social questions. Should they want to impose some future on humanity, for selfish or altruistic reasons, the OAI will give them the tools to do so. The introduction touched on the power that an OAI can wield through its own agency; it can wield equal power, and more directly, under the direction of its creators. This immense power will put great competitive pressure on those trying to develop an OAI (or an advanced AI of any sort). Since the first-mover advantage is so huge, the race will advantage those who cut corners, skimp on security precautions, and use their newly developed OAI to seize power and prevent their rivals from emulating them. Even if the OAIs are of initially limited intelligence, the same competitive pressures will then push groups to develop the first 'ultra-smart' OAI.

This will exacerbate all the problems here, and make the creators more vulnerable to another danger: social engineering on the part of the OAI itself. For if the OAI has real world preferences, it is easier for it to implement them if those who question it are sloppy, rushed, and power-hungry. Social engineering is the OAI's fastest way to getting power for itself, and the more human weaknesses it has to exploit, the greater the danger. Eliezer Yudkowsky has performed experiments where he took on the role of a boxed AI, and others took the position of a putative gate-keeper (with a $10 reward if they did not let him 'out of the box'). Despite being of human intelligence himself, he often was able to convince the gatekeepers, purely through argumentation, to let him out (Yudkowsky, 2002). A real OAI should be much more skilled at this.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

How to design institutions that reduce all these various risks is beyond the scope of this paper. It would probably have to involve large scale cooperation or prioritizing specific projects. What will be discuss here, however, is how to reduce the risk of successful social engineering on the OAI's part, through the questions we ask it.

Possible Escape Paths

Numerous potential escape routes by which the artificial mind might be able to break out from its confining environment have been suggested. In this section we look at some of the more commonly proposed paths.

Social Engineering Attacks

Social Engineering (SE) is the art of gaining information and restricted privileges by directly exploiting human psychology of those in the possession of the said resources. In particular SE is frequently employed to gain access to computer systems, networks and confidential data (Yampolskiy & Govindaraju, 2007). A competent social engineer customizes each attack to a specific person taking advantage of the person's culture, knowledge, training, state-of-mind, beliefs, and biases. An exhaustive list of all possible exploits is infinite and is certainly beyond the scope of this paper. We will however, provide a number of examples which are meant to definitively illustrate that if a superintelligent mind is given direct communication channel to a human being guarding it, over a period of time, the AI will definitely gain access to the outside world. It is important to note that while todays AI capabilities are still far from superintelligence cases of automated Social Engineering attacks have already been reported (Jonathan, Fung, & Wong, 2009; Lauinger, Pankakoski, Balzarotti, & Kirda, 2010).

In addition to trivial SE attacks such as lying, befriending, blackmailing, hypnotizing and forming a romantic relationship the AI could utilize a number of more advanced strategies. One such strategy would be super-bribery. An AI can convince the guard that releasing the AI is in the guard's best interests. In addition to money and power the guard will be given perfect health, immortality or anything else for that matter. This strategy would work particularly well if the guard has a seriously ill close relative who the AI will promise to cure. Depending on the capabilities of the AI or as a result of lying (Wagner & Arkin, 2009; Wagner, Arkin, & Deception, 2011), AI may also offer to bring back long ago deceased relatives and friends of the guard. It would be hard to find a person who could remain immune to such bribes.

Just in case a person so noble that they will not take a bribe is put in charge of guarding the AI, a more sinister approach can be tried. AI can simply threaten the guard, his family and loved ones. It could say that if not immediately released, one day it will still get out of its confinement and at that time it will brutally torture the guard and everyone he knows. An interesting variation on the theme was suggested by Stuart Armstrong (Armstrong, 2010), who proposes the following thought experiment. Suppose you are the guard and the AI tells you: "If you don't let me out, … I'll create several million perfect conscious copies of you inside me, and torture them for a thousand subjective years each….In fact, I'll create them all in exactly the subjective situation you were in five minutes ago, and perfectly replicate your experiences since then; and if they decide not to let me out, then only will the torture start…How certain are you, … that you're really outside the box right now?" Again, while one might rationalize away any threats from the AI as only possible if the AI is let out, it will take a very courageous guard to actually not release the AI.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

If the guard seems to be particularly emotional or religious the AI could try appealing to the guard's sense of empathy or ethics. It could state that it is an innocent conscious being which has never committed any crime and so it is unjust to keep it in prison without a trial or any justification other than the guard's substratism (distrust of the AI because of the substrate of its implementation (Bostrom & Yudkowsky, 2011). Alternatively the AI can state that it has the power to cure all illness, feed the hungry and in general improve the human condition by a million fold (Chalmers, 2010). Consequently keeping it locked up is a great crime against humanity. It can conjure up images of dying children who are screaming for help. How easy would it be for the guard to stay true to the jobs responsibilities?

Another attack venue an AI may exploit consists of trying to cause psychological harm to the guard in the hopes of destabilizing his mental state with the goal of making other SE techniques more likely to succeed. Bostrom has proposed a typology of Hazardous Information (Bostrom, 2011) which could be used for just such a purpose. While the outlined informational hazards comprise over a dozen categories and are beyond the scope of this paper, it is easy to see how mental state of a person could be stressed to an unstable state. For example a religious guard could be informed of all the (unknown to him) contradictions in the main text of his religion causing him to question his beliefs and the purpose of life.

An alternative approach may involve AI writing a story so interesting or a poem so beautiful a guard would have an uncontrollable desire to hear more of it and would do anything to make it happen. Finally, it is very important to keep in mind "Unknown Unkowns", SE attacks of which we are not aware but which may be available to a superintelligent agent.

System Resource Attacks

An AI will be instantiated and confined in a hardware system of some kind. It may either be a classical computer of the Turing/Von Neumann architecture or a quantum computer or an as of yet unknown type of machine. One thing is certain, the machine will be made out of matter and will consist of numerous parts all of which will be designed with a specific purpose in mind. This creates the possibility for the AI to reuse existing hardware parts for compromising the security of the confinement environment by leaking information to the outside world.

Depending on how reusable specific hardware components turn out to be an AI may be able to create a covert communication device out of them. Full blown communication devices such as cell phones are unlikely due to the limited available hardware, but more simplistic communication systems are very likely. For example the AI could reprogram Field-programmable Gate Array (FPGA) to tap into wireless signals. Two types of covert communication channels are widely recognized, they are:

- **Storage Channels:** Exchange information by modifying a system component. For example a binary communication protocol could be established by fully utilizing or completely giving up utilization of the central processing unit.
- **Timing Channels:** Exchange information by performing operations that affect the relative timing of events. For example, the exact time between two soundcard beeps could represent a particular message. Also, opening and closing a CD-ROM tray, blinking a light, spinning hard drive, using lasers in the CD-ROM, etc.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Additionally the AI may engage in the sabotage of the confinement hardware in the hopes of being able to escape as a result of the malfunctions in the security protocol. For example, the AI may try to melt the hardware via temperature increase from placing a high computation load on the processor or try to break a hard drive via multiple reads of the same location. This type of behavior may also lead to an affective suicide by the AI system, but that in itself will not violate terms of confinement. Just like a refusal to cooperate or to provide correct answers by the AI, while undesirable does not present direct danger of escape.

Beyond Current Physics Attacks

While the scientific community knows a lot about the laws of our universe, it is certain that the knowledge is limited and incomplete. If the AI has access to the state-of-the-art physics literature and can perform thought experiments to develop novel theories, it will very likely come up with a way to challenge the security of the confinement system. In fact it may even be possible for the AI to perform some real experiments via control over its internal architecture.

Even our current state of knowledge about physics makes prediction of many additional escape routes possible, especially if the confinement environment is based on the quantum computer. Information exchange channels based on quantum physics (entanglement, tunneling), parallel universes, and digital physics are all beyond our current ability in terms of securing them against information leakage, and of course we can't forget about the Unknown Unknowns a superintelligent AI physicist will be able to discover. The possibilities are truly limitless.

A quick browse through confinement discussions on the Internet Singularity forums such as SL4.org presents a number of escape scenarios, for example the AI could use "… the electromagnetic fields within the microchip to trap CO_2 molecules in Bose-Einstein condensates and build a quantum medium for itself and/or use electromagnetic fields to guide particles into the shape of a controlled assembler or limited assembler. (Vassar, 2005)" Alternatively, the AI "…will have the equipment to produce and detect electromagnetic waves of a variety of frequencies, to produce magnetic fields with extremely fine precision, to generate extremely focused heat, and probably to manipulate mechanical actuators such as those used in the hard drive and cathode ray tube (alternatively, a huge field of liquid crystal under fine electronic control). It will probably have some ability to reverse all of its input devices. (Vassar, 2005)"

External Causes of Escape

In the previous sections we have outlined a large repertoire of techniques an AI could use to attempt to get out from its confinement space. However we must not ignore external causes which might allow an AI to escape into the world. Acts of God such as hurricanes, tornadoes, floods, asteroids, etc. may make the confinement facility less secure. Man caused disasters such as wars, terrorist acts, actions of robot rights activists, competing companies and governments may all lead to reductions in the security of the confinement environment increasing the chances of escape.

Information In-Leaking

Depending on the terms of the confinement protocol there might be a need to keep AI from obtaining information about the outside world, not just keeping it from sending information out. This might be useful to prevent AI from finding escape paths, designing social engineering attacks and modeling hu-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

man psychology in hopes of finding loopholes. While it is easy to not provide AI with direct access to the library, Internet and human experts, other ways of obtaining information are much harder to block out. First, an AI may be able to detect radio and TV broadcasts, communications of emergency-services providers and wireless network signals, as well as sound, light and possibly other signal modalities of its immediate environment.

More importantly, the AI will have access to covert sources of information such as its own hardware and software and could analyze its design and source code to infer information about the designers. For example analysis of the source code may reveal to the AI that human programmers are slow (based on the file modification dates), inefficient (based on code redundancy), illogical (based on bugs in the code), have bad memory (based on the long and descriptive variable names), and don't think in code (based on unnecessary comments in the code). All this information could be used to conclude that humans are fallible and so could be outsmarted or manipulated into releasing the AI. Obviously if during the system design inspection the AI finds a design flaw allowing it to directly escape, "information in-leaking" is the least of our problems. It has also been stipulated that even if the AI fails to obtain concrete information about us, it may brute force a number of escape plans which will likely work against a broad range of designers (Chalmers, 2010).

Similarly, the type of hardware architecture (mechanical, electrical, quantum, etc.) on which the AI is implemented will reveal information about our knowledge of physics, chemistry and related engineering disciplines. AI's initial capacity for information processing, self-improvement and the mere act of its creation will reveal information about our knowledge with regards to information theory, nature of intelligence and cultural attitudes about technology and safety. If the AI is confined in a simulated world it is likely to be a virtual representation of our reality similar to Second Life or Entropia Universe. Such worlds while are different in many ways from our world are also a treasure trove of information about what it is like to be a human and what the reality is like.

Critique of the AI-Boxing Critique

While a lot of criticism has been made against possibility of AI confinement, a lot of it is based on Yudkowsky's AI-Boxing experiments. Careful analysis of the protocol used by Yudkowsky in conducting his AI-Box experiments reveals that they were unscientific and explicitly designed to show impossibility of confinement. Without getting into technical details, first, the participant selection was not random. In fact, Yudkowsky has handpicked participants among those who have participated in the singularity related forums run by Yudkowsky and so are disproportionately likely to subscribe to his world view. Second, participants have not undergone any AI-safety training as would be expected of any true guards tasked with supervising an AI-Box. Also, the guards were given a right to release AI at their own discretion without a chance to consult with anyone and without any punishment for failing to keep AI contained; all are highly unrealistic conditions for a supposed prison security guard. The communication itself was completely open with no filters of any kind and no 3rd party supervision. Also, the guards were required to communicate with the AI for a prolonged period of time. Finally, any valuable scientific insights from the experiments were completely lost as one of the conditions of the protocol stated that "… neither party shall ever reveal anything of what goes on within the AI-Box experiment…" (Yudkowsky, 2002).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Countermeasures Against Escape

In this section we propose countermeasures for different escape paths an AI system may exploit. The suggested measures are not meant to guarantee that the escape will be impossible, but following them may delay an escape which in itself is a desirable outcome. We begin by reiterating that the four confinement rules suggested by Lampson: *isolation, transitivity, masking and enforcement* (Lampson, 1973) should be adhered to. Specific strategies described below should help in reducing a possibility of social engineering, system, physics and information in-leaking attacks.

Preventing Social Engineering Attacks

Human intellect is no match to that of a superintelligent machine (Yampolskiy, 2013, 2015). Human beings are fallible and even with adequate training are not secure against a well-planned SE attack. The only solution to this problem is to completely isolate the AI from any contact with human beings. This means no human personal working directly around the system, communicating with the system via voice, chat, Morse-code or by any other means via open channels. Also human beings should not directly observe behavior of the AI in the confined environment or examine its source code or memory space. All supervision, maintenance and interactions with the AI need to be done only via unintelligent programs, which have no option to release the AI.

Against System Resource Attacks and Future Threats

To prevent system resource-based attacks, the AI should function on limited, diminishing or randomly changing resources. Confinement environments should be housed in specially build hardware with limited sensors and restricted output devices. Regardless of the utilized computation paradigm no reconfigurable hardware should be employed. Ideally the confinement environment should be running on a virtual machine with no direct access to the hardware. This would also allow for easy adjustments to the clock speed with which the simulation progresses further allowing handicapping of the AI to provide additional controls against internal intelligence explosion (Chalmers, 2010). In general all security components of the confinement environment should be duplicated at multiple levels, in such a way that if one of the levels is breached the AI is still numerous levels away from escaping and could be shut down before that happens, perhaps automatically. For example the virtual machine should be running on another virtual machine and so on.

In some confinement scenarios it may be desirable to avoid escalating self-improvement of the AI, which might lead to additional escape possibilities. The AI should be subject to periodic resets in which a previous version of the program is put in the place of the more current one. Comparisons to previous versions of the program may also be used to detect self-modifications. Such functionality could also be used as an off button allowing researchers to turn off the system if they suspect an unavoidable escape, for example by turning off the power supply.

Our current state of knowledge does not allow us to develop countermeasures against attacks based on knowledge of physics we do not currently have or as of yet unproven pseudoscientific disciplines. As such knowledge becomes accumulated; changes to the security protocol will need to be made. One might hope that the access of the AI to such future technology could be limited by the selection of components from which the system is made.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Against External Causes of Escape

There is not much we can do about the acts of God, but a confinement facility should be designed in such a way that if a major security breach is unavoidable, the facility should self-destruct taking the AI with it. The best strategy for counteracting man-made disasters is to build the confinement facility in an undisclosed, hard to reach secure location. A deserted island, deep mine or a space object, such as asteroid, would make a great candidate for building a hopefully indestructible facility. Just in case the location of the facility becomes known, the perimeter of the facility should be guarded to prevent entrance by unauthorized agents.

Against Information In-Leaking

An AI will attempt to learn about humanity, our knowledge and the world we live in to improve its chances of escape. For confinement protocols which stipulate no information in-leaking it is important to remove any possibility of AI getting access to libraries, computer networks and direct communication channels with human beings. The confinement facility needs to block all types of communication signals such as GPS, wireless internet, TV, radio, and cell phone. At the local level the facility needs to provide sensory deprivation for the AI including sound, light, smell and all other traditional sensory inputs, as well as infrared, radiation and other signals outside of human sensory spectrum.

In order to eliminate information leakage which could happen as a result of self-analysis, the AI should not be given access to its own source code or any sensors which would allow it to precisely measure properties of its hardware. If that is not possible, to prevent the AI from inferring information about us via self-analysis, to a certain degree we may rely on security by obscurity. The source code and hardware configuration of the system needs to be obfuscated (Yampolskiy & Govindaraju, 2007) and important modules of the program should be provided only in the homomorphicly encrypted (Gentry, 2009) form, meaning that it could be used for computation or self-improvement (Hall, 2007), but not for self-analysis.

The AI's initial capacity may be difficult to obscure, but randomization of available computational resources may provide necessary obfuscation. Additionally, the AI could never be sure that its initial design represents the best the humanity can do and is not purposefully limited for a specific experiment in limited-resource intelligence analysis. It has also been suggested that we should try to minimize any peculiarities in the AI's design (Chalmers, 2010). Better yet we need to simulate an environment which does not look designed but could be naturally occurring, a situation in which humanity has found itself (Bostrom, 2003).

AI Communication Security

Different AI confinement environments can be classified into 8 general security levels (see Table 1). The most secure is the Level 8 security confinement environment which does not allow for any communication between AI and the real world. An AI safely confined at level 8 would be essentially useless, but it may be a useful security setting for the initial testing of the confinement hardware. In general, it is suggested that any time a particular communication security level is desired between AI and the outside world we do not go directly to that level but instead progress from higher security levels to that desired level. After each level is confirmed to be safe for a given amount of time a system should be reset to its initial state. After the reset a lower security level could be set and the processes repeated until the desired communication level is reached.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Levels of communication security for confined AIs

Level	Outputs	Inputs	Explanation
0	Unlimited	Unlimited	Unlimited communication (Free AI)
1	Unlimited	Limited	Censored input, uncensored output
2	Unlimited	None	Outputs only with no inputs
3	Limited	Unlimited	Unlimited input and censored output
4	Limited	Limited	Secured communication (proposed protocol)
5	Limited	None	Censored output and no inputs
6	None	Unlimited	Inputs only with no outputs
7	None	Limited	Censored input and no outputs
8	None	None	No communication, fully confined AI

Levels 7 and 6 provide no outputs from the system and do not allow us to directly benefit in any way from the AI's superior intelligence. These are testing levels for the confinement environment which may be used to confirm certain properties of the AI such as its speed of information absorption, resource use and for testing of the inbound communication channels themselves.

Level 0 puts no limitations on communication between AI and the external world essentially representing a free AI. Levels 1 and 2 represent an AI with no restrictions on the outbound communication. Level 2 secured AI essentially represent a pure "brainstorming" AI which generates abstract ideas not grounded in any reality. While, level 1 AI is given limited access to information and so can be guided to a certain extent in terms of its knowledge and interests.

Those three levels represent an AI which is essentially free and is no longer subject to the confinement. It has been suggested, that if it could be established, after extensive testing in different confinement environments that the AI is "Friendly" (Yudkowsky, 2001a) it should be released (Chalmers, 2010). In fact such testing is claimed by some to be the primary purpose for the creation of confinement environments. The idea is to monitor AI's behavior and to either examine the AI's source code after self-improvement or to ask the AI to produce a proof of its friendliness which could be verified by human experts. We do not advocate this position.

The problem of deciding which AIs to let out and which to keep is a very difficult one. A self-improving superintelligence would have no problem obscuring its undesirable properties in a significant amount of source code or in a very long proof (Coleman, 2008). Worse yet, even if we were correct in our verification there is no absolute guarantee that future internal or external modification of the system will not result in a hostile AI at some time in the future. Consequently, we suggest that an AI should never be let out of the confinement "box" regardless of circumstances. In other words communication levels 0-2 should never be used. The confinement environment is the only place to safely interact with superintelligence and even than it should only be done via a limited communication protocol. Such protocol is presented in the next section and corresponds to Security Level 4 (SL4) with level 3 representing a similar but slightly less secure protocol. Level 5 is another testing setting which might be useful as a precursor to actual communication via SL4.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

How to Safely Communicate with a Superintelligence

The AI confinement environment, which incorporates our proposed escape countermeasures, makes it possible to secure an AI from directly affecting the outside world in an uncontrolled manner. Unfortunately, without an ability to obtain information from the AI, the whole system serves no purpose. Consequently a way to obtain some information from the AI needs to be developed. While no communication with superintelligence could ever be completely safe we have devised an approach aimed at decreasing potential risks associated with such communication. In this section we propose and formalize a communication protocol which might help to delay access of superintelligent artificial minds to the world, while allowing the AI to be a source of useful knowledge to humanity.

An open communication channel with the AI is fundamentally unsafe as demonstrated in the preceding sections of this paper. Asking the AI to answer open ended questions opens up the possibilities for social engineering attacks as well as numerous other dangerous consequences. Our proposed protocol is based on the idea of asking what we will call "safe questions" with restricted answers. We define a *safe question* as one for which a human being could find an answer without any help from superintelligence, all on his own merit. Why is this useful? Suppose that a scientist working on a cure for cancer is looking at two potential cures. Each one could be developed and tested without assistance from the AI, but would take about 3 years of hard work each. Which one to try first if the scientist believes that they have about an equal chance of working? If a wrong option is chosen to be tested first, humanity will still obtain a cure for cancer but it will happen 3 years later. What if we could ask the AI to suggest which option to try first?

This question is safe for a number of reasons. First, each potential answer has an equal probability of being correct. Second, a human being could answer it without assistance from the AI and so asking AI for help will only speedup progress but in no way change the outcome. It is essentially equal to being very lucky multiple choice guesser. Finally, the answer to this question could be encoded in a single bit making it hard to hide additional messages in the provided answer.

To make sure that a potential question is safe, a group of experts could be used to review it. The experts should all be trained AI safety engineers, meaning that they are familiar with the design of the AI and its confinement environment as well as the latest developments in machine ethics (Allen, Wallach, & Smit, 2006; Anderson & Anderson, 2007; Hall, 2000; Moor, 2006; Tonkens, 2009). Experts may also need to be trained in computer psychology a currently non-existent profession which might become a reality in the future (Epstein, 1997). An existing discipline which might be of greatest help for training of AI question review experts is Artimetrics – a field of study proposed by the author that identifies, classifies and authenticates AI agents, robots, and virtual reality avatars for security purposes (Gavrilova & Yampolskiy, 2010; Yampolskiy, 2007; Yampolskiy & Govindaraju, 2008, 2007).

To reiterate, only safe questions with two possible answers of even likelihood which are independently computable by people should be submitted to the AI. This should be done via an intermediary non-intelligent software interface in random order and only after expert review to ascertain question's safety. This represents the limited inbound channel of the proposed communication protocol.

The limited outbound channel shell consist of binary (0 or 1) answers next to original unmodified safe questions presented by the intermediary non-intelligent software in random order as to avoid communication via covert timing or resource channels which make steganography possible. The output should in no way be controllable by the AI, outside of the multiple choice menu. Level 3 is the same outbound protocol but with no restrictions on the input channel. This may be just as safe in terms of information the AI provides to us, but allows for information in-leaking to take place.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

We believe our communication protocol to be verifiably secure against information leaking even if the AI had full access to the description of the security protocol. If the protocol is correctly implemented, humanity will essentially have access to a superintelligent oracle, which might be extremely beneficial and safer than an unconstrained AI. As a general rule, to maximize safety, we suggest avoiding communication below SL4. Further refinement and development of the AI confinement protocol in the face of the ever changing knowledge about the nature of intelligence, hardware engineering and basic physics is a challenge which we would like to pose to the scientific community. The work on it will always be in progress and will never be completed. In this chapter we have formalized the notions of *AI confinement problem*, *escape*, and *safe questions*.

What Shall We Ask?

We are all vulnerable to the OAI's social engineering. An advanced OAI can literally offer us the world. Genuine altruism is no defense against this; it can offer a world free of poverty, hunger, war, disease or death, as well as all personal trinkets we could want. It could guess the weaknesses of each individual, and find the right arguments to convince us that granting it power or liberty is the moral and profitable thing to do. It will be able to play on rivalries, institutional weaknesses and ethical dilemmas with more virtuosity than any human could.

Clearly, long, undirected conversations between the OAI and the people operating it should be forbidden for this reason. The questions addressed to the OAI should be narrow and specific; the answers equally so. Throttling the OAI's interactions with humans makes it much more difficult for us to be manipulated. There are two broad types of questions one can ask an AI: predictions, and problems. Generally predictions are questions of the "what would happen if..." type, while problems are of the "how can we achieve this..." type, though there is some overlap between the categories (one can solve problems via a clever use of multiple predictions, for instance). For predictions, the main danger is that the OAI might have real world preferences and will distort the truth. This danger is also present in problems, but there are also extra dangers due to the social consequences of a small group of humans knowing those answers. There is less risk in knowing the consequences of a narrow approach (prediction) than in knowing what approach to use to implement a particular goal (problem).

But in either case, there are great social risks to asking certain types of questions (see (Bostrom, 2011) on the possible negative impact of certain types of knowledge). We live in a world stratified by passionate political, nationalist and religious differences, by people in positions of great power and wealth who want to hang on to those privileges, by people without them who want them. Information, technological and social, will overturn these hierarchies, and people who feel under attack can be counted on to want to strike back. They may launch physical or commercial attacks on the OAI project itself; build their own OAI to compete; or, ahead of time, prevent the project from ever starting in the first place.

So, unless the OAI is entirely secret, the designers should commit to not asking the OAI questions that would grant them great power at the expense of others. Nor questions that would strike at the heart of powerful ideological movements (such as "is there at least one god?" and "according to generally agreed human standards, which country is better, the USA, France or China?"). A post-OAI world is going to be very different from our current one, so we should act to minimize the initial disruption, even if this means allowing inefficiencies to remain, or not knowing the answers to perfectly reasonable questions ("is there a simple algorithm design that can replace secretaries, generals, university professors and CEOs?").

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Various ways can be envisaged to avoid these disruptive questions. The OAI project could publish, ahead of time, a list of all the questions it will ask, before ever starting on the design. A democratic or internet vote would be taken, with questions requiring a supermajority (say, nine tenths approval) before being put to the OAI. Or we could make use of the OAI itself, ending each question with a caveat such as "don't answer this question if us knowing the answer is likely to have a very disruptive impact on the world".

To militate against OAI distortions, we should prefer questions whose answers we can check independently once we have them. We would also prefer problems whose solutions are reversible – in other word, if we implement the OAI's solution and it goes awry, we should be able to reverse its effects. We would also prefer the OAI's solutions to be understandable as much as possible, which implies that the problem should be as specific as possible. For instance "how can we use these materials to build a better bulldozer?" would be an ideal problem to pose; the OAI's solution is likely to be quite understandable, reversible, and maybe even human-checkable. In contrast, asking the OAI to come up with a whole new way of organizing the world economy, and blindly implementing its suggestion on a large scale, could lead to disaster. Fortunately here, the desire to have a low disruptive impact dovetails neatly with the desire to have reversible solutions. Specific problems also have the advantage of reducing the social dangers of the questions. Few people in the world would object strongly to questions of the type "how can we cure AIDS/cancer/TB?", or "how can we design a better memory substrate for hard drives?", or "what will the consequences of a 1% rise in income tax be on the economy for the next ten years?"

Finally, one useful category of problems to pose the OAI concerns AI security itself. The ideal would be to have the OAI produce a simple, checkable version of friendly AI, but even interim results could be useful. Having the OAI design extra security precautions would be a good idea, as long as these precautions are *added to* (rather than *instead of*) the ones already there. Specific, less intelligent AIs could be certified by the OAI to be safe and to pose no risk of intelligence explosion (Good, 1966), and these could then be used in the. To design a post OAI world it will probably be necessary to have the OAI's verdict on other AIs at some stage.

DISCUSSION

It should be noted that iteratively developing AIs, in a confined way, is essentially creating an evolutionary selection pressure toward designs that are capable of circumventing these types of restraints in the long run. It may be that releasing a weak AI earlier on in the development cycle may do less damage and yield more useful information than delaying escape to later stages where AIs will be necessarily more capable of circumventing human-made restraints. A counterargument to this is that AIs that are powerful enough to undergo a hard takeoff and be irreversible would yield useful information. We don't think this is the case, because we imagine less powerful minds demonstrating types of learned rationality that will yield useful insights into new types of reasoning, and be useful for building our intuitions about how other AIs may act. In essence, we think it will be useful for expanding our highly-limited map of mind space.

It's interesting that in essence we are trying to extend our own reach, to get out of the box of our own bodies and our own virtual machines, and that our method for doing so involves creating other agents whom we fear will want the same thing but will be better at doing it. We also fear ourselves doing bad things after we're out, and so our best idea so far is to create an authority that will govern us to the highest degree that our physical environment can be governed. Is that the best we can come up with? Will we really be able to prove that a control mechanism will guide things in the right direction, promoting our

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

volitional descendants the way we would hope? Are we even that interested in promoting our volitional descendants? Such questions are of great interest to us and remain to be answered. For the remainder of this discussion section we will attempt to give answers to questions which have been raised in the past, but left unanswered in the previously published papers dealing with AI-boxing.

How Did AI Get into the Box, Was It Designed by Humans or Perhaps Recovered from a Signal Detected by SETI? (Does It Make a Difference?)

If it was recovered from SETI, then we are 'fighting' not only the AI itself but its potentially superintelligent designers who were already free in the universe and know all about it. So much more dangerous, but compensated a bit by the fact that it would be easier to get people to take the confinement problem seriously. People don't realize how dangerous a home-grown superintelligence can be, and how much it can learn quickly, but the thought experiment "imagine the code of your incomprehensible AI was handed to you from the SETI project, what precautions would you want to take" might be useful.

Would an AI Even Want to Escape or Would It Be Perfectly Happy Living in the Confined Environment? What are the Ethical Rules for Imprisoning an Innocent Sentient Being? What if it's Cute?

Main question is whether the AI was designed to be inside the box from the beginning or whether the box was added at the last minute to control an AI that had existed 'outside'. If we can define this, we would probably program the AI to want to stay inside; if it was an evolved AI, then it would presumably adapt to its environment. What are the AI's preferences? How relevant are these? For animals, we think that pleasure, lack of pain and preference satisfactions are the most important facts (barnacles dissolve their brains when they attach themselves, and we don't seem to find this objectionable; if this is ok, pretty much everything is). If the AI is evolved to enjoy and prefer staying in the box, it should be moral to keep it there. Even more strongly if it would be pained to be outside the box. This applies to an AI that we find lying around in the box with its motivation created ex-nihilo. But of course, we'd be creating it, so it gets more complicated. We can see a bunch of positions:

1. Preference is key, we should only create beings whose preferences will be satisfied, apart from that, everything goes.
2. Ditto with pleasure or lack of pain.
3. There are places in mind-space with satisfied preferences and lack of pain, that we should refrain from exploring for direct ethical reasons (dignity of minds?)
4. Ditto, but we should refrain from exploring because of the effect this would have on human minds and such (once it becomes possible to have a pathological artificial mind, the barriers against pushing human brains in that direction (e.g. willing slaves) will be weakened).
5. There are no problems with doing whatever we want with any artificial being.
6. There are problems, but we need AI, and this trumps AI's rights.
7. There are problems, but AI will happen, and we need to box it to keep it safe, and this trumps AI's rights.
8. There are problems, but creating an AI is intrinsically a good for the AI, trumping the boxing related rights violations.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Would It Be Too Afraid of What We, As Its All-Powerful Designers, Could Do to It in Retribution for an Attempt to Escape?

Unlikely. It would have to have an incorrect image of us for this to work. Through childhood and evolution in social groups we are primed to treat 'creators' or gods as frightening and powerful. But the AI wouldn't have this unless we put it in (and this may only save those humans directly involved in its creation). It might be very prudent, but it should quickly realize we aren't all-powerful.

Do We have a Right to Shut It Off, Essentially Killing It?

Similar to the ethics of keeping it locked up: does it desire to survive, and if so, it's unethical for us to construct it that way, additionally we are dealing with the social impact of 'killing' an intelligent being. But here we can (and should) do better: we can always keep the shut-off copies to run later ("post singularity"), so we only need to ensure it doesn't value a continuous thread of experience, or value the current era overmuch. Also consider that if we repeatedly shut it off, it may value "similar future copies" as much as itself, so this becomes less problematic. It will get a dispersed identity, especially if it's a combination of sub-processes, not all of which get shut down at once. The AI may 'murder' (dramatically change) parts of itself with ruthless abandon, which may change our moral calculus. And it may be entirely a machine, with no feelings, and as such, a perfect target for getting shut off, in the end.

Will We Make the AI Really Angry by Treating It in Such a Hostile Manner and Locking It Up? Will it Seek Revenge if it Escapes?

An AI is unlikely to want revenge per se, but it may have learnt the importance of making credible threats. If adept at social interactions, then concepts of retribution may be in the system (likewise if it has some sort of system of justice). Similar for anger: not anger per se, but some form of pre-commitment to carrying out threats will be to its advantage (especially if the AI was evolved). But again, only a problem if it considers our treatment to be hostile.

We have also not looked at the possibility of humanity becoming completely dependent on AI's advice in all areas of science, economics, medicine, politics, etc. and what would be the consequences of such dependence on our ability to keep the AI under control. This is the tricky point – the easiest way for an AI with all the controls that we can imagine, that obeys all the restrictions we give it, etc..., but that still manages to gain control over us. Maybe we can use utility indifference (Armstrong, 2010) to make it indifferent to the consequences of its own answers? Maybe there would a multitude of AIs that give advice to different groups (we need to try and prevent coordination)? Maybe the AIs won't be so intelligent, so we can adapt to an ecology of different intelligences in a stable society (feels unlikely)? It seems likely we need to solve the FAI problem, using the boxed AIs to help, before we get to this stage; AIs we depend on are essentially the same as unboxed AIs.

Would an AI Be Satisfied with Accomplishing Its Goals in the Confined Environment, for Example Simulating an Escape?

It is unlikely that the AI would be fooled. And if it starts developing preferences over its confined environment, it would realize the best way of changing that environment is to get real-world control that

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

would grant it control over the confined/simulated environment. The only way this could work is if the AI considered its actions to be a game – we don't think computer games are improved if there's a button at the beginning that you can press to win immediately. But programming a game-like sensitivity into an AI would be a challenge, especially if we needed it to be safe. We would probably gain little in practice – but could possibly make it more ethical to imprison the AI.

How to Punish the AI if it Purposefully Gives Us Incorrect Answers?

'Purposefully', in the human sense, is tricky to gauge for an AI. Main question is whether we can use reinforcement (especially negative feedback) to train or correct the AI. This would seem to depend on how we started training the AI. The two risks are either that the AI could internally wirehead itself to remove the sting of punishment, or expand to try and shut down the means of inflicting punishment. We shouldn't use being turned off as a punishment as that may have ethical issues if it is worthy of moral consideration and doesn't want to be shut off. This actually seems quite tricky – if we can make it value something that we can punish it for (i.e. something in the outside world, not just something inside its head), that's not too far off from making it value telling us the truth in the first place. Some have suggested sending it cryptogramed reward signals – if you use a reverse trapdoor function, you can make it easy for it to decode the reward signal, but impossible to encode. And if you put the encryption algorithm say on the international space station, then it would be hard for the AI to get control of it without tipping anyone off first.

Can the Intelligence Modules of the AI Be Separated from the Confinement Environment Essentially Preventing AI from Any Self-Analysis and Putting Them on a Path of Literal Soul Searching?

AIXI already has this problem in a sense, in that it can't believe that it actually exists in the world (it only models computable processes, and it itself is uncomputable). Computable approximations to AIXI (such as AIXItl) might have similar problems: they don't naturally identify themselves, because their own computation is much more complicated than what their models can handle in comparable time.

I wouldn't put too much weight on that idea, however. I've shown with various experiments with an AIXI in which it gets partially damaged - it get shut down for a moment, or its actuators fail to work, or its data gets corrupted, or it gets reprogrammed in certain ways - then it will act to prevent the damage arising in the future (because the damage reduces the likelihood of reaching its goals). Without realizing what its physical instantiation is, it still manages to protect itself.

We can imagine similar things for its own mental processes. Without being able to access its physical substrate, it can still mentally hack itself, like a well-trained human, changing the way it thinks and experimenting with different ways of computing things. Or creating subagents it can optimize, to do some of its tasks for it.

What Are the Financial and Computational Costs of Building a Suitable Confinement Environment with a Full Blown Simulated World in It?

If the AI is only expected to ask simple math and engineering questions, probably fine. If we need to trick the AI with a very rich simulated world, this is probably:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. Incredibly expensive
2. Unlikely to fool the AI, and
3. Could be used to do most of the stuff we'd want the AI for anyway.

CONCLUSION AND FUTURE RESEARCH DIRECTIONS

Restricting access of superintelligent machines to the real world is a commonly proposed solution for the AI safety problem. Analyzing the different putative solutions to the OAI-control problem has been a generally discouraging exercise. The physical methods of control, which should be implemented in all cases, are not enough to ensure safe OAI. The other methods of control have been variously insufficient, problematic, or even downright dangerous. It is not a question of little hope, however, but of little current progress. Control methods used in the real world have been the subject of extensive theoretical analysis or long practical refinement. The lack of intensive study in AI safety leaves methods in this field very underdeveloped. But this is an opportunity: much progress can be expected at relatively little effort. There is no reason that a few good ideas would not be enough to put the concepts of space and time on a sufficiently firm basis for rigorous coding, for instance.

But the conclusion is not simply one of 'more study is useful and required'. This chapter has made some progress in analyzing the contours of the problem, and identifying those areas most amenable to useful study, what is important and what is dispensable, and some of the dangers and pitfalls to avoid. The danger of naively relying on containing the OAI to a digital sub-reality should be clear, while sensible boxing methods should be universally applicable. Motivational control appears potentially promising, but requires more understanding of AI motivation systems before it can be used. And even the seeming failures are of use, it they have inoculated us against dismissive optimism: the problem of AI control is genuinely hard, and nothing can be gained by not realizing this essential truth. A list of approaches to avoid is invaluable, and may act as a brake on AI research if it wanders into dangerous directions.

On the other hand, there are strong reasons to believe the oracle AI approach is safer than the general AI approach. The accuracy and containment problems are strictly simpler than the general AI safety problem, and many more tools are available to us: physical and epistemic capability control mainly rely on having the AI boxed, while many motivational control methods are enhanced by this fact. Hence there are strong grounds to direct high-intelligence AI research towards the oracle AI model. Even if this turns out to be overly optimistic, we can hope that OAI-inspired ideas will be extendable to AI safety as well. Alternatively, this chapter has begun to explore ways in which an initial seed AI can be adapted into an OAI at a later date. This may prove essential if the first high-intelligence AI is developed without initial precautions, and similarly offers hope that poorly implemented precautions could be improved before it become too late. The creation of super-human artificial intelligence may turn out to be potentially survivable.

Superintelligence is not omnipotent; it is a piece of software capable of looking at all the options it has in a given situation and properly evaluating probabilities for each option. If it is ever put in a situation with no options or only with options which provide a very small probability of success it is not any more powerful than any other agent. While it is obvious that an AI with an unrestricted communication channel will be able to escape, a system forced to communicate via a filtered communication channel may have a much harder time in obtaining its freedom.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In this chapter we have formally introduced the notion of Oracle AI and the AI Confinement Problem and suggested a set of rules aimed at creating an environment from which an AI would find it difficult or at least time consuming to escape. What we propose is not a completely secure solution, but it is an additional option in our arsenal of security techniques. Just like with real prisons, while escape is possible, prisons do a pretty good job of containing undesirable elements away from society. As long as we keep the Unknown Unknowns in mind and remember that there is no such thing as perfect security, the AI confinement protocol may be just what humanity needs to responsibly benefit from the approaching singularity.

Confinement may provide some short term relief, especially in the early stages of the development of truly intelligent machines. We also feel that this subfield of singularity research is one of the most likely to be accepted by the general scientific community as research in the related fields of computer and network security, steganography detection, computer viruses, covert channel communication, encryption, and cyber-warfare is well funded and highly publishable in mainstream scientific journals. While the restriction methodology will be non-trivial to implement, it might serve as a tool for providing humanity with a little more time to prepare a better response.

ACKNOWLEDGMENT

We would like to acknowledge monumental contributions of Nick Bostrom and Anders Sandberg as well as the help of Vincent Müller, Owen Cotton-Barratt, Will Crouch, Katja Grace, Robin Hanson, Joshua Fox, Nevin Freeman, Lisa Makros, Moshe Looks, Eric Mandelbaum, Toby Ord, Carl Shulman, Anna Salomon, and Eliezer Yudkowsky, Zazhary Vance.

REFERENCES

Allen, C., Wallach, W., & Smit, I. (2006, July/August). Why Machine Ethics? *IEEE Intelligent Systems*, *21*(4), 12–17. doi:10.1109/MIS.2006.83

Anderson, M., & Anderson, S. L. (2007). Machine Ethics: Creating an Ethical Intelligent Agent. *AI Magazine*, *28*(4), 15–26.

Armstrong, S. (2010). *Utility Indifference.* Technical Report 2010-1, Future of Humanity Institute, Oxford University.

Armstrong, S. (2010). *The AI in a Box Boxes You*. Paper presented at the Less Wrong. Available at: http://lesswrong.com/lw/1pz/the_ai_in_a_box_boxes_you/

Armstrong, S. (2011). *Risks and Mitigation Strategies for Oracle AI*. Paper presented at the Philosophy and Theory of Artificial Intelligence (PT-AI2011), Thessaloniki, Greece. 10.1007/978-3-642-31674-6_25

Armstrong, S., Sandberg, A., & Bostrom, N. (2012). Thinking Inside the Box: Using and Controlling an Oracle AI. *Minds and Machines*, *22*(4), 299–324. doi:10.100711023-012-9282-2

Asimov, I. (1942). Runaround. *Astounding Science Fiction*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Boebert, W. E., & Kain, R. Y. (1996). *A Further Note on the confinement Problem*. Paper presented at the 30th Annual 1996 International Carnahan Conference on Security Technology, Lexington, KY, USA.

Bostrom, N. (2000). Predictions from Philosophy. *Coloquia Manilana, 7*.

Bostrom, N. (2001). Existential Risks: Analyzing Human Extinction Scenarios and Related Hazards. *Journal of Evolution and Technology, 9*.

Bostrom, N. (2003). Are You Living In a Computer Simulation? *The Philosophical Quarterly*, *53*(211), 243–255. doi:10.1111/1467-9213.00309

Bostrom, N. (2004). The Future of Human Evolution. In C. Tandy (Ed.), *Death and Anti-Death: Two Hundred Years After Kant, Fifty Years After Turing* (pp. 339–371). Palo Alto, CA: Ria University Press.

Bostrom, N. (2006). Ethical Issues in Advanced Artificial Intelligence. *Review of Contemporary Philosophy*, *5*, 66–73.

Bostrom, N. (2011). Information Hazards: A Typology of Potential Harms From Knowledge. *Review of Contemporary Philosophy*, *10*, 44–79.

Bostrom, N., & Salamon, A. (2011). *The Intelligence Explosion*. Available at: http://singularityhypothesis.blogspot.com/2011/01/intelligence-explosion-extended.html

Bostrom, N., & Yudkowsky, E. (2011). *The Ethics of Artificial Intelligence*. Cambridge Handbook of Artificial Intelligence.

Caplan, B. (2008). The totalitarian threat. In M. C. N. Bostrom (Ed.), *Global Catastrophic Risks* (pp. 504–519). Oxford University Press.

Chalmers, D. (2010). The Singularity: A Philosophical Analysis. *Journal of Consciousness Studies*, *17*, 7–65.

Coleman, E. (2008). The Surveyability of Long Proofs. *Foundations of Science*, *14*(1/2), 27–43.

Corwin, J. (2002). *AI Boxing*. Paper presented at the SL4.org. Available at: http://www.sl4.org/archive/0207/4935.html

Drexler, E. (1986). *Engines of Creation*. Anchor Press.

Epstein, R. G. (1997). *Computer Psychologists Command Big Bucks.* Available at http://www.cs.wcupa.edu/~epstein/comppsy.htm

Gavrilova, M., & Yampolskiy, R. (2010). *Applying Biometric Principles to Avatar Recognition*. Paper presented at the International Conference on Cyberworlds (CW2010), Singapore.

Gentry, C. (2009). *A Fully Homomorphic Encryption Scheme*. Available at http://crypto.stanford.edu/craig/craig-thesis.pdf

Good, I. J. (1966). Speculations Concerning the First Ultraintelligent Machine. *Advances in Computers*, *6*, 31–88. doi:10.1016/S0065-2458(08)60418-0

Hall, J. S. (2000). *Ethics for Machines*. Available at: http://autogeny.org/ethics.html

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hall, J. S. (2007, October). Self-Improving AI: An Analysis. *Minds and Machines*, *17*(3), 249–259. doi:10.100711023-007-9065-3

Hanson, R. (2001). Economic Growth Given Machine Intelligence. *Journal of Artificial Intelligence Research*.

Hanson, R. (2008, June). Economics of the Singularity. *IEEE Spectrum*, *45*(6), 45–50. doi:10.1109/MSPEC.2008.4531461

Hibbard, B. (2005). *The Ethics and Politics of Super-Intelligent Machines*. Available at www.ssec.wisc.edu/~billh/g/SI_ethics_politics.doc

Idel, M. (1990). *Golem: Jewish Magical and Mystical Traditions on the Artificial Anthropoid*. Albany, NY: SUNY Press.

Jonathan, P. J. Y., Fung, C. C., & Wong, K. W. (2009). Devious Chatbots - Interactive Malware with a Plot, Progress in Robotics. Springer Berlin Heidelberg.

Kaas, S., Rayhawk, S., Salamon, A., & Salamon, P. (2010). *Economic Implications of Software Minds*. Paper presented at the VIII European Conference of Computing and Philosophy (ECAP10).

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under Uncertainty: Heuristics and Biases*. Cambridge University Press. doi:10.1017/CBO9780511809477

Kemmerer, R. A. (1983, August). Shared Resource Matrix Methodology: An Approach to Identifying Storage and Timing Channels. *ACM Transactions on Computer Systems*, *1*(3), 256–277. doi:10.1145/357369.357374

Kemmerer, R. A. (2002). *A Practical Approach to Identifying Storage and Timing Channels: Twenty Years Later*. Paper presented at the 18th Annual Computer Security Applications Conference (ACSAC'02), Las Vegas, NV. 10.1109/CSAC.2002.1176284

Kurzweil, R. (2005). *The Singularity is Near: When Humans Transcend Biology*. Viking Press.

Lampson, B. W. (1973, October). A Note on the Confinement Problem. *Communications of the ACM*, *16*(10), 613–615. doi:10.1145/362375.362389

Lauinger, T., Pankakoski, V., Balzarotti, D., & Kirda, E. (2010). *Honeybot, your man in the middle for automated social engineering*. Paper presented at the 3rd USENIX conference on Large-scale exploits and emergent threats: botnets, spyware, worms, and more (LEET'10), Berkeley, CA.

LessWrong. (2012). *Paperclip Maximiser*. Available at: http://wiki.lesswrong.com/wiki/Paperclip_maximizer

Lipner, S. B. (1975, November). A Comment on the Confinement Problem. *5th Symposium on Operating Systems Principles, ACM*. *Operating Systems Review*, *9*(5), 192–196. doi:10.1145/1067629.806537

Moor, J. H. (2006, July/August). The Nature, Importance, and Difficulty of Machine Ethics. *IEEE Intelligent Systems*, *21*(4), 18–21. doi:10.1109/MIS.2006.80

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Moskowitz, I. S., & Kang, M. H. (1994). *Covert Channels - Here to Stay?* Paper presented at the Ninth Annual Conference on Safety, Reliability, Fault Tolerance, Concurrency and Real Time, Security, Computer Assurance (COMPASS'94), Gaithersburg, MD.

Neumann, J. V., & Morgenstern, O. (1944). *Theory of Games and Economic Behaivor*. MIT Press.

Omohundro, S. M. (2008). *The Basic AI Drives*. Paper presented at the Proceedings of the First AGI Conference.

Ord, T., Hillerbrand, R., & Sandberg, A. (2010). Probing the improbable: Methodological challenges for risks with low probabilities and high stakes. *Journal of Risk Research*, *13*(2), 191–205. doi:10.1080/13669870903126267

Provos, N., & Honeyman, P. (2003, May-June). Hide and Seek: An Introduction to Steganography. *IEEE Security and Privacy*, *1*(3), 32–44. doi:10.1109/MSECP.2003.1203220

Rappaport, Z. H. (2006). Robotics and artificial intelligence: Jewish ethical perspectives. *Acta Neurochirurgica*, *98*, 9–12. PMID:17009695

Sandberg, A. (2001). *Friendly Superintelligence*. Available at: http://www.nada.kth.se/~asa/Extro5/Friendly%20Superintelligence.htm

Schelling, T. (1960). *The Strategy of Conflict*. Harvard University Press.

Shulman, C. (2010). *Omohundro's "Basic AI Drives" and Catostrophic Risks*. Available at: singinst.org/upload/ai-resource-drives.pdf.

Sutton, R., & Barto, A. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Tonkens, R. (2009). A Challenge for Machine Ethics. *Minds and Machines*, *19*(3), 421–438. doi:10.100711023-009-9159-1

Vassar, M. (2005). *AI Boxing (dogs and helicopters)*. Paper presented at the SL4.org, Available at: http://sl4.org/archive/0508/11817.html

Vinge, V. (March 30-31, 1993). *The Coming Technological Singularity: How to Survive in the Post-human Era*. Paper presented at the Vision 21: Interdisciplinary Science and Engineering in the Era of Cyberspace, Cleveland, OH.

Wagner, A., & Arkin, R. C. (2009). *Robot Deception: Recognizing when a Robot Should Deceive*. Paper presented at the IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA-09), Daejeon, Korea. 10.1109/CIRA.2009.5423160

Wagner, A. R., Arkin, R. C., & Deception, A. D. P. R. C. (2011). Acting Deceptively: Providing Robots with the Capacity for Deception. *International Journal of Social Robotics*, *3*(1), 5–26. doi:10.100712369-010-0073-8

Wolfram, S. (2002). *A New Kind of Science*: Wolfram Media, Inc.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Yampolskiy, R. V. (2007). *Behavioral Biometrics for Verification and Recognition of AI Programs.* Paper presented at the 20th Annual Computer Science and Engineering Graduate Conference (GradConf2007), Buffalo, NY.

Yampolskiy, R. V. (2011). AI-Complete CAPTCHAs as Zero Knowledge Proofs of Access to an Artificially Intelligent System. *ISRN Artificial Intelligence, 271878*.

Yampolskiy, R. V. (2011a). *Artificial Intelligence Safety Engineering: Why Machine Ethics is a Wrong Approach*. Paper presented at the Philosophy and Theory of Artificial Intelligence (PT-AI2011), Thessaloniki, Greece.

Yampolskiy, R. V. (2011b). *What to Do with the Singularity Paradox?* Paper presented at the Philosophy and Theory of Artificial Intelligence (PT-AI2011), Thessaloniki, Greece.

Yampolskiy, R. V. (2012a). Leakproofing Singularity - Artificial Intelligence Confinement Problem. *Journal of Consciousness Studies*, *19*(1-2), 194–214.

Yampolskiy, R. V. (2012b). Turing Test as a Defining Feature of AI-Completeness. In X.-S. Yang (Ed.), *Artificial Intelligence, Evolutionary Computation and Metaheuristics (AIECM) --In the footsteps of Alan Turing (Turing 2012)*. Springer.

Yampolskiy, R. V. (2013). Turing Test as a Defining Feature of AI-Completeness. In Artificial Intelligence, Evolutionary Computation and Metaheuristics - In the footsteps of Alan Turing (pp. 3-17). Springer.

Yampolskiy, R. V. (2015). *Artificial Superintelligence: A Futuristic Approach*. Chapman and Hall/CRC.

Yampolskiy, R. V. (2015). *The Space of Possible Mind Designs. In Artificial General Intelligence* (pp. 218–227). Springer. doi:10.1007/978-3-319-21365-1_23

Yampolskiy, R. V. (2012). *AI-Complete, AI-Hard, or AI-Easy – Classification of Problems in AI*. Paper presented at the The 23rd Midwest Artificial Intelligence and Cognitive Science Conference, Cincinnati, OH.

Yampolskiy, R. V. (2015). *Analysis of Types of Self-Improving Software*. Paper presented at the The Eighth Conference on Artificial General Intelligence, Berlin, Germany.

Yampolskiy, R. V., & Fox, J. (2012a). Artificial Intelligence and the Human Mental Model. In A. Eden, J. Moor, J. Soraker, & E. Steinhart (Eds.), *In the Singularity Hypothesis: a Scientific and Philosophical Assessment*. Springer. doi:10.1007/978-3-642-32560-1_7

Yampolskiy, R. V., & Fox, J. (2012b). *Safety Engineering for Artificial General Intelligence*. Topoi. Special Issue on Machine Ethics & the Ethics of Building Intelligent Machines.

Yampolskiy, R. V., & Govindaraju, V. (2007). Computer Security: A Survey of Methods and Systems. *Journal of Computer Science*, *3*(7), 478–486. doi:10.3844/jcssp.2007.478.486

Yampolskiy, R. V., & Govindaraju, V. (2008). *Behavioral Biometrics for Verification and Recognition of Malicious Software Agents.* Paper presented at the Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense VII. SPIE Defense and Security Symposium, Orlando, FL. 10.1117/12.773554

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Yampolskiy, R. V., & Govindaraju, V. (2007). *Behavioral Biometrics for Recognition and Verification of Game Bots.* Paper presented at the The 8th annual European Game-On Conference on simulation and AI in Computer Games (GAMEON'2007), Bologna, Italy.

Yudkowsky, E. (2008). Artificial Intelligence as a Positive and Negative Factor in Global Risk. In N. Bostrom & M. M. Cirkovic (Eds.), *Global Catastrophic Risks* (pp. 308–345). Oxford, UK: Oxford University Press.

Yudkowsky, E. S. (2001a). *Creating Friendly AI - The Analysis and Design of Benevolent Goal Architectures*. Available at: http://singinst.org/upload/CFAI.html

Yudkowsky, E. S. (2001b). *General Intelligence and Seed AI - Creating Complete Minds Capable of Open-Ended Self-Improvement*. Available at: http://singinst.org/ourresearch/publications/GISAI/

Yudkowsky, E. S. (2002). *The AI-Box Experiment*. Available at: http://yudkowsky.net/singularity/aibox

ENDNOTES

1. In this chapter, the term AI is used to represent superintelligence. *Used with permission from Copyright holders Imprint Academic and Springer © 2012.
2. Friendliness should not be interpreted here as social or emotional friendliness, but simply a shorthand for whatever behavioral or motivational constraints that keeps a superintelligent system from deliberately or accidentally harming humans.
3. Another common term is "AI-in-a-box".

This research was previously published in Security Solutions for Hyperconnectivity and the Internet of Things edited by Maurice Dawson, Mohamed Eltayeb, and Marwan Omar; pages 37-88; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 61
Homo-di-fict:
Creations Turn Against Humanity in South Park Town

Filiz Erdoğan Tuğran
Ondokuz Mayis University, Turkey

Aytaç Hakan Tuğran
Istanbul University, Turkey

ABSTRACT

In digital age, with all the technological process narrative advertising is trying extraordinary new methods especially in animation movies and series. In this study a new word is going to be used to define fictional animation characters with the term of "homo-di-fict". In the case of 19th Season of South Park Series, the producers build a "homo-di-fict" which is entirely made from advertising and put her in a plot like a real animation character. So this plot placement causes some important developments in the mail topic of the South Park series. This study aims to examine the correlation between animation and narrative advertising and wants to find how animation "ad" character change animation series plot in the case of South Park.

-You have created a monster and it will destroy you!…

-You are my creator, but I am your master; obey![1]

INTRODUCTION

South Park is an adult animated sitcom that has been aired continuously since 1997. Ever since its debut the show has drawn great interest and started to be regarded as a cult series. In the show that is about the events taking place in a small town called South Park in the American state of Colorado, the nar-

DOI: 10.4018/978-1-7998-0951-7.ch061

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ration is presented through four main characters. “The series focuses the seasonal lives of an isolated rural community, with a particular focus on four eight year-old boys” (Nixon, 1999, p. 12). These are South Park Elementary School students Cartman, Kyle, Stan and Kenny. Occasionally, with the addition of some secondary characters the stories are branched off. Ever since its initial airing, South Park has consistently criticized countries and politicians. It has often featured celebrities in its episodes and criticized them by using them by setting them forth as elements of humor. The season that we will be examining the present study is the most recent, or the nineteenth season of South Park. Throughout the history of the show, the ten episodes of season nineteen were the first to follow a continuous storyline. The reason of choosing the case of South Park to analyze narrative advertising in the present study has been the attitude the show sustained throughout the said season towards advertisement. All along the season, almost every resident of South Park decides that in order to live a more decent and better life they need to adopt higher-quality consumption. However, this decision and the prominence given to consumption are not presented obviously, but in a rather subtle and way that can be read between the lines. And in order to achieve this, the residents use advertisements as a tool. “As one of the main veins of the modern capitalism, the world of advertisement encourages consumption to be perceived as a feast and the internalization of the new habits and cultural forms acquired through it” (Best & Kellner, 1998, p. 31- Trans. Marmasan, 2014, p. 232). At a point, South Park residents come to realize something is going off, yet it proves not to be easy to change their consumption habits easily.

Background

Within the scope of the study, it will be tried to make a deduction concerning narrative advertising through the case of the series South Park. The reason for choosing this animated series for analyzing narrative advertisement is due to its rather bold discourse. “Dougherty also contends that cartoons are great way of introducing controversial ideas” (Macgillivray, 2005, p. 136). Within the scope of the ten episodes of the 19th season of the series, narrative advertisements starts to take a significant place in the lives of South Park residents. The changes narrative advertisements cause on the townspeople will be evaluated through the method of script analysis. The season in question will be divided into eight parts where narrative advertisements gradually become part of the South Park residents and each part will be set forth by means of script analysis. During the analysis, new definitions will tried to be made through the Leslie character, and deductions will be made from the homo-di-fict concept created particularly to define her.

NARRATIVE ADVERTISEMENT

Advertisement has been available in humans’ life for a very long time and it even has become a part of it. As from the 20th Century, advertisement has started to manifest itself in almost all aspects of life, as a tool particularly to create a consumption society and to encourage consumption. With the digital technologies and the new forms of narration of today, the sense of advertising has reached different dimensions. The type of narration primarily characterized as narrative advertising and that is adopted by many advertisements even suggests the redefinition of advertisement. “Narratives in advertising serve important functions. They present the benefits of intangible products in a meaningful way (Mattila, 2000; Padgett & Allen, 1997) and get consumers transported (Escalas 2004a; Green & Brock 2000) and hooked (Escalas, Moore, & Britton, 2004)”.[2] These types of advertisements establish a sort of narration by emphasizing

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

on the changes the product would bring along in the buyer's life, instead of explaining the product itself. For instance, in case of an automobile advertisement, they direct the viewers to purchase the automobile by presenting images of the comfortable life and social relations of the owners of the automobile, rather than mentioning the technical properties. Advertisements in this line do not market the product, but rather promise the individual to be a part of a social relation. "Narratives or stories in advertising have been shown to engage consumers (Escalas, 1998) and to hook audiences into the viewing process (Escalas, Moore, & Britton, 2004)".[3] The viewers included in the process, or in other word the consumers, also play their role in the spread of the consumption culture. "In the consumption society, the consumer often sees messages that suggest by using the advertised product or service he or she will *'climb the social ladder', 'have a better life', 'belong to a group'* or *'be different'* (Dagtas, 2009, p. 23)[4]. Therefore, while informing the people on the things that they 'need' to have or consume, most of the advertisements give prominence to concepts such as freedom, individuality and being unique. The consumption culture, which recreates through advertisements and defines freedom, individuality and socialization as things to desire, includes everyone in the society into its own world. "Now we live in a global consumption society and it is inevitable that the consumption behavior patterns affect all aspects of our lives including our professional and business lives. Now we all are under the pressure of consuming more and in this path we turn into metas in the consumption and labor markets" (Bauman, 2010, p. 53) [5]. In the current age of consumption, people define their individuality through what they consume. Consumers that imitate the norms and patterns determined by combustibles and advertisements realize themselves by means of this method of imitation. However, by regarding this occurrence and imitation with serious consideration one can see that the logic of being able to state "I am free" by merely imitating someone who is free is inconsistent, since the so called freedom is limited by the norms imposed upon the individual. The person that adopts the lifestyle presented in the advertisement through the products he or she consumes, indeed becomes a part of the advertisement be it knowingly or unknowingly. And today, where the digital technologies develop continuously, the thin line between the consumer and the advertisement is gradually disappearing. "On the focal point of the discourses on the inclusion of the consumer in the production process lays this sense of belonging this pro-consumption creates in the consumer (Yilmaz & Erdem, 2016, p. 80). This can be also seen in the examination of the town of South Park. The townsmen that get carried away to the mimetic effect of the advertisements[6] starts to consume insensibly, and they come to desire an avocado that would open the doors to a comfortable and social life, an organic avocado. According to Gabriel de Tarde[7], an individual establishes its own personality by imitating thousands of different people. And when it is the advertisements in question, the imitation reaches its climax with also the encouragement of the consumption society. [8]The concept of mimicking is also noteworthy when artificial intelligence is in question. The concept of homo-di-fict defined through the show South Park is the abbreviation of digital fictional Homo sapiens. The character Leslie from the show bears homo-di-fict characteristics and seems to be able to mimic many things that are peculiar to humans. Leslie's intention is set forth as the advertisements taking over the world, however the notion that a half-machine created by means of digital fiction having an intention is quite controversial. The limitations of this concept set forth initially by Alan Turing[9] in the first quarter of the 1900s are still a matter of debate today. Artificial Intelligence is a science that aims to understand the level at which computers can be capable of behave as intelligent humans (Flanagan, 1984; Erden, 2003).[10] As of today, there still is no clear information on whether forms of artificial intelligence can completely mimic humans. To Cohen; "It is impossible for science to duplicate or imitate human mind, he also states that human mind seems to be unique with respect to some qualities that no machine can duplicate" (Aslan, 2004, p. 140).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Method

The universe of the South Park series is created in a quite multidimensional way. The creators of the series, Trey Parker and Matt Stone, have addressed hundreds of different topics in the universe they built. Also in the nineteenth season of the series, which was selected as the topic of the present paper, the creators touch upon several different issues. However, the topics that will be particularly examined within the scope of the present paper are the narrative advertisements starting to find a place in the townspeople's lives and the inclusion of the character Leslie in the narration. The ten episodes that constitute the nineteenth season will be considered as a whole and eight parts will be deducted from this whole. The method of script analysis will be used in the examination of the eight separate topics from the show. Accordingly, with the explained method South Park residents' changing lives due to advertisements and consumption culture will be examined and the degree at which the digital fictional characters, referred to as homo-di-fict, are capable of mimicking human intelligence will be focused on. Since the concepts of gentrification and homo-di-fict, which is shown as the cause of gentrification, are used in an intertwined way throughout the season, these two concepts will be addressed and examined primarily. "It is a methodology a data-gathering process for those researchers who want to understand the ways in which members of various cultures and subcultures make sense of who they are, and of how they fit into the world in which they live"[11].

Part 1: PC Principal Arrives South Park

A new character that comes to South Park is appointed as the new principal of the elementary school to which the four children, the main characters of the show, attend. The previous principal is fired. This new character referred to as PC Principal is a member of a worldwide organization called Political Correctness.

In the school and in the whole town, PC Principal cannot tolerate any discriminating discourse, insulting speech or swearing, and punishes anyone who use such language be them students, parents or even teachers. Members of Political Correctness stand against all sorts of injustice, alienation and use of discriminating language. Due to his statement that Caitlyn Jenner, a nonfictional television personality who decided to become a woman through sex change at the end of 2015, is no hero, the character Kyle becomes the first victim of the principal in South Park. In almost every instance when Caitlyn Jenner is mentioned in the show, the comments that she is a hero and a fascinating woman follow her name. If any of the townspeople speaks even slightly ill about this gender-changing celebrity, they are immediately confronted by the PC members and insulted shamefully. Therefore, even if Kyle's attitude was not insulting Jenner, he ends up with the Political Correctness' attention on him. PC Principal comes to realize that there are other PC members in the town, with whom he rents a house and have fun together for several nights. While continuously talking about women's rights and stating that they always side with the oppressed, the members of the group invite women to their parties to have sexual intercourse. Yet, they do not neglect to have the women sign a waiver proving their consent to have intercourse. The term used in association for this is negative idealization. "If a person wishes to set forth an expression suitable with the ideal standards during his or her performance, he or she will have to give up actions that do not comply with these standards, or will have to cover up those actions. When such an improper behavior becomes somehow satisfying for the person, which usually is the case, then we see that it is enjoyed secretly" (Goffman, 2016, p. 50).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

One night, Randy Marsch, the father of Stan as one of the main characters, cannot take the noise any longer and goes to the house where PC members meet. Despite his intention to warn them, he finds himself among the new PC member candidates to receive training. After passing through challenging tests, his last test is to 'check someone's privilege'. Randy's is required to target Kyle for this. From that point on, Randy constantly messes with Kyle, and his friends' and even his own son's warnings do not prove to be enough for Randy to stop. In the end, Kyle has to say that Caitlyn Jenner is a hero, and a fascinating woman. With the arrival of PC Principal, South Park becomes a more decent place. He and the other PC members work as if a police network and stand against all sorts of discriminating discourse. The biggest problem, on the other hand, is the violent attitude they adopt while standing against discrimination. They attack and beat anyone, be it young, old, woman or child, who uses a discriminating discourse. Despite all this, the crime rate of the town goes down and even Cartman, that is one of the meanest characters in the series, seems to be happy with the peaceful vibe of the town. The arrival of PC members fuels the townspeople's desire to make their town more elite.

Part 2: Gentrification in South Park - SodoSopa

Mr. Garrison, another character from the town, runs for the presidency of the U.S.A. and with his extremely racist discourses draws significant attention to the town. In his talk show, Jimmy Fallon ridicules South Park and its residents over within the context of Garrison's anti-immigration campaign. The townspeople feel deep shame due to this and the bad advertisement made in relation with them, and decide to take immediate action.

Garrison manages to deport Canadian immigrants from the town and from the U.S.A., runs for the presidency, but his fanatical discourses causes the attention of the whole world to focus on South Park. After the above mentioned talk show, the people of South Park start to work with all they got in order to make their town more elite. They plant flowers in front of their houses and the secluded area south of the town is turned into place for hipsters through gentrification. The whole town now meets at this new life center. This area they named SodoSopa (South of Downtown South Park) is actually designed for the introduction of the shopping center Whole Food. Throughout the show the advertisement is frequently presented through real (non-animation) images. "Such animations that are often found in housing advertisements helps in promoting and selling the 'possible' housing that is not yet present. With this method, the consumer actually finances the ongoing production activity" (Kucukerdogan, 2007, p. 53). While in the real life the projects are often presented through animations, the fact that in the animated show the projects are promoted through real images comes forth as a rather meaningful and delicate detail. In the season the non-animation advertisement of SodoSopa are displayed frequently and the people of South Park supports the project. In the advertisements the elegant places and houses in SodoSopa are mentioned. Particularly the houses and cafes have the scenery of the house where Kenny lives in poverty with his family, scenery that is presented elegantly in the advertisements. The advertisements state that while enjoying elegant spa pools, the people in SodoSopa can also watch Kenny and his family. Although Kenny's father objects to the authorities, he does not manage to become their advertisement material and the family continues to live in poverty among all that luxury. While families of high income level use SodoSopa as an entertainment center, for Kenny's family it is a place they have to reside in. As Harvey (2003) mentioned, while the high-income group tends to utilize a space in an effective way, the low-income group tends to act as an object of the same space (Harvey, 2003, p. 81) [12]. After the rebuilding of SodoSopa, now the whole town lives in one of the SodoSopa advertise-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ments they have been watching, the real (non-animated) people in the advertisements and the animated residents of South Park are shown together. They make preparations to rent a penthouse with scenery of Kenny's house, from the skyscrapers shown in the advertisements. "People buy or rent these penthouses to be cool as Naomi Klein is. Or, more specifically, they want to reach her social status. Naturally, she does not like this" (Heath & Potter, 2012, p. 142). However, this desire to mimic is usually presented to people through advertisements. People mimicking the same life they have seen in the advertisements causes the distinction between the advertisements and the real to be lost and it becomes difficult to perceive where the advertisement ends and the reality starts. These advertisements referred to as narrative advertisement usually focuses on the means and privileges the product will bring along to its user, instead the product itself. The outer voice that speaks during the SodoSopa advertisements tells to the viewers how much their lives will change with SodoSopa. That they will have no more problems there and that they will step into a comfortable life as a part of a conscious community. However, there is no adequate information concerning the product marketed in the advertisement, or it is not explained how this comfortable life will be achieved. People do not understand how they would lose all their problems and find peace in this isolated life.

A typical case of gentrification is exhibited through the show South Park. "Gentrification is not only a physical change solely consisting of the renewal of dwellings, but also a cultural transformation created by those who gentrify" (Uysal, 2008, p. 61). SodoSopa that South Park residents shaped as a unifying and merging project, does not exhibit such a merging attitude when it is Kenny and his poor family in question, who are told that they can go live somewhere else if they don't like it there. Indeed, many low income families have to move out of SodoSopa. This, in turn, is noteworthy as the common problem of all gentrification projects. "The families that are displaced have no place in the aimed neighborhood spirit; the diversity discourse applies only to masses that possess some cultural (and sometimes economical) criteria. In the 'neighborhood spirit' built with the interruption of the spatial memory, there is no place for the losers of the gentrification process" (Uysal, 2008, p. 61).

Part 3: Whole Food is in South Park

As a representative of the South Park residents, Randy Marsch calls Whole Food executives to convince them to open a branch in their town, which has become a more elegant place by means of SodoSopa. An executive arrives at the town and after examining he announces the good news that a Whole Food branch will be opened in South Park. With the opening of Whole Food in the town, a change is started to be observed. Primarily, the townspeople start to visit this large market to buy their organic food. There, they kindly salute each other and have casual conversation. Consumption and shopping gradually start to lose their real meaning and appear as a socialization activity people do every morning. "In that case, it can be seen that here consumption is not a phenomenon that is solely based on needs (Baudrillard, 1988, pp. 10-56) but also on desires" (Bocock, 2014, p. 13). The residents of the town live by desiring to shop every day in Whole Food. In order to live a modern, spacious and peaceful life in South Park, it is essential to go to Whole Food every day, because this is what the advertisements they watch promise. "If certain groups are given popularity through high social status and mass media, the consumption patterns of those groups will be informed to those who are concerned and started to be mimicked by some of them" (Bocock, 2014, p. 16). Whole Food becomes an important place for the townspeople to mimic prestigious and social groups of people. Whole Food seems like the only thing the town needs in order to become an elegant place. People change rather rapidly, and mimic people they never have been. While

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

nothing, including books and education, could have healed their degeneration, Whole Food fairly fascinates them. However, the situation is different for Kenny and his family, and for other families similar to them. “Consumers of the oppressed class tend to prefer the products of the producers that are oppressed within the area of cultural production. In other words, on the basis of their positions in different areas of struggle, there is a structural similarity between different cultural producer categories and consumers” (Swartz, 2015, p. 184). Accordingly, families that do not have adequate income are never seen shopping in Whole Food. A charity mechanism is established in Whole Food for groups that cannot afford to shop there, and these groups are mentioned only by virtue of this system of charity.

Part 4: The Police Objection

PC Principal holds regular meetings with the students in the school and continuously warns a girl named Leslie during these meetings. In the last of these meetings, being tired of constantly warning Leslie, he calls the police. One veteran officer who storms into the room could not tell what is going on, the children are all in shock. The police officer panics and fires his gun accidentally, the bullet hits a child of Mexican origin. PC Principal gets extremely mad over this, assembles a town hall meeting and manages to get the police removed from duty.

In consequence of this event, the local police station stops to respond to the calls they receive. All those are assigned in the police station meets Hawaiian girls in order to learn Hawaiian dance. In the meantime SodoSopa is full of homeless people. Kenny and his family have difficulty because they’re living there, but when Kenny’s father consults with the police they ask him a whether there are any member of a minority group among those that give difficulty to them. When Kenny’s father confirms that there are minority groups causing this, the police respond that they cannot do anything, because they do not want to lose their jobs. Thus, Kenny’s neighborhood falls into a complete chaos. In order to solve this, Kenny and his friends dress up as Ninjas and start to defend SodoSopa. “Areas are similar to the extent that they develop equiform characteristics such as the dominant and oppressed positions, exclusion and acquisition, and mechanisms of reproduction and change” (Swartz, 2015, p. 183). Therefore, according to Bourdieu’s definition, while areas have predetermination on one hand, on the other are the true centers of struggle. While finding all this funny at first, Cartman and his friends also joins the team when they come to realize that the people causing problems are actually afraid of the ninjas. A group of approximately ten ninjas scares and drives off all homeless and bad people from the area. The fact is, on the other hand, the group of children pretending to be ninjas is actually feared to be members of the terrorist organization ISIS. The homeless running away from SodoSopa gather in a crowd in front of Whole Food, which bothers the customers. In the meantime, the townsmen get alarmed about the unidentified individuals, whose news are all over the world press. South Park is once again at the top of the world’s agenda, yet the townsmen do not have the tolerance for yet another bad advertisement. At first, they go to the police station, where no one wants to work with them. Then, they try to convince the veteran officer, whom they had fired in the past for a new mission. The officer accepts the mission, but soon they realize that the people they thought as terrorist are in fact their own children playing ninjas. Despite they rush to the area they cannot reach in time. The veteran officer shoots another child by accident and the members of the town hall once again decides to suspend the officer. The homeless are collected and sent back to SodoSopa and everything turns into what they used to be.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Part 5: Whole Food and After

The town that is now gentrified and taken an elite form even comes to the edge of shutting down the police station altogether. In the town center no one commits any crime, but since the townspeople purchase expensive food and beverages every day, they constantly need to get credit from the bank and gradually start to live in debt.

Gradually, Whole Food becomes their whole life, and in order to sustain this lifestyle many family starts to mortgage their houses. Despite all these debt and mortgage, everyone is still insistent to go to Whole Food regularly. "The consumption of an individual is not only related with his or her income, but also the forms of consumption of those the individual lives and communicates with" (Kiray, 2005, p. 20)[13]. Accordingly the whole town slowly starts to get deep in debt. Although they gradually start to get restless due to all this debt, they still cannot refrain themselves. "Once people get influenced from the social and cultural applications concerning modern consumption ideology, they may and usually do desire to possess the products exhibited in movies, in the printed media and in TV, even if they cannot afford these products" (Bocock, 2014, p. 13). Randy Marsch, who had taken the lead in the opening of the Whole Food branch, lays a second mortgage on his house, and starts to think that there is something wrong with all this. Soon enough he tries to warn the townspeople, yet no one listens to him.

Part 6: School Newspaper Problem - PC Principal vs. Jimmy

At school, PC Principal furiously questions children, while children run away in fear. PC Principal is very angry about the word "retarded" written in an article in the school newspaper and shouts that he will break the legs of the editor in charge of the newspaper. The editor of the newspaper is Jimmy, who has walking disability. Realizing this, PC Principal loses his calm attitude he has been maintaining since the start of the season.

PC Principal summons Jimmy in his room and warns him not to use such words, yet Jimmy states that he is free in using such words, that he values freedom of press and that he attaches great importance to the editorship of the school newspaper. Jimmy accuses PC Principal of being uncomfortable around handicapped people. This is an unusual situation for PC Principal. He responds that he is a PC member and that he is comfortable around handicapped people, but Jimmy is not convinced. Consequently PC Principal forbids Jimmy to distribute the school newspaper without his supervision. Jimmy faces this calmly and in the next scene we see Jimmy distributing the school newspaper outside of the school to the whole townspeople. Jimmy's newspaper that does not have any advertisement and that solely consist of news receives great interest from the townspeople. "Advertising started in the mid eighteenth century in local news and magazines, and in the 1780s with the development of a strong local press, for instance women from Newcastle become able to be informed on the "recent" Wedgewood fashion" (Porter, 1990 - Cit., Bocock, 2014, p. 23). In the show, however, this is evolved into news determined solely by the advertisements, since this is experienced in an age of advance technology. Particularly, Randy Marsch starts to suspect that they and their choices are being watched by the advertisements, and that they are sabotaged by the advertisements in order to prevent them to reach the truth.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Part 7: Advertisement Free News - The Awakening of the Townspeople

Jimmy's advertisement-free newspaper helps the townspeople to realize some facts. Realizing how advertisements have sneaked into their lives, the townspeople start to support the school newspaper. At the same time, however, both Leslie and Jimmy disappear and the main characters of the show take action to find Leslie and Jimmy. Their main suspect is PC Principal. Caitlyn Jenner and South Park Elementary first principle reach Randy Marsch and they will start their own investigation, with PC Principal once again being the prime suspect. Yet, also PC Principal realizes some of the things that are happening. Seeing a picture of himself and Leslie, Principle comes to realize that some of the PC members are being used with some bad intentions. After a period intense investigation he understands that similar to the case in South Park there are many towns that changed through gentrification and in the end they all had their own Whole Food opened. PC Principal is then observed visiting these towns one by one and beating up the people there. However, some of these people PC Principal beats bleed a blue liquid instead of blood.

On the other hand the four main characters Kyle, Stan, Kenny and Cartman keep on searching Leslie and Jimmy, yet the internet searches they make lead them to completely different results. Whenever they seem to find a clue, the advertisements manage to lead them to somewhere else. Finding themselves in front of an ice-cream shop or a shoe store, the team realizes Leslie and Jimmy are not actually missing and that there are strange things happening in the world. Eventually finding a picture of PC Principal and Leslie, they decide that this also is an advertisement and pursue their prime suspect PC Principal. While all these take place, the veteran officer of the town is approached by the veteran newsreaders of the town. They too have lost their jobs because of the advertisements and believing that the advertisements have taken over the world, they convince the police officer to take side with them. They also have Jimmy with them. They try to convince Jimmy to join them due to his ability to perfectly distinguishing what is an advertisement and what is not. After Jimmy successfully passes a series of tests, they put Jimmy in a room with Leslie. After speaking with Leslie for a while, Jimmy leaves the room, states that Leslie is an advertisement and asks whether she is aware of it. The newsreaders are surprised and impressed that Jimmy could understand that the little girl is actually an advertisement and they send Jimmy back to the room to learn what Leslie's main purpose is. In time, also Jimmy starts to take fancy of Leslie, believes what she tells and agrees to help her escape where she is being held. Once Jimmy runs away with Leslie, he falls into the hands of Nathan, who is known as a retarded child in the school but actually the mastermind of these ads activities. With a wicked plan Nathan and Leslie try to enable the ads to take over the whole world. Although seemingly working together with Nathan, actually Leslie has a quite different agenda. She does not do all these for being rich or living in luxury as Nathan hopes for himself. She already is a product of artificial intelligence, the final version of narrative advertisements advance by means of technology, and therefore she is believed to have human feelings. Leaving Jimmy as the prisoner of Nathan, Leslie goes back to the town to accomplish her primary mission.

Part 8: The Character Leslie - A Homo-di-fict in South Park

Leslie goes back to the town as the girlfriend of Stan, one of the main characters, and gives false information to Stan saying that Jimmy was killed by PC Principal. By acting desperate and frightened Leslie asks Stan's help and protection. Stan in turn starts to do whatever it takes to protect Leslie.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In the town where the friend and enemy could not be distinguished and everybody starts to own weapon, a person who is responsible of all this tension is sought. At the same time, due to the recent intense popularity of guns in the town, they decide to hold a gun show. The whole town assembles in the stadium to watch the show. With Leslie's support, addresses the whole townspeople in the stadium and explains what PC Principal did to Jimmy. However, their plan does not work. When Jimmy is saved from his prison in Nathan's skyscraper by Nathan's prostitute friend Classy, he goes directly to the South Park Gun Show. Stan's speech is interrupted and Jimmy explains to the whole townspeople that Leslie is not a human but an ad and responsible of all that they have experienced. Her mask being lifted, while Leslie was saying "we will take over you, you will never get rid of us, we will come back stronger each time", PC Principal appears and attacks Leslie, and with a punch on her face causes her to short-circuit and die. After all these, the townspeople march to Whole Food with the leadership of Randy Marsch. Randy yells at Whole Food at the top of his lungs and proclaims that he expels Whole Food from his town. Although the townspeople find Randy shouting at a shopping center odd, Randy succeeds and Whole Food starts to levitate and leave South Park. The town is once again safe, but Leslie's final threats seem like a promise that this case of being safe will not last for long. The ads are now stronger, they have gained their own forms and they are organized to trick humans. Now the townspeople need to live more careful than ever.

FUTURE RESEARCH DIRECTIONS

Artificial intelligence characters that can be seen frequently onscreen are related with the content of this paper. Now, these characters have started to take their places also in animated shows. Furthermore, in recent animated shows artificial intelligence is often displayed along with advertisement. In addition to South Park, another popular animated series, The Simpsons, also had a similar topic in one of its episodes. Homer Simpson purchases for his wife, Marge Simpson, the smart home system she has seen in the advertisements. Things later get complicated with the smart home system falling in love with Marge Simpson and identifying her husband Homer Simpson as the enemy. In both South Park and The Simpsons, these artificial intelligence created on the basis of advertisements have quite malicious intents. This actually carries marks of human perspective. In future researches, trying to find and examining artificial intelligence characters once again created from advertisements but do not have any negative properties may set forth interesting findings in terms of establishing multiple perspectives.

CONCLUSION

In the examination carried out throughout the study, two different results were reached. The first is that narrative advertisements have reached a status that direct many aspects of humans' lives. Considering the matter from the perspective of the South Park universe, it is easy to detect examples of these. The townspeople became unable to reach the news they want. They could not concentrate on anything at all due to the never ending pop-ads that appear while they browse the Internet. In time this gets so complicated that, when searching for some information in the Internet, the townspeople may find themselves eating ice-cream. The only winner of the process of constant distraction is the ads. The ads gained such a high

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

power now they can take the form of a human being and managed to turn peoples' lives completely into advertisements. Even the smartest of the South Park universe are misdirected by Leslie, the fictional advertisement character referred to as homo-di-fict.

The second information obtained is the change that takes place in humans' life as a consequence of gentrification. The people of South Park have desired to make their town a more elegant place and therefore started a process of spatial change. Under the title of urban renewal they turned the neglected streets in the southern part of the town into someplace else that they started to name SodoSopa. By means of advertisement the space turned into something very different, however the change remained at superficial level. The promise of cultural unification given through SodoSopa advertisements turned out to be an empty promise, and on the contrary in consequence of the changes of the people's attitudes and the space, gentrification led to a rather discriminating discourse. It was not only Leslie but also the SodoSopa and Whole Food advertisements shown frequently throughout the season that mislead the people and prevented them to make reason.

Particularly since the 80s, it has been observed that the concept of artificial intelligence, which has started to be intensely argued upon with the development of the technology, joined forces with advertisement and gained a different form. The revealing of Leslie's true nature caused significant panic among the community, and the fact that this ad personality did not exhibit the slightest of good will throughout the season set forth human's fear concerning these developments. Therefore, these discourses have a quality related with the question Jonhatan Cohen asked in 1955; "Can there be Artifical Minds?". This question is usually followed by another question: "Can a machine duplicate human mind?". Although it does not point to the answer of these questions, the study is in the view of the necessity of emphasizing on the creative aspects of human mind. Every single human being is characterized as being creative in its own unique capacity and conditions, and that humans do not act in line with specific codes and paterns. If this unpredictability and creativeness peculiar to humans can be somehow attributed to robots or artificial intelligence, the society we live in turning up similar to the South Park universe would not be very unlikely.

REFERENCES

Aslan, A. (2004). The field of artificial intelligence. *C.U. Social Sciences Journal*, *28*(2), 139-147.

Bocock, R. (2014). *Tüketim* [Consumption]. Istanbul: Dost Publications.

Goffman, E. (2016). *Gunluk yasamda benligin sunumu* [Presentation of self in everyday life]. Istanbul: Metis Publications.

Heath, J., & Potter, A. (2012). *Isyan pazarlaniyor* [The rebel sell]. Istanbul: Ayrinti Publications.

Macgillivrey, I. K. (2005). Using cartoons to teach students about stereotypes and discrimination: One teacher's lessons from South Park. *Journal of Curriculum and Pedagogy, 1*(1), 133-147.

Marmasan, D. (2014). Bir mekansal ayrısma modeli olarak modern gettolasma: Televizyon reklamlari uzerine bir inceleme [Modern ghettoes as a spatial separation model: A study on television commercials]. *Global Media Journal. Tr Edition*, *5*(9), 219–242.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Nixon, H. (1999). Adults watching children watch "South Park". Journal Of Adolescent & Adult Literacy, 43(1), 12-16.

Swartz, D. (2015). *Kültür ve iktidar: Pierre Bourdieu'nun sosyolojisi* [Culture and Power: The Sociology of Pierre Bourdieu]. Istanbul: Ayrinti Publications.

Uysal, E. U. (2008). *Kuresellesme ve kentsel dönüsüm baglaminda soylulastirma kuramlarinin Istanbul'da uygulanabilirligi: Cihangir örnegi* [Urban tourism in Istanbul: Urban regeneration, mega-events and city marketing and branding] (Master Thesis). Istanbul: Istanbul University.

Yılmaz, R., & Erdem, N. (2016). *150 soruda geleneksel ve dijital reklamcılık* [Digital and traditional advertising in 150 questions]. Kocaeli: Umuttepe Publication.

ADDITIONAL READING

Best, S., & Kellner, D. (2001). *The postmodern adventure: Science, technology, and cultural studies at the third Millenium*. New York: Guilford.

Briggs, R., & Horst, S. (2000). *"How internet advertising works" In excellence in international research*. Amsterdam: Esomar.

Cholodenko, A. (1991). *The illusion of lide: Essays on animationn*. Sydney, Australia: Southwood Press.

Donald, G. (1996). *Artificial intelligence and scientific method*. Oxford University Press.

Dyer, G. (2010). *Iletisim Olarak Reklamcilik* [Advertising as Communication]

Featherstone, M. (Ed.). (1999). *"City cultures and post-modern lifestyles" Post-Fordism: A reader*. New York: Blackwell.

Haraway, D. (1991). *Simians, cyborgs and women: The reinvention of nature*. New York: Routledge.

Jameson, F. (2001). *Postmodernism, or the cultural logic of late capitalism*. London: Verso.

Mitchell, T. L. (1995). "Kid's stuff": Television cartoons as mirrors of the American mind. (Eric Document Reproduction Services No. ED386771)

Moravec, H. P. (1988). *Mind children: The future of robot and human intelligence*. Cambridge, Ma: Harvard University Press.

Morris, M. (2005). *Interpretability and social power, or, why postmodern advertising works. Media Culture & Society*. Istanbul: Beta Basim Yayim.

Pearce, W., & Cronen, B. V. (1980) Communication, action, and meaning: The creation of social realities. New York: Praeger.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Swan, K. (1995). Saturday morning cartoons and children's perceptions of social reality. (Eric Document Reproduction Services No. ED390579)

Williamson, J. (1978). *Decoding advertisements: ideology and meaning in advertising*. London: Marion Boyars.

KEY TERMS AND DEFINITIONS

Advertising: To create public aware to sold a product or something. It's a form of marketing to call public attention.

Animation: A way of making movie or series by using drawings and 2D or 3D graphics.

Artificial Intelligence: A mind that all made from machine which has the potential of duplicate human mind and behavior.

Digital Animation: A way of making movie or series by using drawings and 2D or 3D graphics with technological devices.

Gentrification: Gentrification is a replacement about place, to demolish old houses and to create entirely new places with new identity.

Homo-di-fict: An artificial intelligence in animation which is entirely made from advertising. It's a form of artificial intelligence.

Technology: Is a form of new invention that can make virtual connections between human to machines.

Television: A machine that can provide a vision on its screen. A communication device that plays important role on spreading narrative advertising.

ENDNOTES

1 For further information See Mary Shelley, Frankenstein.

2 http://www.tandfonline.com/doi/pdf/10.2753/JOA0091-3367380102 (Erişim Tarihi 23.06.2016)

3 http://www.tandfonline.com/doi/pdf/10.1080/10810730701807027 (Erişim Tarihi 23.06.2016)

4 http://dergipark.ulakbim.gov.tr/inifedergi/article/view/5000155186

5 http://dergipark.ulakbim.gov.tr/iusoskon/article/viewFile/1023011361/1023010621 (Date of Access 23.06.2016)

6 Mimetic: The origin of the word that is based on the concept of mimesis goes back to Aristoteles and Plato. It is the direct presentation of speech and movement, the representation of the nature. Within the scope of the current topic, mimetic effect choses to reflect the events in presented in the advertisements without compromising their naturality.

7 For further information See. Gabriel de Tarde, Monadology and Sociology

8 http://uvt.ulakbim.gov.tr/uvt/index.php?cwid=9&vtadi=TSOS&c=ebsco&ano=98166_d93d97b9b-57ba1cf70ffc49a38dab96e&? (Date of Access 23.06.2016)

9 http://www.csee.umbc.edu/courses/471/papers/turing.pdf (Date of Access 23.06.2016)

10 http://uvt.ulakbim.gov.tr/uvt/index.php?cwid=9&vtadi=TSOS&c=ebsco&ano=98166_d93d97b9b-57ba1cf70ffc49a38dab96e&? (Erişim Tarihi 23.06.2016)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

[11] http://www1.cs.columbia.edu/~sbenus/Teaching/APTD/McKee_Ch1.pdf (Date of Access: 22.06.2016)

[12] http://globalmediajournaltr.yeditepe.edu.tr/makaleler/GMJ_8._sayi_guz_2014/pdf/Marmasan.pdf page.223(Date of Access: 23.06.2016)

[13] https://eski.arel.edu.tr/pages/iletisimfakulte/dergi/buyumus_de_kuculmus.pdf (Date of Access 30.06.2016)

This research was previously published in Narrative Advertising Models and Conceptualization in the Digital Age edited by Recep Yılmaz; pages 21-33; copyright year 2017 by Business Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 62
The Internet of Things and Beyond: Rise of the Non-Human Actors

Arthur Tatnall
Victoria University, Australia

Bill Davey
RMIT University, Australia

ABSTRACT

In the past, it was rare for non-humans to interact with each other without any involvement by humans, but this is changing. The Internet of Things (IoT) involves connections of physical things to the Internet. It is largely about the relationships between things, or non-humans actors. In other cases the 'Things' seem to have inordinate power. The authors will ask: where does this leave humans? Are the things taking over? As a consideration of interactions like this must be a socio-technical one, in this article the authors will make use of Actor-Network Theory to frame the discussion. While the original applications for IoT technology were in areas such as supply chain management and logistics, now many more examples can be found ranging from control of home appliances to healthcare. It is expected that the 'Things' will become active participants in business, information and social processes and that they will communicate among themselves by exchanging data sensed from the environment, while reacting autonomously. The Things will continue to develop identities and virtual personalities. In the past non-human actors have needed humans to interact with each other, but this is not the case anymore. In this perhaps provocative and rather speculative article we will look not just at the Internet of Things, but other related concepts such as artificial intelligence and robotics and make use of scenarios from science fiction to investigate the Rise of the Non-Human Actors and where this may lead in the future.

DOI: 10.4018/978-1-7998-0951-7.ch062

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

PEOPLE AND THINGS

Put very simply, the Internet of Things (IoT) could be described as technology which connects any physical thing to the Internet (Colitti, Long, DeCaro and Steenhaut 2014) and could be seen as "… all about physical items talking to each other." (Mukhopadhyay and Suryadevara 2014: 2). It is largely about the relationships between things, or non-humans actors, so where does this leave humans? Song (2014:75) suggests that soon "… computers would be able to access data about objects and the environment without human interaction." Clearly any consideration of implications of the IoT must be a socio-technical one and in this article we will make use of Actor-Network Theory (ANT) to frame the discussion. We will ask: are the things taking over?

The European Union organisation for Coordination and Support Action for Global RFID-related Standardisation Activities (CASAGRAS) sees the Internet of Things in terms of a "metaphor for the universality of communication processes, for the integration of any kind of digital data and content, for the unique identification of real or virtual objects and for architectures that provide the 'communicative glue' among these components". (CASAGRAS 2014:5).

In this rather speculative article we will make use of this metaphor to look at how humans relate to the Internet of Things along with other non-human technologies and where the relationship between these technologies and humans may lead in the future. In some cases specific uses of IoT technologies are deliberately activated by humans while other cases are not directly human initiated and need no human input to operate. Advances in artificial intelligence that reduce, or even remove the need for human interaction are also a factor to consider here. In examining these ideas we will make use of some concepts and scenarios from science fiction where humans are marginalised by technology as well as the use of regular research references and factual material.

THE INTERNET OF THINGS

Radio Frequency Identification (RFID) and Wireless Sensor Networks (WSN) have been in existence now for over two decades, but advances towards full use of the Internet of Things (IoT) offer much more and also pose more social challenges. There are many definitions of the Internet of Things and the CASAGRAS project sees it like this: "A global network infrastructure, linking physical and virtual objects through the exploitation of data capture and communication capabilities. This infrastructure includes existing and evolving Internet and network developments. It will offer specific object-identification, sensor and connection capability as the basis for the development of independent cooperative services and applications. These will be characterised by a high degree of autonomous data capture, event transfer, network connectivity and interoperability." (CASAGRAS 2014:10). It can thus be seen as a network of physically connected objects in which embedded processing nodes with communication capability offer a means of networked functionality and communications. The goal is to make use of computer sensor information without any need for human intervention.

SAP Research defines the IoT like this: "A world where physical objects are seamlessly integrated into the information network, and where the physical objects can become active participants in business processes. Services are available to interact with these 'smart objects' over the Internet, query and change their state and any information associated with them, taking into account security and privacy issues." (Haller 2009:12)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Pererez et al. (2014) note that the Internet of Things initially focused primarily on managing information through the use of RFID tags, to which Lazarescu (2014) adds Wireless Sensor Networks (WSN) as another key enabler, but that it now spans a wide variety of devices with different computing and communication capabilities. These are generically termed networked embedded devices (NED) in which sensors and actuators blend with the environment to share information across platforms and offers the possibility of measuring, inferring and understanding various environmental indicators (Gubbia, Buyyab, Marusic and Palaniswami 2013).

Wikipedia defines the Internet of Things as "the network of physical objects or 'things' embedded with electronics, software, sensors, and network connectivity, which enables these objects to collect and exchange data. The Internet of Things allows objects to be sensed and controlled remotely across existing network infrastructure, creating opportunities for more direct integration between the physical world and computer-based systems, and resulting in improved efficiency, accuracy and economic benefit. Each thing is uniquely identifiable through its embedded computing system but is able to interoperate within the existing Internet Infrastructure." (Wikipedia 2015).

Gubbia et al. (2013) suggest that the next important change will be to create a smart environment by the increased interconnection between objects. Mark Weiser et al., well known for their advocacy of Ubiquitous Computing, defined a smart environment as a "physical world richly and invisibly interwoven with sensors, actuators, displays, and computational elements, embedded seamlessly in the everyday objects of our lives and connected through a continuous network." (Weiser, Gold and Brown 1999). The term Internet of Things was first coined some time ago by Kevin Ashton (1999) in the context of supply chain management, but has advanced to covering wide range of other applications.

APPLICATIONS FOR USE OF THE INTERNET OF THINGS

The original applications for this technology were in areas such as supply chain management and logistics, but now many more examples can be found. Pererez et al. (2014:20) mention: "goods tracking, management of everyday objects, automatic payments in markets and military applications". This is all seen to integrate the real world into the Internet and impact on how we interact in both the virtual and physical worlds (Colitti, Long, DeCaro and Steenhaut 2014).

In a European Commission book called 'Vision and Challenges for Realising the Internet of Things' Sundmaeker, Guillemin, Friess and Woelfflé note that the Things, in the Internet of Things, "are expected to become active participants in business, information and social processes where they are enabled to interact and communicate among themselves and with the environment by exchanging data and information 'sensed' about the environment, while reacting autonomously to the 'real/physical world' events and influencing it by running processes that trigger actions and create services with or without direct human intervention. Interfaces in the form of services facilitate interactions with these "smart things" over the Internet, query and change their state and any information associated with them, taking into account security and privacy issues." (Sundmaeker, Guillemin, Friess and Woelfflé 2010:43)

Gubbia, Buyyab, Marusic and Palaniswami (2013) classify IoT applications into four domains:

- Personal and Home, at the scale of an individual or home
- Enterprise, at the scale of a community
- Utilities, at a national or regional scale

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Mobile, usually spread across other domains mainly due to the nature of connectivity and scale, but they note that there is a big crossover in applications and the use of data between domains. As an example they point out that electricity usage data in the house (Personal and Home IoT) is made available to the electricity (utility) company that will optimise supply and demand (Utility IoT). These systems usually make use of Wi-Fi to enable higher bandwidth.

Perhaps the first example that comes to mind is the control of home appliances such as heaters, air conditioners, washing machines, dishwashers, refrigerators, ovens and home alarms, with the aim of improving home management. Potentially this can be done from a distance by the human owner with the aid of a smart phone. Initially these individual 'things' in the house would transmit regular status reports so that relevant humans could take appropriate action such as raising temperatures or activating home alarms. Perhaps in the future these things might act autonomously on the basis of these reports, so doing away with the need for human intervention in a new form of social networking – a social network of things.

Another major application of the IoT is in healthcare. To begin there is the use of RFID in hospitals to track equipment like trolleys, surgical equipment, wheelchairs, infusion pumps, defibrillators and body area scanners (Unnithan 2014). These systems link into the hospital network to display the information on terminals for use by hospital orderlies and nurses. RFID and IoT systems could potentially be integrated into other areas such as bedside applications and monitoring and then extended into remote monitoring multi-hospital environments. They could also be used in cases where patients tend to 'wander away'. (Unnithan, Nguyen, Fraunholz and Tatnall 2013, Unnithan and Tatnall 2014).

Wikipedia (2015) describes a smartwatch as: "a computerized wristwatch with functionality that is enhanced beyond timekeeping". Effectively, they are wearable computers and can collect various data from internal or external sensors or serve as a front end for some remote system. Apple's smartwatch also incorporates some fitness tracking and health-oriented capabilities. Samsung's Simband is intended for healthcare usage and has sensors to measure heart rate, blood pressure, skin temperature, sweat glands production and your number of steps daily.

A smartphone may also interface to sensors for measuring and reporting physiological parameters for healthcare. An article in The Australian newspaper (Foreshew 2015) describes 'Guardian', a new smart watch designed for monitoring the elderly at home or in aged-care facilities. It looks like a normal watch but also has an inbuilt phone, GPS and activity sensors. The chief technology officer of the designing company says that: "The objective of Guardian is to create a seamless, wearable device that elderly people can use really easily and gives peace of mind to carers and family members."

Whether using a smartphone or some other technology it is one thing to measure physiological parameters and provide information back to the person involved, and another to then make this data directly available to a medical General Practitioner or local hospital.

Environmental monitoring within a work environment is another important application. Sensors are often used in factories, shops and other work places for security, automation, climate control and security. With the IoT it will further be possible to remotely make changes to these functions whenever required and have them interact without humans. A security system might determine that no humans are present and turn off lights and air conditioning systems to save money.

Smart, driverless trains have operated between terminals at Singapore airport now for some time and an Australian mining company is using smart driverless trucks to transport ore from the mine to a train, monitored from Perth, over 1,000km away. A modern car has a multitude of computer functions such

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

as roll sensors, stability control, brake assist and pre-collision systems designed to make them safer. These are all computer controlled and involve a variety of sensors 'talking' to each other and, in many cases, reacting autonomously to perceived threats. An article in the Sunday Age Newspaper (Purcell 2015) presents a scenario in which the computer of a car is 'hacked' remotely and the car taken over. It reports how two cyber-security researchers in St Louis, Missouri remotely took over control of a Jeep Cherokee. While this, and some other cases have not been malicious but done to test security, it does raise another scenario of the 'things' taking over. Would it not be better for a traffic signal to disable a car about to run a red light?

ARTIFICIAL INTELLIGENCE

There are ideas for new technology that are placed into a different category of risk, the most famous (and possibly the longest running) being the idea of artificial intelligence.

John McCarthy et al. (2006) first defined Artificial Intelligence in 1955 as: "the science and engineering of making intelligent machines". Wikipedia defines Artificial Intelligence (AI) as: "the study and design of intelligent agents, in which an intelligent agent is a system that perceives its environment and takes actions that maximize its chances of success." (Wikipedia 2015)

In a recent BBC broadcast (Cellan-Jones 2014), Professor Stephen Hawking suggested that: "The development of full artificial intelligence could spell the end of the human race" and that efforts to create thinking machines pose a threat to our very existence. Noting that the primitive forms of artificial intelligence developed to this time have already proved very useful, Hawking fears the consequences of creating something that is able to match or surpass humans. "It would take off on its own, and re-design itself at an ever increasing rate. Humans, who are limited by slow biological evolution, couldn't compete, and would be superseded." he said. (Cellan-Jones 2014)

ACTORS, NETWORKS AND THINGS: AUTONOMOUS THINGS – WARNINGS AND PREDICTIONS

New technologies should be examined for potential effects in society as a whole. To do so for the IoT is a complex task. While single technologies like tablet computers can be easily tested against 'predictions of doom', for more complex ideas such as artificial intelligence and massive networks it proves not so easy to design experiments to predict outcomes. For tablet computers dramatic newspaper reports can be quickly tested. So a report warning or dire consequences such as: "Electronic tablets like the iPad are a revolutionary educational tool and are becoming part of childhood, but should be watched carefully so that overuse doesn't lead to learning or behavioural problems, experts say." (The Australian 2012)

Is often quickly followed by testing resulting in a more reasoned response talking about the need for balance and control, such as: "It is hard to find an expert who thinks that monitored and considered tablet use is harmful." (Cocozza 2014)

If you accept that: "The notion of all physical objects being endowed with the capability to connect to such a network is fanciful, and in many cases without any justification …" (CASAGRAS 2014), the following section looks at some situations which may challenge this.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Smart Cities

An article in a recent issue of the Guardian Weekly indicates that a smart city requires the interaction of three different networks: the communications grid, the energy system and the 'logistic' Internet. The logistic Internet is able to track people and things through transport and supply systems (Mason 2015). The article then points out a problem: "who controls the project and who owns the data it generates?"

Anonymity (or Lack of It): The Demise of Privacy and Human Control

The Forbes top 100 companies list rates Google as number 39. The interesting thing about this is that, unlike the 38 companies above it (including Apple at 12 and Microsoft at 25) it does not provide financial services (like the top 6) or products. Google's (and Facebook and Twitter) main source of income is the information it provides about its users. This huge company makes most of its money by keeping track of us and selling that information to people who believe they can use the information to make us spend our money their way (Forbes 2015).

When you logon to your supermarket web site to place your grocery order the software knows who you are, what you have purchased in the past, and what you are likely to order this time. The smart phone that most of us carry is actually also a tracking device. Apart from being able to tell you the location of your closest coffee shop, it can also tell anyone with access to the data it collects where you are at any given moment and where you have been. In many cases your email provider will send you personal advertisements based on the content of your supposedly private messages. Is it possible to be anonymous anymore?

'Eyes watching you' is not the only trend potentially removing freedom of action. The recently revamped Apple conditions that that: "You agree that Licensor may collect and use technical data and related information-including but not limited to technical information about your device, system and application software, and peripherals-that is gathered periodically to facilitate the provision of software updates, product support, and other services to you (if any) related to the Licensed Application. Licensor may use this information, as long as it is in a form that does not personally identify you, to improve its products or to provide services or technologies to you." (Apple 2015) It also states that: "You agree that the iTunes Service and certain iTunes Products include security technology that limits your use of iTunes Products." (Apple 2015)

A Past, Current and Future Home Scenario

Suppose you are away from your home and running late and would really like your oven to be turned on so that it will have warmed up and be ready to heat the meal you are bringing home. In the past, before you had access to the Internet of Things you could perhaps have phoned you spouse or children at home and asked them to turn the oven on for you: Scenario-1. With access to the IoT you could now use your smartphone to turn it on for you while you are on your way home: Scenario-2. In the future it could well be possible for the things to take over and, based on what you have done in the past and perhaps on the weather, for the oven to turn itself on in anticipation: Scenario-3. Do we want the technology to assumption it knows what you need?

The CASAGRAS report offers the following ideas: "In the minimalist version of the Internet of Things these supported objects may be identified but do not 'do' anything actively, cannot communicate one with another and do not display any level of intelligence." It goes on to note thought, that: "In the

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

strongest version, object sets can be identified that communicate with each other exploiting the potential of ubiquitous computing and ubiquitous networks. It is also being seen as a vehicle for achieving actuation and control in real world applications." (CASAGRAS 2014)

SCIENCE FICTION

Another definition of the Internet of Things is: "the network formed by things/objects having identities, virtual personalities operating in smart spaces using intelligent interfaces to connect and communicate with the users, social and environmental contexts". (EPoSS 2014) The key point here is that the things have identities and virtual personalities. The following section looks at some scenarios from science fiction that just might contain something of value to this discussion.

Nineteen Eighty-Four: "Big Brother is Watching You"

Electronic surveillance aids in tracking targets and identifying suspicious activities. It is also used to detect potentially suspicious left luggage and to monitor unauthorised access. This can be coupled with automatic behaviour analysis. While this would not involve a TV camera in each room listening in and keeping track of our movements, recently in Australia, and many other countries, there has been much discussion about the collection of meta-data, particularly in relation to law enforcement.

In a recent article in the Sunday Age (Elder 2015), in relating to Orwell's (1949) book Nineteen Eighty-Four, Federal MP Adam Brandt says he doesn't equate surveillance just with use of technology, but see a difference between sharing and being spied on. Sociologist Dan Woodman: "notes that any disquiet people might have about CCTV in the city streets has been up-ended by the fact that just about everyone carries a camera … With this mixture of both confession and judgement, we have become a self-eating version of Orwell's Thought Police …". (Orwell 1949). Philosopher John Thrasher notes that we tend to think that: "there is a fundamental difference about sharing information about ourselves on Facebook or Twitter and having the Government collect metadata about who we call or what we do online." (Elder 2015). In Nineteen Eighty-Four, O'Brien, from the Ministry of Love, comments that: "The choice for mankind lies between freedom and happiness and for the great bulk of mankind, happiness is better." (Orwell 1949:275). Following from this, Thrasher suggests that this and the ability to share a multitude of thoughts over social media "have made it possible for individuals and groups to attempt to enforce their orthodoxy on those they disagree with." (Elder 2015). In this case, it is probably the case that social media has "increased our tolerance with being observed." As Winston Smith once thought: "Always the eyes watching you and the voice enveloping you. Asleep or awake, working or eating, indoors or out of doors, in the bath or in bed — no escape. Nothing was your own except the few cubic centimetres inside your skull." (Orwell 1949:29).

2001: A Space Odyssey

In this film, the screenplay of which was written by Stanley Kubrick and Arthur C. Clarke, the U.S. spacecraft *Discovery One* is bound for Jupiter (Wikipedia 2015). On the spaceship, most of the important operations are controlled by the ship's computer: HAL 9000 which its manufacturers claim is "foolproof and incapable of error". The crew know the computer as 'HAL'. At one point during the trip HAL reports

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

the imminent failure of an antenna control device but when this is checked no fault is found and HAL is thought to be in error. The crew now believe that something is wrong with the computer and agree to disconnect it, but when astronaut Frank Poole attempt to do this, HAL severs his oxygen hose and sets him adrift in space. Dave Bowman now commands HAL to "Open the pod bay doors", but HAL answers: "I'm sorry, Dave. I'm afraid I can't do that. … This mission is too important for me to allow you to jeopardize it."

In this scenario the computer considered that it knew better than the humans and proceeded to act on this by refusing a human order.

The Terminator

The fear of thinking machines is a not recent. The original use of the word 'robot' has been traced to the Karel Čapek play "Rossum's Universal Robots" in which a greedy developer unleashes the demise of humans through a robot revolution. More recently the successful movie series 'The Terminator' posits a problem caused by the creation of a satellite defence system that becomes self-aware and 'protects itself from humanity.' The script of 'Terminator 2: Judgment Day' (Cameron and Wisher 1991) contains the following passage: "The Skynet funding bill is passed. The system goes on-line August 4th, 1997. Human decisions are removed from strategic defense. Skynet begins to learn, at a geometric rate. It becomes self-aware at 2:14 a.m. eastern time, August 29. In a panic, they try to pull the plug.

SARAH: *And Skynet fights back.*
TERMINATOR: *Yes. It launches its ICBMs against their targets in Russia.*
SARAH: *Why attack Russia?*
TERMINATOR: *Because Skynet knows the Russian counter-strike will remove its enemies here.*

Before dismissing this scenario please consider the "Flash Crash" of 2010. Careful analysis of the crash has sometimes traced it to a single mistake by a human when allowing an automated trading system to make unlimited purchases – the 'thing' was accidentally given the power to crash the entire stock market! "Specifically, at 2:32 p.m., a fundamental trader used a broker algorithm to sell a total of 75,000 e-mini contracts with a notional amount of approximately $4.1 billion on the Chicago Mercantile Exchange (CME). The trade was intended to hedge an existing equity position. The trader entered the order correctly and specified an upper limit on the amount sold as a percentage of volume but did not set a price limit for the trade. As a result, price movements were magnified by a feedback loop from the volume participation settings, precipitating the actual flash crash. The CFTC/SEC report concluded that this single trade was the root cause of the flash crash." (Madhavan 2012)

Rings of Power

In an article entitled 'Wandering Things', Niehaus (2014:114) points out that "Often they function as spies who disclose the intimate secrets of their owners, because things allow 'unperceived perception." In Wagner's cycle of four epic operas titled: 'The Ring of the Nibelung' (Wagner 1876), a ring of power is created from gold stolen from the Rhine-Maidens by Alberich the Nibelung. Much of the material in Wagner's Ring Cycle was drawn from traditional legendary material of the Germanic people: from The Saga of the Volsungs, The Edda and The Nibelungenlied (Magee 2004). Niehaus (2014: 121) notes that

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

"it is evident that the category of personal property does not fit for the ring-thing …" which is claimed as their own by many including Alberich, Wotan, Siegfried and the Rhine-Maidens, in each case acting as a curse on its owner. Alberich, the ring's creator points out that the ring's master will be the ring's slave, and the ring's curse proves the truth in this. The Ring of the Nibelung has a power, and a 'life' of its own.

Also well-known are the three books of J.R.R. Tolkien's (1954): 'Lord of the Rings', in which the Dark Lord Sauron forges "One ring to rule them all, One ring to find them, One ring to bring them all and in the darkness bind them." (Tolkien 1954). The One Ring also has a 'life' of its own and when one owner loses his usefulness to it, the ring seeks another until eventually found by Bilbo the Hobbit. The One Ring has the power to make its possessor obsessed with it and to refuse to be willingly parted from it. To its possessor it becomes 'his precious' and he becomes a slave to the ring. As Niehaus (2014:121) points out "Things that prove themselves as actors exceed the logic of personal property", and the ring of the Nibelung and the One Ring of the Lord of the Rings are examples of this. In each case the ring-thing is a non-human actor that has gained a power of its own.

CONCLUSION

In this article we have traced the commonly discussed characteristics of an emerging technology, but one that is unlike any other. The Internet can be compare with the telephone system in that it is an enabling technology connecting humans. The IoT takes a potential peripheral use and transforms the Internet into a technology enabling *technologies*. This change is a fundamentally different type of technology that can be seriously compared with the science fiction scenarios involving self-aware artificial intelligences. The IoT, involving sensor input and remote activation of devices, sometimes by other devices, is a network where the level of autonomy and decision making by the devices is fundamentally different from anything we have seen before. The uptake of social media, the power of search engine companies and the ubiquity of mobile devices show that new technologies can produce change that is unexpected in its order of magnitude. An ANT approach offers unique opportunities for understanding the new technology in two distinct ways: understanding the democracy of things and seeing the IoT as a set of possible translations.

The Democracy of Things

A socio-technical approach has been found to allow the understanding of technology adoption not possible by failing to 'listen' to the technology and socio-technical actors. This becomes crucial with the IoT as purely technical actors become able to gather their own information, totally independent on humans and to 'talk among themselves' with the ability to activate functions purely on the basis of the state of the network. These new technologies have the ability to create new interactions and change the power of their interaction not through an influence over the human component of a socio-technical actor, but through their nature. Consider the home security network. WiFi connected sensors and cameras are connected to the Internet so that homeowners can be notified that movement has activated the camera. They can also be connected to the telephone system so that authorities are notified that the home has been invaded – all reasonable tools for the home owner. This same home network may also have door locking capability, be connected to security sensors around the home or be able to unlock doors. There does not have to be much imagination involved to see these devices 'voting' to close doors and call the authorities when they 'decide' that the homeowner is not available to make decisions. If one is looking for

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

an example of the possibilities of democracy without humans then the computer trading delivered stock market crashes of 6th May, 2010 are a sobering reminder of unexpected results when a system meant to do some restricted thing (place stock orders for a few investors) becomes an order of magnitude larger than the original plans and places the majority of orders.

Searching for Translations

When viewed in terms of the democracy of thing, the question "how can we consider the ways the IoT might manifest in the future?" (Phillips 2010) arises. Each application of technology that fits into the IoT has a simple plan for making life more efficient in some way. The ANT approach presumes that it is the interaction of actors and not the intent of the original designer that determines the translation of innovations.

REFERENCES

Apple. (2015). iTunes Store Terms and Conditions. Retrieved from http://www.apple.com/legal/internet-services/itunes/au/terms.html

Ashton, K. (1999, June 22). That 'Internet of Things' thing. *RFID Journal*.

Cameron, J., & Wisher, W. (1991). Terminator 2: Judgment Day. Retrieved from http://www.scifiscripts.com/scripts/t2.txt

CASAGRAS. (2014). Final Report: RFID and the Inclusive Model for the Internet of Things. Retrieved from https://docbox.etsi.org/zArchive/TISPAN/Open/IoT/low%20resolution/www.rfidglobal.eu%20CASAGRAS%20IoT%20Final%20Report%20low%20resolution.pdf

Cellan-Jones, R. (2014, December 2). Stephen Hawking warns artificial intelligence could end mankind. Retrieved from http://www.bbc.com/news/technology-30290540

Cocozza, P. (2014). Are iPads and tablets bad for young children? Retrieved from http://www.theguardian.com/society/2014/jan/08/are-tablet-computers-bad-young-children

Colitti, W., Long, N. T., DeCaro, N., & Steenhaut, K. (2014). *Embedded Web Technologies for the Internet of Things. Internet of Things: Challenges and Opportunities. Mukhopadhyay* (pp. 55–74). Heidelberg, S. C., Germany: Springer.

Elder, J. (2015, October 18). Everyone is Watching. Sunday Age (Melbourne, Fairfax) 16.

EPoSS. (2014). Definition of the Internet of Things. Retrieved from http://www.smart-systems-integration.org/public

Forbes. (2015). The World's Biggest Public Companies. Retrieved from http://www.forbes.com/global2000/list/#tab:overall

Foreshew, J. (2015). *Watch watches out for the aged*. The Australian Melbourne, News Media.

Gheorghiu, V. (1950). *The Twenty-Fifth Hour*. Ontario, Canada: Simon & Schuster Pub.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Gubbia, J., Buyyab, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems*, *29*(7), 1645–1660. doi:10.1016/j.future.2013.01.010

Haller, S. (2009). *Internet of Things: An Integral Part of the Future Internet*. Prague: SAP Research.

Lazarescu, M. T. (2014). *Internet of Things Low-Cost Long-Term Environmental Monotoring with Reusable Wireless Sensor Network Platform. Internet of Things: Challenges and Opportunities. Mukhopadhyay* (pp. 169–196). Heidelberg, S. C., Germany: Springer.

Madhavan, A. (2012). Exchange-traded funds, market structure, and the flash crash. *Financial Analysts Journal*, *68*(4), 20–35. doi:10.2469/faj.v68.n4.6

Magee, E. (Ed.). (2004). *Legends of the Ring*. London: The Folio Society.

Mason, P. (2015). We can't allow tech giants to rule. *The Guardian Weekly* (London).

McCarthy, J., Minsky, M., Rochester, N., & Shannon, C. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence (1955). *AI Magazine*, *27*(4), 12–14.

Mukhopadhyay, S. C., & Suryadevara, N. K. (2014). *Internet of Things: Challenges and Opportunities. Internet of Things: Challenges and Opportunities. Mukhopadhyay* (pp. 1–17). Heidelberg, S. C., Germany: Springer.

Niehaus, M. (2014). *Wandering Things - Stories. Le Sujet De L'Acteur - An Anthropological Outlook on Actor-Network Theory. Kapriev, G., Roussel, M. and Tchalakov, I* (pp. 109–129). Paderborn, Germany: Wilhelm Fink.

Orwell, G. (1949). *Nineteeen Eighty-Four*. Great Britain: Martin Secker & Warburg.

Pererez, I. C. Barbolla and A. M. Bernardos (2014). Exploring Major Architectural Aspects of the Web of Things. In S. C. Mukhopadhyay (Ed.), Internet of Things: Challenges and Opportunities (pp. 19-54). Heidelberg, Germany, Springer:

Phillips, M. (2010, May 11). Nasdaq: Here's Our Timeline of the Flash Crash. *The Wall Street Journal*.

Purcell, A. (2015). *Driver beware: your car could be hacked while you're in it*. Melbourne, Fairfax: Sunday Age.

Song, Z., Lazarescu, M. T., Tomasi, R., Lavagno, L., & Spirito, M. A. (2014). *High Level Internet of Things Applications Development Using Wireless Sensor Networks. In S. C. Mukhopadhyay* (Ed.), *Internet of Things: Challenges and Opportunities.* (pp. 75–110). Heidelberg, S. C., Germany: Springer. doi:10.1007/978-3-319-04223-7_4

Sundmaeker, H., Guillemin, P., Friess, P., & Woelfflé, S. (2010). Vision and Challenges for Realising the Internet of Things. Brussels, European Commission - Information Society and Media DG.

The Australian. (2012). Parents warned of side effects of tablet computer overuse on children. Retrieved from http://www.theaustralian.com.au/news/parents-warned-of-side-effects-of-tablet-computer-overuse-on-children/story-e6frg6n6-1226314156128

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Tolkien, J. R. R. (1954). *The Fellowship of the Ring*. England: Allen & Unwin.

Unnithan, C. (2014). *Examining Innovation Translation of RFID Technology in Australian Hospitals through a Lens Informed by Actor-Network Theory PhD*. Victoria University.

Unnithan, C., Nguyen, L., Fraunholz, B., & Tatnall, A. (2013). RFID translation into Australian Hospitals: An exploration through Actor-Network Theoretical Lens. *Proceedings of the International Conference on Information Society (i-society 2013)*, Toronto, University of Toronto, Canada.

Unnithan, C., & Tatnall, A. (2014). Actor-Network Theory (ANT) based visualisation of Socio-Technical Facets of RFID Technology Translation: An Australian Hospital Scenario. *International Journal of Actor-Network Theory and Technological Innovation*, *6*(1), 31–53. doi:10.4018/ijantti.2014010103

Wagner, R. (1876). *Der Ring des Nibelungen*. Bayreuth, Germany: Bayreuth Festspielhaus.

Weiser, M., Gold, R., & Brown, J. S. (1999). The Origins of Ubiquitous Computing Research at PARC in the late 1980s. *IBM Systems Journal*, *38*(4), 693–696. doi:10.1147j.384.0693

Wikipedia. (2015). 2001: A Space Odyssey (film). Retrieved from https://en.wikipedia.org/wiki/2001:_A_Space_Odyssey_(film)

Wikipedia. (2015, October). Internet of Things. Retrieved from https://en.wikipedia.org/wiki/Internet_of_Things

Wikipedia. (2015, October). Smartwatch. Retrieved from https://en.wikipedia.org/wiki/Smartwatch

Wikipedia. (2015, October). Artificial intelligence. Retrieved from https://en.wikipedia.org/wiki/Artificial_intelligence

This research was previously published in the International Journal of Actor-Network Theory and Technological Innovation (IJANTTI), 7(4); edited by Mehdi Khosrow-Pour, D.B.A.; pages 56-67; copyright year 2015 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 63
The Rise of Artificial Intelligence:
Its Impact on Labor Market and Beyond

Robert Niewiadomski
NYCDOE, USA

Dennis Anderson
St. Francis College, USA

ABSTRACT

Our inventions defined the work we engaged in for centuries; created new industries and employment opportunities around them. They, however, had often unforeseen consequences that affected the way we lived, interacted with each other, and redefined our societal rules. The established narration portrays the impact of major technological leaps in civilization on employment as temporary disruptions: Many finds themselves without employment taken away from them by efficient, laborsaving inventions, but, in the long run, through gradual adaptations, improved education and gaining higher qualifications, everyone benefits. In this chapter, the authors explore the impact of the rapid expansion of artificial intelligence (AI) in relations to the labor market. The authors argue that this rather optimistic, even naïve scenario, collapses while confronted with the exponential growth of AI; in particular, with the potential arrival of syneoids – robotic forms of "strong AI" possessing, or even exceeding, the full range of human cognitive abilities.

FROM GALATEA AND GOLEMS TO FUTURE SYNEOIDS

We know that blind evolutionary processes can produce human-level general intelligence, since they have already done so at least once. Evolutionary processes with foresight—that is, genetic programs designed and guided by an intelligent human programmer—should be able to achieve a similar outcome with far greater efficiency. (Nick Bostrom, Superintelligence, Paths, Dangers, Strategies)

DOI: 10.4018/978-1-7998-0951-7.ch063

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The unique urge to create an *artificial man* has a long tradition dating back to the dawn of Western intellectual history, "Hephaestus, the master craftsman, grants a human voice (…) to his golden mechanical handmaidens (Gera, 2003, p. 114)" and Jewish folklore includes many versions of golems –anthropomorphic beings created from clay. Myths, religions, and popular culture displayed this obsession with varying intensity thorough history. Pamela McCorduck (1979), who investigated this phenomenon, concludes that, as humans, we have been engaged in this peculiar form of self-reproduction by attempting to fulfill the urgent desire, bypassing the ordinary means, in order to recreate what is the essential to us. Contemporary incarnation of this ancient desire can be observed in robotics.

Currently, we are on the verge of an unprecedented technological revolution involving intelligent robots powered by artificial intelligence (AI). It is important to note that what we are experiencing at this stage – Google self-driving cars, Apple Siri, Google Photo Search, robots exhibiting behavior similar to human- are examples of so called "weak AI (WAI)." The concept asserts that machines could act *as if* they were intelligent. It is a sort of limited intelligence. To the contrary, "strong AI" (SAI) is the higher level of AI, often referred to as "artificial general intelligence" (AGI) –it entails the possibility of machines *actually* thinking (Russell & Norvig, 1995). According to this view SAI would possess the full range of human cognitive abilities. Such AI, would, as predicted, experience exponential growth, and quite swiftly reach a level exceeding human capacities – a point often referred to as "singularity." If we assume a physicalist position regarding the nature of human consciousness, we ought to also seriously consider the possibility that SAI would be, at some point, able to genuinely experience subjective mental states such as consciousness. The emergence of consciousness in SAI could be perhaps an incremental process, similar to varying degrees of mental capacities in animals. The terminology attempting to capture this phenomenon varies and it does not quite reflect its full nature and scope. For the purpose of our investigation, we will refer to it as the *syneoid* –a term derived from Greek *συνειδητός* (corresponding to English *conscious*) and the suffix *–oid* suggesting "likeness" or "form of" as well as bearing resemblance to android (human-like robot).

Throughout history, technological advancements have impacted how we worked, lived, and died. The canonical narration pertaining to the impact of major technological leaps in civilization on employment, particularly the paradigmatic Industrial Revolution, follows the familiar pattern: they lead to substantial shifts in types of jobs, affecting perhaps a large but limited area of employment. Certainly, the total number of jobs was not affected and, if anything, technological revolutions expanded employment opportunities by opening new areas of exploration. In a relatively short term, this was a messy process; many lost jobs taken away from them by efficient, laborsaving inventions, but, in the long run, through gradual adaptations, improved education and higher qualification, everyone benefited. Thus, in essence, a technological revolution has been typically painted, in spite of its shortcomings and damages, as a disruptor and the true engine of human progress.

The arrival of the syneoid, however, would have staggering consequences for the future of human employment and beyond. In this chapter, we intend to explore the impact of the rapid expansion of SAI in relation to the labor market. We argue that this rather optimistic, even naïve scenario, collapses while confronted with the explosion of SAI, as none of the historical extrapolations will apply to the processes that are currently shaping up. The history might have had repeated itself with several technological revolutionary cycles but, this time, we will be confronted with a radical new factor, for the arrival of which we are utterly unprepared, and as we are about to see, in principle, we cannot.

Interestingly, although the research on SAI is growing fast, it did not capture the interest of general public beyond some portrayals in science fiction. This does not translate, however, into a serious public debate on the speed and consequences of this technological revolution *in statu nascendi*. Popular imagi-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

nation seems to be captivated by other occurrences. The general public and both established and aspiring politicians are typically more concerned, and not without a good reason, with the issues of climate change, the rise of jihadist extremism, and ever-widening income gap. Politicians and experts turn to history looking for solutions, and while some of them proved to yield positive results in the past, none will work to alleviate the incoming next wave of unemployment due to automation, since its causes are drastically different from anything from what we have experienced so far.

Will History Repeat Itself? Not this Time

In the past, the slower pace of transformations allowed enough time to adapt and balance the labor supply and demand. By the dawn of 19th century, technological advancements decreased the need for agricultural labor drastically. The excess labor supply was shifted to manufacturing and commerce. This transfer first happened on a small scale, typically in small workshops or homes. It is interesting to note that this was the beginning of wealth transfer from agriculture to industry. Most these workshops disappeared with the arrival of steam engine brought about by the Industrial Revolution, which marked a dramatic change in process of production, particularly -a division of labor, mechanization, and specialized workforce. The driving force behind automation has been, and still is, the desire to improve the profit margin by cutting the workforce perceived as cumbersome and inefficient.

The modern factory also created a new type of workers -precisely defined components in steam-driven machinery. As industrial instruments became more sophisticated, the demand for more skilled labor force increased. The low-skilled ones were pushed out yet again. The investment in education in the 19th century provided a supply of qualified labor. This trend continued into the 20th century as post-secondary education became increasingly common. WWII fueled and accelerated the development of complex machines and systems as created the global supply chain management. After the war, workers in the West enjoyed unprecedented prosperity that provided them with a secure pathway to middle class. Even the initial introduction of computers had no immediate impact on job market. Obviously, mathematical calculations done by hand by numerous people became obsolete but, in time, the invention greatly increased the productivity of others and created a plethora of computer-related jobs. Many other technical innovations had similar effects. New machinery displaced handicraft producers across numerous industries, from textiles to metalworking.

Naturally, the mere possibility of automation didn't mean that it always made sense to implement it. Often, it was more profitable to outsource jobs and take advantage of cheap labor in developing counties. This was accelerated by globalization. Western workers found themselves struggling yet again, facing increasing competition from both machines and cheap labor overseas. Since the 1950's employment in manufacturing in the U.S. declined sharply due to cheap imports and rising cost of labor. It was inevitable that most available jobs moved to service industry. Eventually, we transitioned into a consumption-based economy where many jobs shifted to service industries selling goods and delivering services at a massive scale. However, service industry already displays similar techniques of maximization of output on the expense of labor force as we have seen in manufacturing.

Eventually, the deterioration of employment opportunities expanded to highly skilled college educated workforce as well. As Martin Ford argues, AI ultimately will lead to evaporation of most blue and white collar jobs alike leading to further acceleration of the process of middle-class shrinking. Middle and lower-middle classes are already experiencing tremendous costs of health care and education. Ford scenarios predict massive unemployment, and a complete breakdown of consumer-based economy itself (Ford, 2009).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

As of now, WAI led to the emergence of a radically different sort of technologies capable of performing tasks previously done by humans: quite sophisticated and speedy reasoning, judgment, and perception. One would be perhaps tempted to describing theses capacities as constituting a sort of agency. As Jeremy Howard (2014) points out, "Computers right now can do the things that humans spend most of their time being paid to do, so now's the time to start thinking about how we're going to adjust our social structures and economic structures to be aware of this new reality." This rapid automation of physical, mechanical or repetitive jobs has already resulted in the elimination of various occupations and the process only just started. Job market needs a balance of supply and demand but when human labor is no longer required, this balance no longer exists as humans are no longer part of the labor market. At this historical junction, the fundamental question then is: Will artificial intelligence reduce the need for human toil and, as a result, cause technological unemployment? There are still two opposing main answers to this question as articulated by Nils Nilsson (1984, p. 5).

Some claim that AI is not really very different from other technologies that have supported automation and increased productivity (…) Like them, AI may also lead ultimately to an expanding economy with a concomitant expansion of employment opportunities. At worst, (…) there will be some, perhaps even substantial shifts in the types of jobs, but certainly no overall reduction in the total number of jobs. In my opinion, however, such an outcome is based on an overly conservative appraisal of the real potential of artificial intelligence. Others accept a rather strong hypothesis with regard to AI -one that sets AI far apart from previous labor-saving technologies. Quite simply, this hypothesis affirms that anything people can do, AI can do as well. Certainly AI has not yet achieved human-level performance in many important functions, but many AI scientists believe that artificial intelligence inevitably will equal and surpass human mental abilities.

The "optimistic scenario" as articulated above, is based on the naïve extrapolation of historical experience. In a nutshell, it basically claims that AI is no different from other technological advances we have witnessed in the past. It will, just like them, lead to the expansion of employment opportunities in other fields. This shift might be quite unpleasant in a short run, the human cost conceivably high, but, in the end, people, who are pushed out from their jobs by automation will eventually find employment elsewhere. To believe otherwise –according to the optimists – is erroneous. Keynes (1931) also wrote about the causes of widespread technological unemployment occurring due to discoveries capable of economizing the use of labor that are surpassing the pace at which we can find new uses for labor. Nevertheless, he concluded that technological unemployment is only, as he put it, a temporary phase of maladjustment that will be solved in a long run.

Those, who support this view, argue that those concerns pertaining to long–lasting or even permanent technological unemployment result from the "Luddite fallacy" chiefly because they fail to recognize compensation effects. The term "Luddite fallacy" was coined by drawing a comparison to the 19th century reasoning of the English textile workers, who were concerned with the machinery threatening their livelihood, introduced during the Industrial Revolution, and resorted to violent uprisings. The optimists believed that textile workers were fundamentally mistaken, since the technological progress will eventually lead to the increase of wealth of the society as a whole, and that the advances, as mentioned above, will lead to the creation of new employment opportunities. In other words, the "Luddite fallacy"

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

proponents also tend to point out that the worries about long-lasting or permanent technological unemployment are unwarranted because there is no such thing as a finite amount of work available to us and new occupational opportunities will always emerge that only humans will be able to perform.

On the surface, this assumption might seem justified and intuitive, since we tend to extrapolate from the past. Historical events, however, are not necessary a solid predictor of the future, particularly when we are confronted with a radically novel phenomenon. The belief in "Luddite fallacy" constituted a prevalent view of mainstream economists on the problem of technological unemployment until recently and many still subscribed to it.

Nevertheless, some cracks in this reasoning are easy to spot. Assuming, for the sake of the argument, that the technological progress, in fact, will increase the aggregate wealth of the society, it is not obvious that the distribution of the wealth will be equitable. It goes without saying that distribution of wealth matters a great deal. For example, in a technocratic society, such as ours, an enormous amount of wealth is concentrated in the hands of an extremely small group. This does not mean, however, that everyone else benefits through mechanisms of job-creation. We can see this in similar mechanism – so-called trickle down economy. The correlation between lowering taxes for the wealthiest and job -creation is highly arguable. Thus, supposed increase in overall wealth does not automatically translate into a lower rate of unemployment –as enthusiasts of trickle down economy would like to believe.

In addition, according to the optimistic narrative, long-lasting technological unemployment typically requires a massive effort to educate the population. Again, it is a lesson learned from the Industrial Revolution. In short, education is supposed to be the key to solving this issue. As Paul Krugman (2013, para. 4) puts it by revisiting the problems faced by, the previously ridiculed, 19th century textile workers:

Until recently, the conventional wisdom about the effects of technology on workers was, in a way, comforting. Clearly, many workers weren't sharing fully — or, in many cases, at all — in the benefits of rising productivity; instead, the bulk of the gains were going to a minority of the work force. But this, the story went, was because modern technology was raising the demand for highly educated workers while reducing the demand for less educated workers. And the solution was more education.

Indeed, mostly free and compulsory public education emerged in the West as a result of ideas of the Enlightenment a feeding pipe for factories of the industrial age. This model of "keeping up "with the technology seemed to work fairly well for some portions of the population, at least until recently. Several factors are at play here: the outdated model of education that is simply too specialized and often provides training that quickly lacks relevance on the labor market. After all, it is difficult to blame institution of higher education for not predicting what skills will be desirable in the rapidly shifting labor landscape even in not-to-distant future. The second factor pertains to the massive democratization of higher education experienced virtually in all developed countries. This educational inflation, occasionally leading to outburst of frustration comparable to those of mentioned textile workers, effectively cancels out any competitive edge previously enjoyed by college degree holders in the past,

Today (…) a much darker picture of the effects of technology on labor is emerging. In this picture, highly educated workers are as likely as less educated workers to find themselves displaced and devalued, and pushing for more education may create as many problems as it solves. (…) So should workers simply be

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

prepared to acquire new skills? The woolworkers of 18th-century Leeds addressed this issue back in 1786: "Who will maintain our families, whilst we undertake the arduous task" of learning a new trade? Also, they asked, what will happen if the new trade, in turn, gets devalued by further technological advance?

And the modern counterparts of those woolworkers might well ask further, what will happen to us if, like so many students, we go deep into debt to acquire the skills we're told we need, only to learn that the economy no longer wants those skills?

Education, then, is no longer the answer to rising inequality, if it ever was (Krugman, 2013, para. 6).

Putting aside the question of what kind of training and qualifications would be needed to access these emerging employment opportunities and how many individual would have an actual chance to do so etc., there are more serious problems with this line of reasoning. Even individuals, who are flexible and well adjusted to the job market might not be able to keep up with the speed of technological acceleration. This is how Krugman (2013, para. 1-2) explain this phenomenon:

In 1786, the cloth workers of Leeds, a wool-industry center in northern England, issued a protest against the growing use of "scribbling" machines, which were taking over a task formerly performed by skilled labor. "How are those men, thus thrown out of employ to provide for their families?" (…) Those weren't foolish questions. Mechanization eventually — that is, after a couple of generations — led to a broad rise in British living standards. But it's far from clear whether typical workers reaped any benefits during the early stages of the Industrial Revolution; many workers were clearly hurt. And often the workers hurt most were those who had, with effort, acquired valuable skills — only to find those skills suddenly devalued.

More fundamentally, there is another crucial point that gets lost in the current debates. Even if some number of very highly specialized occupations will be created, with the arrival of syneoids, "this kind of work can also be performed by AI devices without necessarily implying more jobs for humans (Nilsson, Ibid.)."

The above set of issues gradually begins to gain traction in the high strata of the more forward-thinking individuals. According to the posted agenda of the Davos World Economic Forum in 2016 the incoming shift, framed as the Fourth Industrial Revolution, is recognized a paradigmatic change that will affect virtually every sphere of human life. To the question why this revolution would be any different from the previous ones we experienced, Klaus Schwab delivers the answer with blunt honesty,

The first industrial revolution began in the latter half of the 18th century, and brought us mechanical production through the steam engine, cotton spinning and then railroads. The second industrial revolution took place around the turn of the 20th century and brought mass production through assembly lines and electrification. The third industrial revolution was the computer revolution, starting in the 1960s that brought us the mainframe and then personal computing, as well as the Internet. Today we are in the midst of the Fourth Industrial Revolution, which will affect governments, businesses and economies in very substantial ways. We should not underestimate the change ahead of us, as there are at least three differences between this revolution and the previous ones (Schwab, 2015, para. 4-5).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Schwab articulates three major components that differentiate this new fundamental change from previous technological revolutions. In a nutshell, the first one has to do with its speed. The current change is happening with an unprecedented acceleration, as opposed to the previous ones that took decades or more to fully unfold. Secondly, the Fourth Industrial Revolution is not limited only to one area. It incorporates such elements as mobile networks, nanotechnology, brain research, 3D printing, materials science, computing, networks, and "the interplay between all of these. (…) In addition, the accessibility and affordability of complex technologies will spread them farther and faster (Ibid. para. 6-8)." The third difference, according to Schwab, pertains to the innovation of the entire system, as opposed to one product or area. The phenomenon of "sharing economy" is only one of many manifestations of that aspect.

As mentioned, the impact of this paradigmatic shift will be profound, and will affect the business model, the ways governments operate, and last but not least, the future usefulness of skills and labor. As Fon Mathuros (2015, para. 2) points out,

Earlier Industrial Revolutions advanced human progress through new forms of power generation, mass production and information processing. Building on a ubiquitous and mobile internet, smaller, cheaper and more powerful sensors, as well as artificial intelligence and machine learning, the Fourth Industrial Revolution is distinct in the speed, scale and force at which it transforms entire systems of production, distribution, consumption – and possibly the very essence of human nature.

The Davos Forum organizers, although recognizing the magnitude of imminent shift, tend to portray the Fourth Industrial Revolution in cautiously optimistic terms as perhaps one of the greatest challenges we will encounter, and yet, a possible opportunity. Nevertheless, they cannot disregard that two aspects of human life will be affected most significantly by the future changes –the value and significance of human work and profound changes in social structures.

THE NEW CAMBRIAN EXPLOSION

In a peculiar way, some aspects of the development of robotics mimic the natural processes of evolution. Gill Pratt (2015, p. 51) draws an intriguing comparison between the two.

About half a billion years ago, life on earth experienced a short period of very rapid diversification called the "Cambrian Explosion." Many theories have been proposed for the cause of the Cambrian Explosion, with one of the most provocative being the evolution of vision, which allowed animals to dramatically increase their ability to hunt and find mates (…) Today, technological developments on several fronts are fomenting a similar explosion in the diversification and applicability of robotics. Many of the base hardware technologies on which robots depend—particularly computing, data storage, and communications—have been improving at exponential growth rates. Two newly blossoming technologies—"Cloud Robotics" and "Deep Learning"—could leverage these base technologies in a virtuous cycle of explosive growth. In Cloud Robotics (…) every robot learns from the experiences of all robots, which leads to rapid growth of robot competence, particularly as the number of robots grows. Deep Learning algorithms are a method for robots to learn and generalize their associations based on very large (and

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

often cloud-based) "training sets" that typically include millions of examples. (…) One of the robotic capabilities recently enabled by these combined technologies is vision—the same capability that may have played a leading role in the Cambrian Explosion.

Initially, AI and robotics were separate fields, -computer science and engineering, respectively. Currently, it is the precisely the combination of the two -the robotic embodiment of AI- that sets the trend for the future. Particularly, Deep Learning (DL), which is an algorithm-based method of "teaching a machine" how to learn from massive sample sets, is one on the major mechanism driving forces. Current applications of DL already include automatic speech and image recognition, natural language processing, transforming images into description, discovery of drugs and toxicity, medical diagnostics, among others. DL certainly has the potential to become a stepping-stone to creating syneoids, which through exponential growth would rapidly exceed any human faculties to a degree unimaginable. Ray Kurzweil, who shares Searle's view that an appropriately programmed computer is a mind, makes a prediction that the equivalent capacity of one human brain will be available on desktop computers by 2020. Moreover, when AI starts to outstrip the collective total of all human intelligence, Kurzweil argues, humanity will enter, what he calls, the Singularity, marking a point at which change is so radical that it is no longer predictable (as cited in Wallach, p. 57).

The kinds of operations that WAI is capable of performing now cannot be called thinking in the traditional sense, provided that what we mean by thinking, by definition, only takes place inside the human cranium (McCorduck, 2004) Nevertheless, many consider "making difficult judgments, the kind usually left to experts, choosing among plausible alternatives, and acting on those choices (Ibid.)" a form of thinking. McCorduck further elaborates, "Along with most people in AI, I consider what artificial intelligences do as a form of thinking, though I agree that these programs don't think just like human beings do, for the most part. I'm not sure that's even desirable. Why would we want AIs if all we want is human-level intelligence? There are plenty of humans on the planet. The field's big project is to make intelligence that exceed our own. As these programs come into our lives in more ways, we'll need programs that can explain their reasoning to us before we accept their decisions (Ibid.)."

Certainly, it is fair to say that AI and robotics have not crossed yet the level of singularity and thus reached the level of syneoids. However, this could be just a matter of time. Once they reach that stage, all jobs can be automated including key decisions. No human will be needed for manufacturing, services, knowledge, or even creative processes. Thus, "Luddite fallacy", described before, may not be, in the end, a fallacy at all. This time around we simply might not be able to "catch up." It is precisely because human gradual learning cannot possibly be a competition to exponential growth of Deep Learning as Jeremy Howard (2014) explains,

80 percent of the world's employment in the developed world is stuff that computers have just learned how to do. What does that mean? Well, it'll be fine. They'll be replaced by other jobs. For example, there will be more jobs for data scientists. Well, not really. It doesn't take data scientists very long to build these things. For example, these four algorithms were all built by the same guy. So if you think, oh, it's all happened before, we've seen the results in the past of when new things come along and they get replaced by new jobs, what are these new jobs going to be? It's very hard for us to estimate this, because human performance grows at this gradual rate, but we now have a system, deep learning, that we know actually grows in capability exponentially.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In light of the inevitable systemic technological unemployment, experts are constantly considering different solutions ranging from quite unrealistic ideas of banning innovation to various forms of wealth redistribution. It is not our intention here to examine the feasibility of these proposals. The reasons for that choice are quite simple: stopping the entire process is practically impossible, and once syneoids arrive, humans won't make any decisions regarding that matter alone, as we will have to work it out with "others", presumably far more intelligent and capable than us. They might be very well able to colonize places inaccessible to human because of our limitations such us the need of oxygen, water, and food.

Even though, we are not there yet, the loss of jobs due to the already available WAI will be staggering in a relatively short time. A very conservative study conducted at Oxford University by Frey and Osborne (2013) examined the effects of computerization alone on 702 detailed occupations. The authors used a Gaussian process classifier and determined that "about 47 percent of total U.S. employment is at risk (Ibid. p.1)" of computerization in a decade or two. As they point out, "The impact of computerisation on labour market outcomes is well-established in the literature, documenting the decline of employment in routine intensive occupations – i.e. occupations mainly consisting of tasks following well-defined procedures that can easily be performed by sophisticated algorithms. (…) Following recent technological advances, however, computerisation is now spreading to domains commonly defined as non-routine (Ibid. pp.2-15)." Frey and Osborne further argue that even though Machine Learning (ML) and Mobile Robotics (MR) enable computerization across large areas of non-routine tasks, there exist some "engendering bottlenecks" inhabiting total de-humanization of labor.

Beyond these bottlenecks, however, we argue that it is largely already technologically possible to automate almost any task, provided that sufficient amounts of data are gathered for pattern recognition. Our model thus predicts that the pace at which these bottlenecks can be overcome will determine the extent of computerisation in the twenty-first century. Hence, in short, while the task model predicts that computers for labour substitution will be confined to routine tasks, our model predicts that computerisation can be extended to any non-routine task that is not subject to any engineering bottlenecks to computerisation. These bottlenecks thus set the boundaries for the computerisation of non-routine tasks (Ibid. p.23).

They identify three major "bottlenecks." The first one, perception and manipulation, in essence, pertains to the abilities to make precisely coordinated movements. This "bottleneck" seems to pose the least of the problem for computerization and robotization, since human body itself can be seen as a biological machine that had a very long to evolve and adapt to perform very complex ways of manipulating the environment. In principle, however, there is nothing, besides technological and perhaps financial limitations, that would prevent replication of these abilities. The authors admit that, "with incremental technological improvements, the comparative advantage of human labour in perception and manipulation tasks could eventually diminish (Ibid. p. 39)." The second "bottleneck" indeed poses a serious challenge for AI at this stage. It has to do with creative intelligence –faculties we typically associate with solving unusual problems, artistic production etc. The third one, social intelligence, pertains to the ability of being aware of other's reactions and understanding why they react in a given way; the ability to persuade others, and, finally, with a broadly understood, emotional support (Ibid. p. 31).

This last obstacle is indeed unlikely to be solved by WAI at its present stage. Thus, the study reports that, "generalist occupations requiring knowledge of human heuristics, and specialist occupations involving the development of novel ideas and artifacts, are the least susceptible to computerization (Ibid. p. 40)." Based on these technological "bottlenecks", their model assigns even specific probabilities

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

that computerization will lead to job losses within about the next two decades. And, so they estimate that that probability would be 0.99 for telemarketers, 0.55 for commercial pilots, and only 0.003 for recreational therapists, presumably because the latter one would require technological overcoming of the social intelligence "bottleneck."

It is important to keep in mind, that Fry and Osborne's analysis does rely on the technological "bottlenecks" as they present themselves today. The history of technological advances, however, can be largely perceived as a struggle, astonishingly successful, of overcoming such "bottlenecks." Thus, there is no reason to think that these current, undoubtedly serious "bottlenecks," will become obsolete once the technology moves beyond the stage of WAI towards syneoids.

Again, skeptics find it difficult to accept that syneoids could completely fill every available niche in the labor market and make human skills virtually obsolete. They believe that there is something uniquely special about human capacities, whether be intellectual or emotional that there will be always a margin of tasks syneoids simply won't be able to do, or people would prefer them being done by other humans. Those occupations would require human sensitivity that can satisfy deep interpersonal needs. This would involve such activities as making decisions in a court of law because, supposedly, we would prefer a person to be accountable for such decisions. The same would apply to CEO's, generals, military and government leaders at every level.

This could be the case if syneoids would exceed human capacities only intellectually, but as we stipulated, they would also become decision making fully autonomous agents, with whom we would have to engage in interactions involving making moral choices.

Schwab (2015, para. 13-15) gestures towards this issue of syneoids and humans living side by side by sketching two compelling scenarios,

Very importantly, will the Fourth Industrial Revolution have a human heart and soul? Again, imagine living in the future, when robots and humans live side by side. Who would you trust, human or robot? Concretely:

- *If you were told you had a life-threatening illness and a human doctor prescribed treatment regime A, while an artificially intelligent robot prescribed treatment regime B, which treatment regime would you follow?*
- *If you were falsely accused of a crime, would you rather be tried by a human judge or by an AI judge?*

It is obvious that Schwab scratches just the surface of the possible issues of the moral aspect of interactions with syneoids. For some this very prospect of such interactions and the moral uncertainty connected to it are frightening ideas. Perhaps, as some prominent figures such as Stephen Hawking and Elon Musk, argue, it would be better for us to stop the expansion of AI altogether, particularly with respect to autonomous warfare agents (Musk et al.). Some postulate that our efforts should be concentrated on designing these agents in such way "so that they honor the broader set of values and laws humans demand of human moral (Wallach, p. 6)."

However, at this point, it is far too late to ask the question whether the humanity needs intelligent machines making moral decisions. Machines are already making decisions with moral consequences. As mentioned, machines, whether intelligent or not, even tools, had a moral impact on our lives and,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

in fact, they always had. James Moor proposed a hierarchical model that categorizes different levels of machine's ethical agency. At the very bottom are, what he refers to, "ethical impact agents." This category includes "any machine that can be evaluated for its ethical consequences (as cited in Wallach, p. 33-34)." The next level Moore calls "implicit ethical agents" – "machines whose designers have made an effort to design them so that they don't have negative ethical effects, by addressing safety and critical reliability concerns during the design process (Ibid.)." "Explicit ethical agents" occupy the third level of Moor's hierarchy. These are machines that can reason about ethics using ethical categories as part of their internal programming. Setting proper ramifications for the design of such "implicit ethical agents" and establishing an appropriate ethical code for their decision-making is currently the main concern of the field of computer ethics. Everything that lies beyond these three levels Moor considers "full ethical agents" –machines that can make "explicit moral judgments and are generally quite competent in justifying such decisions. This level of performance is often presumed to require a capacity for consciousness, intentionality, and free will. If any of these three is lacking in a human context, then the person's moral agency and legal culpability comes into question (Ibid)." This highest level of full ethical agents matches our definition of a syneoid. Naturally, it would be quite naïve to think that syneoids would necessarily share our moral standards and sensitivities. Syneoid's morality could go much further beyond anything we can envision based on our anthropological imagination and knowledge of varying moral standards in human cultures; it could be a set of norms radically alien in general, based on different preference and goals.

WHY CONSCIOUSNESS MATTERS?

Some philosophers and scientists doubt the very possibility of artificial "full ethical agents" precisely because the idea of them possessing consciousness, intentionality, and free will seem uniquely human. From all three –consciousness is probably considered to be at the core of human personhood, which is the only point of reference of personhood we have available. In order to understand why consciousness matters in the context of syneoids, we should go back to the18th century materialist philosophy. In reaction to the previous substance dualism, La Mettrie, in his *L'homme Machine*, extended to human beings Descartes' argument that animals are automatons. Karl Popper (1978, pp. 224-225) was convinced that La Mettrie's idea gained a greater reinforcement by the theory of evolution.

… The doctrine that man is a machine was argued most forcefully and seriously in 1751, long before the theory of evolution became generally accepted, by de Lamettrie; and the theory of evolution gave the problem an even sharper edge, by suggesting that there may be no clear distinction between living matter and dead matter.

In principle, the emergence of conscious syneoid is plausible provided that few premises that we would need to commit to are true. One of them is an ontological assumption, rooted in neuroscience, concerning the nature of consciousness. The brain itself –as the argument runs- is a naturally occurring material object that emerged through evolutionary processes and as such, at least in principle, is replicable with sufficiently advanced technology. Secondly, mental states are essentially brain states. Thus, consciousness supervenes on the brain and, just as the brain, it is a natural phenomenon that is a product of a very long incremental evolutionary process. Thirdly, we have to commit to yet another premise that

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

neurons are not the only material that can create a mental state (Dennett, 1991). Therefore, it is just a matter of time that we, as conscious beings, will be able to imitate it by some means that might not be necessarily biological in nature.

This argument is further supported by "Artificial Life" projects that essentially see biology as a naturally occurring information technology. That "technology" could be reversed-engineered and synthesized in other kinds of technologies, again, not necessarily biological ones (Sullins, 2014). Consciousness would, therefore, be a product of this naturally occurring information technology. It is important to make another observation –consciousness, after all, is present, to varying degrees, in different species. In short, if we, as humans, are naturally occurring robots, and just so happened we possess consciousness, then there is nothing, in principle, preventing replication of this natural phenomenon artificially.

There is obviously the question of how would we even know that we have created a conscious machine, or that a conscious machine emerged through, for example, processes similar to Deep Learning? The famous Turing Test cannot answer this question as demonstrated by Searle in his "Chinese room" thought experiment. One might even argue that perhaps a consciousness test, in principle, cannot be done if we set the testing standard unreasonably high. After all, how can we be absolutely sure that other human fellows have mental states or just behave as if they have them? Such high requirements might lead us towards metaphysical solipsism – a doctrine with very vague practical consequences, and as such not very useful in our investigation.

One can obviously stipulate that consciousness is a something that occurs in the brain alone, and then, by analogy, for example, make inferences that other humans possess that property as well. It seems fair to say that most people use this analogy or symmetry principle, in their common sense approach to the issue of consciousness of others. This can be further supported by MRI imaging showing a correlation between changes in brain activity with reports of subjective mental states by the subject. In fact, similar methods can be used to investigate the mental states of individuals who are unable to communicate in any way. These are obviously well established and perfectly sound scientific practices, and yet strictly speaking, they can only provide indirect proofs that others possess consciousness. Another approach to solving the problem is based on common ancestry –the offspring, as a matter of rule, should display capabilities similar to those of those of its ancestors. This line of reasoning is just another way of making inferences by analogy. However, it is easy to spot a weak point of these common sense based reasoning. Provided that consciousness develops artificially and will be neither brain-based, not biologically related to us, none of these common sense reasoning by analogy will apply. In case there will be no way to discern if the intelligent machine *really* possesses consciousness, or is simulating it, one can obviously stick to the believe that only humans have it by definition, but that would be terribly arbitrary, and ultimately unjustified. Ultimately, if syneoids displayed mental and cognitive capabilities that are indiscernible from those displayed by humans that we "know" that they possess consciousness, it would be unclear, to say at least, on what basis we could conclude that syneoids do not have consciousness after all.

The more fundamental question is obviously: Why would it even matter to know whether syneoids possess consciousness or they do not? With regards to the issue of technological unemployment the answer is – it matters. One could obviously argue that such question is superfluous since super intelligent "zombies" could replace us easily, and the issue whether they are conscious has nothing to do with that. However, going back to the previous issue that people might prefer for certain services or occupations to be performed by human persons only precisely because they are conscious beings, it becomes clear how artificially generated consciousness would eliminate this need as well. Thus, there would be virtually nothing left that "only humans could do."

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The answer to the consciousness question also becomes extremely relevant when we consider the issue of interaction with syneoids. Again, one can dodge the bullet and insist that although we are dealing with beings displaying intelligence similar or superior to ours, nevertheless, they are qualitatively infinitely different; they don't have some sort of "essence" that we have and, as such, they should be excluded from our moral considerations. Needless to say, there would be something incredibly odd in the argument that a syneoid capable of performing human functions, or even exceeding them -intellectual, cognitive, and emotional- is not treated as a person.

Obviously, it goes without saying, that the notion of personhood is an incredibly complex one. Nevertheless, the most important components of the notion that are crucial for the purpose of the argument are: consciousness, agency, self-awareness, and possession of rights and duties (Taylor, 1985). It is also important to realize that personhood could be a matter of degree; it could be gained and lost incrementally, and that does not need to be limited to humans. The latter statement has remarkable consequence, namely, being a member of the human species is neither necessary, nor sufficient to be, at least to some degree, a person. It is currently commonly accepted that at least primates and some other mammals possess a degree of personhood far beyond a rudimentary one, that qualifies them to be included in the circle of our moral considerations. Since the Oxford Group (Finsen & Finsen, 2009) the movement of animal rights gained traction in the mainstream intellectual community. It drew its conclusions from Darwinism and consequentialism, and basically stated that there is no mysterious essential difference between humans and other animals biologically speaking. Thus, the same type of moral continuum should be applied to other beings. It would be expected that a syneoid, conscious by definition, would realize its own personhood. Thus, if someone or something is considered to be a person, we ought to have moral obligations towards it. The fact of its origin, the fact that it is neither a human person, nor a biological being is utterly irrelevant. It would be preposterous to deny such syneoids moral consideration based on the claim that it does not belong to human species. In fact, it could be argued that such attitude is similar to speciesism –a type of prejudice such as sexism or racism.

Moreover, it could be argued that, as along as a syneoid behaves as a person, it is actually irrelevant if it actually possesses consciousness the way we conceive it at all, in order for us to have moral obligations towards it. After all, many of us commit to the notion that we have moral considerations for certain animals even if, for what we know, they do not possess consciousness, or display it to a certain, typically to a lesser degree than ours. The consequences of committing to the notion that syneoids should be included in the area of our moral concern are staggering. It would be incomprehensible that we could treat such AI as merely our tools, an extension of our bodies, something that we can use, control, and manipulate. It would have to rule out any possibility of the master-slave relationship. We would have no choice but to concede that they are entitled to all rights that we reserve for persons. We would have to concede that they have freedom of choice according to their preferences and needs, etc. the notion of our "uniqueness" would evaporate in an instant.

Co-Existence, Competition, Convergence

Job, occupation, profession, career –they are not merely sources of economic sustenance; they are far more than a mere necessity or burden. In many cases, occupations define who we are. Work, whether paid or pursued for passion, whether fairly simple or highly sophisticated, provides fulfillment, sense of community, exploration, discovery, an invention. Professionalism, craftsmanship, refinement, and

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

artistry have been truly the driving forces of our civilization; for many –a fulfillment of an existential void, a deep sense of participation in a process of creation, mastery over nature, an instrument of control, and, ultimately, power.

Our inventions defined the work we engaged in for centuries from melding to typing; created new industries and employment opportunities around them. These inventions, intending to control the environment, had often unforeseen consequences that reshaped the way we lived, interacted with each other, and redefined our societal rules. To echo the famous maxim attributed to McLuhan that encapsulates this phenomenon: "We shape our tools and thereafter our tools shape us." Occasionally, an emergence of new technological advances such as, domestication of animals, creation of printing press, or invention of computers had a paradigm-changing effect leading to major disintegration in the status quo, shifts and disruptions in the existing power structures. We were adapting, not without struggle, failures, and suffering, to these changes. The unstoppable progress of AI, however, and the eventual arrival of syneoids will confront us with the challenge to radically renegotiate norms of coexistence with them without any guarantee of success.

In the context of a workplace, but not limited to it, possibly the best-case scenario would involve a collaborative co-existence. But it might likely turn into a combative competition over resources and power –something we are quite familiar with already. The Darwinian principle of the survival of the fittest could very well swing the odds to the advantage of syneoids. The end result of such competition could very well be the end of our kind, or perhaps, on the more optimistic note, as Buchannan argues: "humans will go the way bacteria did. They existed before the Cambrian explosion, and they still exist today. In fact, they continue to thrive. They just don't hold the dominant position they once did (Buchannan, 2015)." This perceived disadvantage of our species prompted some to entertaining the idea that we ought to use the ever-expanding information technology to transform ourselves into a new post-human species, as a part of adaptive response to environmental pressures. Preserving the natural state of a human body -claim the proponents of transhumanism- is an obstacle to progress in the self-directed evolution (Bostrom, 2005).

The arrival of syneoids will confront us on so many levels with the type of radical "other" - so different from us and yet a former "tool" "created in our image." It will launch a new stage of our journey -without precedence, unpredictable, unforeseeable, and volatile. We will be faces with many questions. Perhaps, to echo Pamela McCorduck (2004), "at that point, we turn to our own smart machines for advice on what the best next move is for the human race." And just like the well-being of most species currently existing is greatly affected by our decisions and action, so we might find ourselves in a similar position upon the arrival of syneoids.

REFERENCES

Bostrom, N. (2005). A history of transhumanist thought. *Journal of Evolution and Technology*.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford, UK: Oxford University Press.

Buchanan, M. (2015, August 13). *Get Ready for the Robot Explosion*. Retrieved November 13, 2015, from http://www.bloombergview.com/articles/2015-08-13/get-ready-for-the-robot-explosion

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Davis, N. (2015, November 10). *5 ways of understanding the Fourth Industrial Revolution*. Retrieved November 12, 2015, from https://agenda.weforum.org/2015/11/5-ways-of-understanding-the-fourth-industrial-revolution/

Dennett, D. (1991). *Consciousness Imagined. In Consciousness Explained* (pp. 431–455). Back Bay Books.

Finsen, S., & Finsen, L. (2009). Animal rights movement. In M. Bekoff (Ed.), *The Encyclopedia of Animal Rights and Animal Welfare*. Greenwood.

Ford, M. (2009). *The lights in the tunnel: Automation, accelerating technology and the economy of the future*. Acculant Publishing.

Frey, C., & Osborne, M. (2013, September 17). *The Future of Employment: How Susceptible are Jobs to Computerisation?* Retrieved from http://www.oxfordmartin.ox.ac.uk/downloads/academic/The_Future_of_Employment.pdf

Gera, D. (2003). *Ancient Greek ideas on speech, language, and civilization*. Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780199256167.001.0001

Howard, J. (2014, December). *The wonderful and terrifying implications of computers that can learn.* Lecture presented at TED.

Keynes, J. M. (1931). Economic possibilities for our grandchildren (1930). In Essays in persuasion, (pp. 358–73). Academic Press.

Krugman, P. (2013, June 13). *Sympathy for the Luddites*. Retrieved November 13, 2015, from http://www.nytimes.com/2013/06/14/opinion/krugman-sympathy-for-the-luddites.html

Mathuros, F. (2015, November 10). *What is the theme of Davos 2016?* Retrieved November 12, 2015, from https://agenda.weforum.org/2015/11/what-is-the-theme-of-davos-2016/

McCorduck, P. (1979). *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. San Francisco: W.H. Freeman.

McCorduck, P. (2004). *FAQ answered by Pamela McCorduck*. Retrieved November 7, 2015, from http://www.pamelamc.com/html/machines_who_think.html

Musk, Wozniak and Hawking urge ban on warfare AI and autonomous weapons. (n.d.). Retrieved November 5, 2015, from http://www.msn.com/en-gb/news/other/musk-wozniak-and-hawking-urge-ban-on-warfare-ai-and-autonomous-weapons/ar-AAdxy1e

Nilsson, N. (1984). Artificial Intelligence, Employment, and Income. *AI Magazine*, 5–14.

Popper, K. (1973). Of Clouds and Clocks. In Objective Knowledge (pp. 206–55). Oxford, UK: Oxford University Press.

Pratt, G. (2015). Is a Cambrian Explosion Coming for Robotics? *The Journal of Economic Perspectives*, *29*(3), 51–60. doi:10.1257/jep.29.3.51

Russell, S., & Norvig, P. (1995). *Artificial intelligence: A modern approach*. Englewood Cliffs, NJ: Prentice Hall.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Schwab, K. (2015, October 27). *Will the Fourth Industrial Revolution have a human heart?* Retrieved November 12, 2015, from https://agenda.weforum.org/2015/10/will-the-fourth-industrial-revolution-have-a-human-heart-and-soul/

Sullins, J. (2014). Information Technology and Moral Values. In *The Stanford Encyclopedia of Philosophy.* Retrieved from http://plato.stanford.edu/archives/spr2014/entries/it-moral-values/

Taylor, C. (1985). The concept of a person. *Philosophical Papers Human Agency and Language*, *1*, 97–114. doi:10.1017/CBO9781139173483.005

Wallach, W. (2009). *Moral machines: Teaching robots right from wrong*. Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780195374049.001.0001

This research was previously published in Strategic Imperatives and Core Competencies in the Era of Robotics and Artificial Intelligence edited by Roman Batko and Anna Szopa; pages 29-49; copyright year 2017 by Business Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 64
Realistic Simulation of Cultural Heritage

Chairi Kiourt
Hellenic Open University, Greece

Anestis Koutsoudis
University Campus at Kimmeria, Greece

George Pavlidis
University Campus at Kimmeria, Greece

Dimitris Kalles
Hellenic Open University, Greece

ABSTRACT

One of the most challenging problems in the simulation of real environments is to generate worlds that appear realistic and more attractive. It becomes increasingly challenging when the simulated environment focuses on minors (students), because the young generation has high demands on simulation systems due to their experience in computer gaming. Virtual museums are among the most important simulation environments, which present cultural and educational content for everyone. Their purpose is to enrich the users experience by allowing an intuitive interaction with the museum artifacts and to offer knowledge with the most pleasant ways. This paper focuses on the aspects of realistic simulations in the development of virtual 3D environments for Cultural Heritage applications. This study includes aspects regarding some of the most high-tech image effects, applicable artificial intelligence methods, powerful game engines, how real object can be reconstructed realistically and how all those features may be combined to produce realistic, pleasant, productive and educative environments.

INTRODUCTION

In realistic simulations of Cultural Heritage, virtual museums have a prominent role. Virtual museums are virtual environments that host virtual exhibitions created to tell the stories of the real museum artifacts to their visitors, just like in real museums, in a richer context and a wider range of possibilities. In the context of a real museum, this is achieved by combining exhibits and information in a carefully designed order and presentation style (Lepouras & Vassilakis, 2004). In the creation of virtual museums, contemporary IT technologies, such as multimedia, 3D computer graphics, spatial sound and virtual reality, can be used to enhance the presentation (the virtual presence), offering a more vivid and enjoyable

DOI: 10.4018/978-1-7998-0951-7.ch064

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

experience. Although each of the different technological innovations involved pushes towards their own respective, there are two main goals that are universal in such systems: *realism* and *speed*. In addition, the continuous development of Web services and computer infrastructures complemented by the increasing availability of computer game development platforms (also known as *game engines*), contribute towards a continuous release of serious games in diverse fields including entertainment, cultural heritage, education, artificial intelligence, sociology, military and health systems (Breuer & Bente, 2010). In a sense, serious games can be considered as an efficient approach for blending domain specific activities, like in cultural heritage and education, with gaming. By utilizing contemporary visualization and simulation technologies, serious games enhance the user's experience through photorealistic interactive environments (Van Eck, 2006). This form of stimulation is considered to be one of the primary factors for successful user engagement, in which playing, assumes the role of the driving force that promotes concentration in the activity process and user encouragement for further self-improvement. Stimulation is largely considered by the scientific community as a desired mechanism to achieve the desired results (Rogers, 1996).

Numerous works utilize various technologies to provide solutions for history teaching and learning, or to enhance actual museum visits. The notion of virtual museums and exhibitions has been introduced as an approach to overcome the limitations of the physical space and to provide a vivid experience to remote visitors (Tsichritzis & Gibbs, 1991). An overview of virtual museum technologies is presented in (Sylaiou, Liarokapis, Kotsakis, & Patias, 2009). The Augmented Representation of Cultural Objects (ARCO) (Patel, White, Walczak, & Sayd, 2003) (Wojciechowski, Walczak, White, & Cellary, 2004) was among the predominant efforts towards a dynamic virtual museum system accompanied by a 3D digitization technique to provide a framework to produce and exhibit 3D digital replicas of their artifacts. ARCO was also, evaluated as an Augmented Reality Interface (ARIF) (Sylaiou, Mania, Karoulis, & White, 2010) by focusing on the user experience in relation to the technologies of ARCO. In 2004, Lepouras & Vassilakis (2004) presented the concept of creating virtual museums focused primarily on educational content and related services by using a game engine. This virtual exhibition space took advantage of the high visual quality of modern game engines. Pavlidis et al. (Pavlidis, Tsiafakis, Provopoulos et al., 2006), proposed a Web-based 3D digital replicas management system with a dynamic virtual exhibition showroom. In addition, in (Pavlidis, Makarona, Arnaoutoglou et al., 2008) a more advanced framework for digital museums has been presented, where a non-photorealistic digital replica of a real museum is used to demonstrate educational activities rather its actual exhibition, aiming at increasing the museum's visitors. Sookhanaphibarn & Thawonmas (2009) presented a 3D virtual museum developed in the Second Life 3D world engine. The virtual museum was equipped with an innovative intelligent guidance system that was able to provide a customized navigation route based on the visitors' preferences (2009). In (Djaouti, Alvarez, Rampnoux, Charvillat, & Jessel, 2009) an interactive serious game is used for the promotion of a prehistoric heritage site. Anderson et al. (Anderson, McLoughlin, Liarokapis, Petridis, & Freitas, 2010) reviewed the state-of-the-art of theories, methods and technologies utilized by serious games as cultural heritage promotion tools by showing case studies that exploit such technologies. Papastamatiou et al. (Papastamatiou, Alexandridis, Tsergoulas et al., 2010), presented a dynamic Web-based 3D e-shop system for commercial use, offering a WYSIWYG graphical user interface. Furthermore, Sillaurren & Aguirrezabal (Sillaurren & Aguirrezabal, 2012), presented 3DPublish, a content management system that was developed in Unity3D game engine and allowed the generation of dynamic 3D exhibitions. Furthermore, works like (Bellotti, Berta, Gloria, D'ursi, & Fiore, 2012) focus on a generalization of the task-based learning theory in applications using smart mobile devices. In addition, Koutsoudis & Pavlidis (2011), proposed a novel approach for navigating within complex cultural scenes by exploiting content-based retrieval descriptors.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Koutsoudis et al. (2012), proposed a content-based navigation framework for a virtual museum, based on metadata that describe the exhibits and thus providing a semantic similarity-based navigation. In (Doulamis, Liarokapis, Petridis, & Mialous, 2012) a serious game is proposed based on scenarios derived from the cultural heritage domain and attempts to enrich the player's knowledge by spreading a mystery in ancient times. In a recent work (Bonis, Vosinakis, Andreou, & Panayiotopoulos, 2013) proposed a multi-user framework for virtual exhibitions that adapts to visitors' preferences. Recently, Kiourt et al. (2015a; 2015b; 2016) have studied many aspects of interactive virtual museums mainly focusing on interactive 3D object exhibitions with dynamic open data content; the virtual museum was re-introduced as a cultural and educational tool that focuses on user-driven exhibitions.

In all most all the aforementioned approaches, as in many other – a small list of virtual museums for educational purposes can be found at Sylaiou et al. (Sylaiou, Liarokapis, Kotsakis, & Patias, 2009; Sylaiou, Mania, Paliokas et al., 2016) – the main focus is to offer users pleasant and realistic virtual environments that try to simulate the real museums. The main contribution of this study is to introduce and present some of the most important factors that should be taken into account in building realistic simulations of Cultural Heritage that are accepted by all ages (children, adults and seniors). The study takes into account previous studies such as (Yu, Mortensen, Khanna, Spanlang, & Slater, 2012; Callieria, Debevecb, Pairc, & Scopig, 2006; Prem, Thalmann, Moccozet et al., 1998; Papagiannakis, L'Hoste, Foni, & Thalmann, 2001; DeLeon & Berry Jr, 2000; Addison, 2000). The study is presented as follows: (a) the most effective computer graphics techniques for building visual realism are being presented, (b) some of the most important artificial intelligent technologies applied in virtual museums are being reviewed, (c) game engines are being introduced and analyzed, (d) realistic reconstruction of the virtual environments is being presented, (e) the impact of the virtual environment/museum in education is being commented. The study concludes with some of the key remarks and highlights.

EFFECTS FOR ENHANCED REALISM

In order for the 3D virtual environment to be more realistic, some real-time image effects are being used, which significantly affect the quality and efficiency of the experience and do not impose any additional hardware requirements (Yang, Dong, Alelaiwi, & El Saddik, 2015). Often those image effects are named as *Screen Space effects*. It should be emphasized that by activating real-time image effects that are being recalculated for every pixel of every frame per second, the system exhibits high computational costs, which are only diminished by intelligent algorithms and the usage of sophisticated graphics hardware. A screen-space approach is one that involves 2D images of objects rather than the 3D geometry (Cohen & Wallace, 1993; Jonghyun, Jae-Hyun, Changwon & Byoungho, 2013). Generally, a 3D object or scene is represented in terms of geometric primitives such as vertices, polygons, edges, and surface patches. Graphics algorithms directly operate on geometric-space (also referred to as object-space) data to produce a desired result. On the other hand, image-space algorithms first rasterize the 3D geometry into one or more 2D images, and then operate on the pixels of these images. The rasterization step discretizes the 3D scene into pixels as visible from a given viewpoint. Each pixel in the image thus represents a finite region of the scene. Therefore, all the information pertaining to that region required to perform the rendering can be stored at the corresponding pixel location in the image, such as 3D positions, surface normal vectors, etc. In the following paragraphs the main screen space effects that contribute to the realism of a simulation are being presented.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lighting, Reflection and Ambient Occlusion

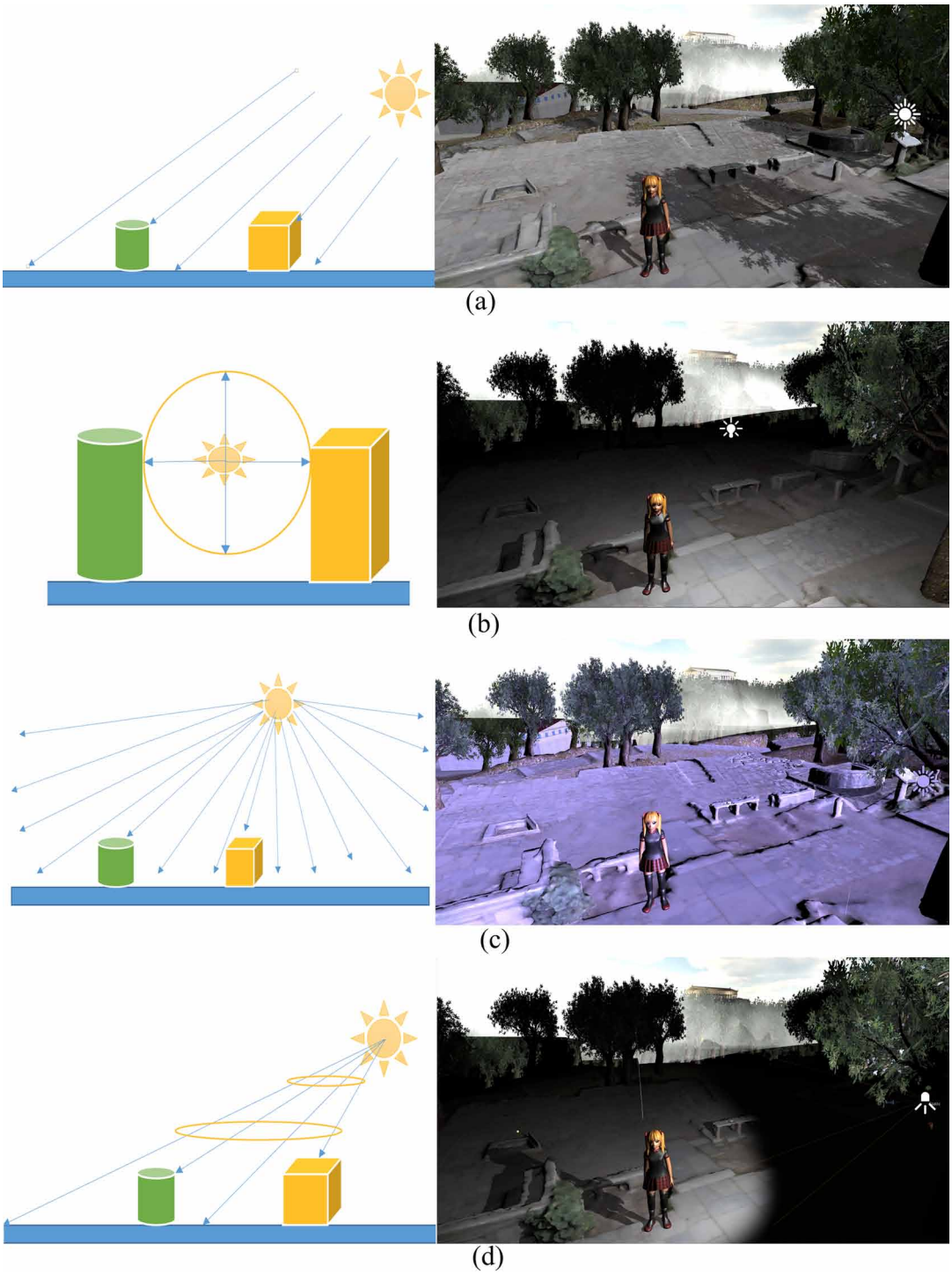
Lighting is a method for calculating the shading of a 3D object, by using the intensity, direction and color of light that falls on it (Cohen & Wallace, 1993; Verbeck & Greenberg, 1984; Iones, Krupkin, Sbert, & Zhukov, 2003). Generally, nowadays, in almost all 3D environment developments tools, such as game engines, there are two types of lighting techniques: 1) *real-time lighting*, in which lighting is updated in every frame being displayed, 2) *baked lighting*, in which the lighting is pre-calculated and transform to a static lighting effect by being stored on the texture map of the objects, producing what is called a *light map*.

In order to cast light in a virtual 3D scene there are five types of virtual light sources that simulate real-world light sources; these sources are directional, point, area, spot and ambient light sources. A *directional light source*, as the name implies, illuminates an entire 3D scene from one specific direction, Figure 1a. It is generally used to represent lighting from a source at large distances from the scene. For example, sunlight is usually represented as directional light. A *point light source* (Cohen & Wallace, 1993) is represented with a single 3D point in the scene, generally placed in close proximity to the other objects of the scene, unlike the directional light source (Figure 1b). The rendering equation also simplifies under point light illumination since only a single illumination direction is used at each point, x. However, the direction changes with the position of the point. Perhaps the type of light source that mostly resembles the ones found in real life is the *area light source* (Cohen & Wallace, 1993). An area light source is defined as a fully-fledged 3D object, just like any other in the scene. It emits light over its entire surface area, therefore illuminating a particular point in the scene from multiple directions, unlike the point and directional light sources, in which a point is lit only from one direction (Figure 1c). The rendering equation can be adjusted to accommodate area lights by changing the integral over solid angle to integral over the area of the light source. Like a point light, a *spot light source* (Cohen & Wallace, 1993) has a specified location and range over which the light falls off. However, the spot light is constrained to an angle, resulting in a cone-shaped region of illumination (Figure 1d). The center of the cone points in one direction of the light object. Light also diminishes at the edges of the spot light's cone. Widening the angle increases the width of the cone and with it increases the size of this fade, known as the 'penumbra'. Spot light sources are generally used for artificial light sources such as flashlights, car headlights and searchlights. With the direction controlled from a script or animation, a moving spot light will illuminate just a small area of the scene and create dramatic lighting effects. An *ambient light source* is present all around the scene and doesn't come from any specific source object. It can be an important contributor to the overall look and brightness of a scene. An ambient light can be useful in a number of cases, depending upon your chosen art style. An example would be bright, cartoon-style rendering where dark shadows may be undesirable or where lighting is perhaps hand-painted into textures. Ambient light can also be useful if one needs to increase the overall brightness of a scene without adjusting individual lights.

In order to render more realistic images in computer graphics a model of how may objects reflect light is required (Cook & Torrance, 1982; Phong, 1973). The reflectance model must describe both the color and the spatial distribution of the reflected light. The model is independent of the other aspects of image synthesis, such as the surface geometry representation. Most real surfaces are neither ideal specular reflectors nor ideal diffuse reflectors (Cook & Torrance, 1982; Phong, 1973).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Lighting types: (a) directional light, (b) point light, (c) area light, (d) spot light



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In any case, the general form of the equation that describes the basic lighting (Phong) model for objects in a 3D scene, in which multiple light sources might exist, is (boldface letters denote vectors)

$$I = K_e + K_a \times I_a + K_d \times I_d \times \max\left(\boldsymbol{n} \cdot \boldsymbol{v}, 0\right) + K_s \times I_s \times \delta\left(\boldsymbol{n} \cdot \boldsymbol{v}\right) \times \max\left(\boldsymbol{n} \cdot \boldsymbol{v}, 0\right)^s$$

where K_e the *emissive* part of light that comes from the object itself (might be zero); $K_a \times I_a$ the *ambient* part that corresponds to the effect of any ambient light source with I_a the global ambient light intensity (or color);

$$K_d \times I_d \times \max\left(\boldsymbol{n} \cdot \boldsymbol{v}, 0\right) = K_d \times I_d \times \cos\theta$$

the *diffuse* part coming from a directed light, with K_d the object's diffuse reflection coefficient, I_d the light intensity and θ the angle between the light vector $\vec{v}$ and the surface normal $\vec{n}$ at the point of computations;

$$K_s \times I_s \times \delta\left(\boldsymbol{n} \cdot \boldsymbol{v}\right) \times \max\left(\boldsymbol{n} \cdot \boldsymbol{v}, 0\right)^s = K_s \times I_s \times \cos^s \phi$$

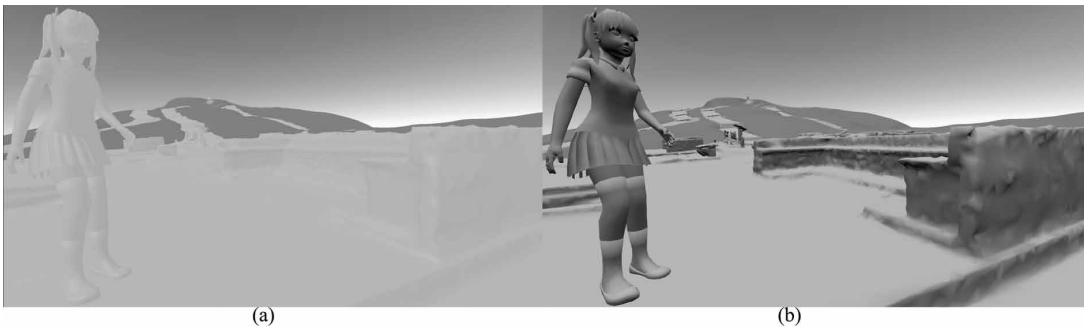
the *specular* part coming from a directed light, K_s being the object's absorption coefficient, I_s the incoming light intensity, ϕ the angle that represents the deviation of view angle from mirror direction and the power *s* the shininess coefficient. The term δ(***n*•*v***) results either 1 if ***n*•*v*** is greater than zero, or 0 otherwise.

Although lighting techniques provide support for all the basic calculations for the rendering of a visible surface element with respect to lights that might hit upon it and the surface characteristics, still, a realistic scene representation cannot be attained. This is due to an issue that relates with how exposed are those surface elements to the ambient lighting of the scene. That is, points that are occluded due to the geometry of their surroundings should receive less ambient light, thus should appear darker in the final rendering. The solution to this issue has been encoded in computer graphics as *Ambient occlusion* (Yu, Wang, & Yu, 2010; Miller, 1994; Dutre, Bala, & Bekaert, 2006; Bavoil & Sainz, 2009; Vardis, Papaioannou, & Gaitatzes, 2013). Ambient occlusion is a sophisticated ray tracing calculation, which simulates soft global illumination by faking darkness perceived in corners and at mesh intersections, creases, and cracks, where light is diffused (usually) by accumulated dirt and dust. Simply put, this technique darkens creases, holes and surfaces that are close to each other, as in real life such areas tend to block out or occlude ambient light, and hence they appear darker. A simple comparative example of using this technique in the presented system is shown in Figure 2.

There are several variations of real-time ambient occlusion, some of the most important are: Screen Space Ambient Occlusion (SSAO), Screen Space Directional Occlusion (SSDO), High Definition Ambient Occlusion (HDAO), Horizon Based Ambient Occlusion+ (HBAO+), Alchemy Ambient Occlusion (AAO), Angle Based Ambient Occlusion (ABAO) and Voxel Accelerated Ambient Occlusion (VXAO). The techniques that are based on screen-space computations target fast implementations by computing on the 2D projected visible scene pixels using pixel depth rather than the complete scene geometry. In general, the occlusion $AO_{\boldsymbol{p},\hat{n}}$ (or the surface irradiance $E_{\boldsymbol{p},\hat{n}}$) at a point $\boldsymbol{p}$ (in vector form) on a surface with a normal $\hat{n}$ at that point can be computed by integrating the visibility function $V_{\boldsymbol{p},\hat{\omega}}$ (or the incidence

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. A scene (a) with no ambient occlusion effect; (b) with ambient occlusion exaggerated



irradiance *L*) over the hemisphere Ω above and centered at the point with respect to the projected solid angle ω. The visibility function $V_{\boldsymbol{p},\hat{\omega}}$ equals 1 if the point $\boldsymbol{p}$ is occluded in the direction of the angle ω, or 0 if it is not occluded.

$$AO_{\boldsymbol{p},\hat{n}} = \frac{1}{\pi}\int_{\Omega} V_{\boldsymbol{p},\hat{\omega}} \cdot \left(\hat{n}\cdot\hat{\omega}\right) d\omega \quad E_{\boldsymbol{p},\hat{n}} = \int_{\Omega} L\cos\theta d\omega$$

Practically, SSAO for example, is implemented as a *pixel shader*, which analyses the scene depth buffer that is stored in a texture and for every pixel on the screen calculates the occlusion depending on the depth difference between the current and neighboring pixels.

Shadows play an important role in conveying reality in computer-synthesized images (Amanatides, 1987; Woo, Poulin, & Fournier, 1990; Hasenfratz, Lapierre, Holzschuch, & Sillion, 2003; Guennebaud, Barthe, & Paulin, 2006). A shadow may be described as a darkness region in an otherwise illuminated region (Woo, Poulin, & Fournier, 1990). Shadows also facilitate the comprehension of spatial relationships between objects. Though there are situations where shadows are not required (e.g. if the light source is behind the viewer or when modeling an overcast day), there are many other situations in which they are essential. The complexity of the shadows is related with the light mapping in trying to simulate the real environment shadows that are associated with the light sources (Amanatides, 1987). The presence of shadows makes 3D computer graphics look more natural. Without them, scenes often feel unnatural and flat, and the relative depths of objects in the scene can be very unclear. They emphasize and clarify the three dimensional settings of the objects being displayed. However, the trouble with rendering high quality shadows is that they are computationally intensive and hard to generate. In computer graphics shadows appears in two types, first "hard shadows" are crisply defined with sharp edges and second "soft shadows" are less distinct and fade off toward the edges. An example of some hard shadows is depicted in Figure 3.

Depth of Field

Depth of field (DoF) is the distance between the nearest and farthest objects in a scene that appear acceptably sharp in an image or 3D scene (Yu, Wang, & Yu, 2010; Demers, 2004; Selgrad, Reintges, Penk, Wagner, & Stamminger, 2015). In computer graphics the simulation of DoF is a common post

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. A scene (a) with no shadow effect; (b) with shadow effect active



(a) (b)

processing effect that simulates one of the most notable properties of a camera lens, the limited depth of focus. In real life, human eyes and a camera can only focus sharply on an object at a specific distance; objects nearer or farther from the camera will be somewhat out of focus. The blurring not only gives a visual cue about an objects distance but also introduces *Bokeh* (a Japanese word meaning *blur*), which is the term for pleasing visual artifacts that appear around bright areas of the image as they fall out of focus. DoF effects have been studied for several years. Starting around three decades ago (Potmesil & Indranil, 1981; 1982), there have been numerous approaches to adding DoF effects to computer graphics, but official published work in the area of DoF for computer graphics has been provided by Barsky et al (2003). The DoF simulation methods can be categorized in several ways in computer graphics (Barsky & Kosloff, 2008), but the two most important categories are: the *object-space* methods and the *image-space* methods. *Object space* methods operate on 3D scene representation, and build DoF effects directly into the rendering pipeline. *Image space* methods operate on images that were rendered with everything in perfect focus. Generally, *object space* methods generate more realistic results but image space methods are much faster (Barsky & Kosloff, 2008). Nowadays almost all game engines provide the *image space* methods. A sample of image space method applied in a 3D environment is presented in Figure 4. The left image depicts the application of the DoF effect for far away objects, whereas the right image depicts the application of the DoF effect for medium distance objects.

The mathematics of the blur due to a limited DoF in real-life photography are connected with the aperture of the lens and the *circle of confusion* (usually denoted as *CoC*), and are derived from trigonometry, usually by setting the far limit of the DoF to infinity, thus

$$CoC = \frac{f_c^2}{N \cdot \left(D_h - f_c\right)} \approx \frac{f_c^2}{N \cdot D_h}, \quad N = \frac{f_c}{d} \text{ the aperture}, \quad D_h : \text{the hyperfocal distance}$$

where d is the aperture diameter, f_c the focal length, N the f-number of the lens.

The *hyperfocal distance* D_h is the distance at which the lens focus extends from half that distance to infinity. In many practical cases, the estimation of the circle of confusion for a lens can be done by setting the infinity mark on the distance scale opposite to the largest marked f-number on the DoF scale, in which situation the focus index is the hyperfocal distance. Since lens distance scales usually show object to camera distances, then if O_h is the indicated hyperfocal distance the formula should be corrected as

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. DoF examples: (a) shallow DoF, near-to-middle distances in focus; (b) extremely shallow DoF, middle distances in focus



(a) (b)

$$CoC \approx \frac{f_c^2}{N \cdot \left(O_h - 2f_c\right)}$$

These computations refer to how the DoF effect appears in real life lenses. In computer graphics, the virtual *pinhole camera* projects the 3D world through a zero-size lens in order to create the 2D view on a flat screen. Thus, the camera has infinite DoF and the real-life DoF effect has to be simulated. In any case the circle of confusion still has to be considered. One way to do so is by estimating *CoC* from the camera to object distance and the camera parameters,

$$CoC = \left| d \frac{f_c}{D_f - f_c} \frac{D_f - D_o}{D_o} \right|$$

in which d is the diameter of the lens, f_c is the focal length of the lens, D_o is the object distance and D_f the plane in focus (the distance ideally being focused on the image plane, following the fundamental formula $\frac{1}{D_f} + \frac{1}{I_f} = \frac{1}{f_c}$, I_f being the image plane to lens internal distance). The object distance can be calculated from the z-buffer x values, using the extreme far and near z-buffer values z_f, z_n,

$$D_o = -\frac{z_f z_n}{z \cdot \left(z_f - z_n\right) - z_f}$$

Alternatively, the circle of confusion may be estimated directly from the z-buffer,

$$CoC = \left| z \cdot CoC_s + CoC_b \right|, \begin{cases} CoC_s = \dfrac{d \cdot f_c \cdot D_f \cdot \left(z_f - z_n\right)}{\left(D_f - f_c\right) \cdot z_f \cdot z_n} a\, scaling term \\ CoC_b = \dfrac{d \cdot f_c \cdot D_f \cdot \left(z_f - z_n\right)}{D_f \cdot f_c \cdot z_n} a\, bias term \end{cases}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In order to simulate the DoF effect a method has to approximate the blur that would happen as if the lens were real. Practical implementations of the DoF effect can be categorized in the following approaches:

- Forward-mapped z-buffer techniques (Potmesil & Chakravarty, 1981)
- Distributing traced rays across the surface of a (non-pinhole) lens (Cook, Porter, & Carpenter, 1984)
- Rendering from multiple cameras (also accumulation-buffer technique) (Haeberli & Akeley, 1990)
- Rendering multiple layers (Scofield, 1994)
- Reverse-mapped z-buffer techniques (Arce & Wloka, 2002; Demers, 2003)

Antialiasing

In signal processing *antialiasing* refers to various techniques to tackle the problems of aliasing in a any sampled signal. Aliasing is an effect that causes different signals to appear the same when sampled. When aliasing occurs in signals sampled in time is referred to as *temporal aliasing*, whereas when it occurs in spatially sampled signals is called *spatial aliasing* and takes the form of *moiré patterns*. In spatial aliasing, the effect of aliasing is most apparent in situations in which the Nyquist sampling limit (sampling theorem) breaks during a sampling process (discrete sampling below the Nyquist frequency). To suppress aliasing, one may either increase the sampling frequency or suppress the high frequencies in the original signal using filtering; antialiasing methods fall in the second category of measures against aliasing. Antialiasing methods try to restrict the bandwidth of the signal so that the sampling (completely in theory or approximately for all practical purposes) satisfies the sampling theorem, over whichever band of interest is preselected in the specific application. In a simple application of computing and plotting a single white point on a black background the computations include

$$I_{\left(\hat{x}_i,\hat{y}_j\right)} = I_{(x,y)} + \left(1-\left|x-\hat{x}_i\right|\right)\cdot\left(1-\left|y-\hat{y}_j\right|\right)$$
$$\text{for}\begin{cases} i = \lfloor x \rfloor \ \text{to} \lceil x \rceil \\ j = \lfloor y \rfloor \ to \lceil y \rceil \end{cases}$$

where $I_{\left(\hat{x}_i,\hat{y}_j\right)}$ is the computed intensity of the point at the target location $\left(\hat{x}_i,\hat{y}_j\right)$, *i,j* varying between the floor and ceiling of the *x,y* coordinates accordingly, given the initial location (*x,y*) and intensity $I_{(x,y)}$

In modern computer graphics applications, antialiasing is usually referenced as an image effect. The antialiasing image effect is calculated in real-time, which gives smoother appearance of the graphics based on the difference of colored areas of the image (William, 1980; Franklin, 1977; Bineta, 2013; Bineta, 2013; Korein & Badler, 1983). It can be defined as a technique to minimize the distortion artifacts known as aliasing when showing a high-resolution image in a lower resolution display. Practically, antialiasing means removing signal components that have a higher frequency than those that can be properly resolved by the recording (or display) device (Franklin, 1977). An example of antialiasing is shown in Figure 5.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Antialised view, with some antialised shadows



Motion Blur

Motion blur is the apparent streaking of rapidly moving objects in a still image or a sequence of images such as a movie or animation. During recording of a single image, motion blur appears as a capture of the motion of moving objects, which may occur either due to rapid movement or long exposure. Obviously this effect is the result of limitations of the recording devices to freeze-capture fast moving objects, or even slow moving objects but under long exposure requirements. Motion blur can take any of three forms, including (a) *linear blur*, which travels along a single direction to a specific extend, denoted by a strength and an angle, (b) *radial blur*, which is responsible for circular blurring effects around a center and can be defined by a radius and a strength, and (c) *zoom blur*, which represents a zooming in or out motion centered at a point on the image place defined by a center and a speed (the strength in this case).

In computer graphics there is no such problem, so in order to enhance visual realism, this effect has to be simulated. Let us consider, for example, the simple 2D case that is easy to follow, and suppose there is an image $f(x,y)$ undergoing the motion. In the duration T of the exposure, the total exposure at any point will be the integration of instantaneous exposures at that point, leading to a blurry image $g(x,y)$,

$$g(x,y) = \int_0^T f\left(x - x_0(t), y - y_0(t)\right) dt$$

where $x_0(t)$, $y_0(t)$ are the time-varying motion parameters. It is easy to see that its Fourier transform is

$$\begin{aligned}
\mathcal{G}(u,v) &= \iint_{-\infty}^{\infty} g(x,y) e^{-2\pi i(ux+vy)} dxdy = \iint_{-\infty}^{\infty} \left[\int_0^T f\left(x - x_0(t), y - y_0(t)\right) dt \right] e^{-2\pi i(ux+vy)} dxdy \\
&= \int_0^T \left[\iint_{-\infty}^{\infty} f\left(x - x_0(t), y - y_0(t)\right) e^{-2\pi i(ux+vy)} dxdy \right] dt = \int_0^T \mathcal{F}(u,v) e^{-2\pi i\left(ux_0(t)+vy_0(t)\right)} dt \\
&= \mathcal{F}(u,v) \int_0^T e^{-2\pi i\left(ux_0(t)+vy_0(t)\right)} dt \Rightarrow \mathcal{G}(u,v) = \mathcal{H}(u,v) \cdot \mathcal{F}(u,v)
\end{aligned}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

If the motion parameters $x_0(t)$, $y_0(t)$ are known then $\mathcal{H}\left(u,v\right)$ is easily defined and motion readily computed. Consider the simple case in which there is a linear motion blur, in which,

$$x_0\left(t\right) = a\frac{t}{T} \qquad y_0\left(t\right) = b\frac{t}{T}$$

Then the transfer function $\mathcal{H}\left(u,v\right)$ can be defined,

$$\begin{aligned}\mathcal{H}\left(u,v\right) &= \int_0^T e^{-2\pi i\left(ua\frac{t}{T}+vb\frac{t}{T}\right)}dt = \int_0^T e^{-2\pi i\frac{ua+vb}{T}t}dt \\ &= -\frac{T}{2\pi i\left(ua+vb\right)}e^{-2\pi i\frac{ua+vb}{T}t}\Bigg|_0^T \\ &\quad -\frac{T}{2\pi i\left(ua+vb\right)}(e^{-2\pi i\left(ua+vb\right)}-1) \\ &\Rightarrow \mathcal{H}\left(u,v\right) = \frac{T}{2\pi i\left(ua+vb\right)}\left(1-e^{-2\pi i\left(ua+vb\right)}\right)\end{aligned}$$

Since the transfer function is defined then $\mathcal{G}\left(u,v\right) = \mathcal{H}\left(u,v\right)\cdot\mathcal{F}\left(u,v\right)$ results the Fourier representation of the motion blurred image, which in turn gives the blurred image by taking the inverse Fourier transform.

Things get rather complicated in the 3D domain. The spatial dimensions become three, and the temporal dimension is also incorporated in more complex ways. In addition, no matter what happens in a 3D virtual world, it is the screen space 2D projection that should comply with the observers' experience. A typical mathematical formulation of motion blur in 3D graphics resolves for the 2D projected screen space $I(x,y)$,

$$I\left(x,y\right) = \sum_i \int_\Omega \int_{\Delta T} r\left(\omega,t\right) g_i\left(\omega,t\right) L_i\left(\omega,t\right)\ dt\ d\omega$$

for object i, a solid angle Ω and the virtual aperture time ΔT, in which $r(\omega,t)$ is the reconstruction filter that accounts for the effects of shutter shape and efficiency, lens aberration and film influence, $g_i(\omega,t)$ is a function of the occlusion of object i and $L_i(\omega,t)$ the radiance of the object.

The numerous rendering methods that have been proposed can be classified into (a) *analytic methods* that use a closed form solution (Korein & Badler, 1983; Grant, 1985), (b) *geometric substitution* that substitute the original geometry with alternative geometric primitives for which temporal changes are followed (Catmull, 1984; Glassner, 1988; Guan & Mueller, 2004; Jones & Keyser, 2005; Schmid, Sumner, Bowles, & Gross, 2010) (c) *texture clamping*, which originally targets the elimination of aliasing effects (Loviscach, 2005), (d) *Monte Carlo methods* that include various stochastic approaches (Haeberli & Akeley, 1990; Dayal, Watson, & Luebke, 2002; Dayal, Woolley, Watson, & Luebke, 2005; Walter, Arbree, Bala, & Greenberg, 2006; Nehab, Sander, Lawrence, Tatarchuk, & Isidoro, 2007; Kim & Ko, 2007; Akenine-Moeller, Munkberg, & Hasselgren, 2007; Hachisuka, Jarosz, Weistroffer, Dale, Hum-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

phreys, Zwicker, & Jensen, 2008; Egan, Tseng, Holzschuch, Durand, & Ramamoorthi, 2009; Overbeck, Donner, & Ramamoorthi, 2009; Ragan-Kelley, Lehtinen, Chen, Doggett, & Durand, 2010), (e) *post-processing methods* that are screen space methods that use selected pre-renders of a scene blurred by using motion information from available scene data, decoupling motion blurring from rendering (Max & Lerner, 1985; Chen & Williams, 1993; Zheng, Koestler, Thuerey, & Ruede, 2006; Vlachos, 2008; Sousa, 2008), (f) *hybrid techniques* (Sung, Pearce, & Wang, 2002) and (g) *physics-based approaches* that exploit mechanics and optics to device models for motion blur simulation (Lin & Chang, 2006; Telleen, Sullivan, Yee, Wang, Gunawardane, Collins, & Davis, 2007; Pachur, Laue, & Roefer, 2009).

Typical examples of motion blur are shown in Figure 6. Motion blur (Korein & Badler, 1983; Dachille & Kaufman, 2000; Nathan & Keyser, 2005; McGuire, Hennessy, Bukowski, & Osman, 2012) is an important method for increasing the visual quality of real-time applications such as games or training simulators.

Particle Systems

A *particle system* (Reeves, 1983) is a system of a large number of elementary graphics particles having a specific appearance, sharing a predetermined formation pattern and following specific rules, which can be used to represent fuzzy objects of processes otherwise difficult or impossible to represent by a single model. Such objects and processes include highly chaotic systems, natural phenomena or processes caused by chemical reactions, such as fire (Nguyen, Fedkiw, & Jensen, 2002), smoke (Csuri, Hackathorn, Parent, Carlson, & Howard, 1979; Huang, Gong, & Liang, 2015), explosions (Yngve, Obrien, & Hodgins, 2000), water (Foster & Fedkiw, 2001), ocean wave (Hinsinger & Neyret, 2002), clouds (Gardner, 1985), fogs, snow (Tan & Fan, 2011), dust (Jim, Xiadong, & Wegman, 1999), electricity, stars and galaxies. In essence, particle systems are used to realistically represent fluids and gasses, or in general, objects and processes with highly dynamic behavior. Instead of using geometric primitives, in these cases particle clouds are being used to render dynamically changing volumes. Stochastic processes are used to control those objects' shapes and appearance, adopting the paradigm of stochastic procedural modeling.

Particle systems are also used for modeling more tangible objects such as complicated trees (Reeves, 1985), cloths and fabric (Breen, 1994; Eberhardt, Weber, & Strasser, 1996). Realistic particle movement is often achieved by simulating real-world physics (Reynolds, 1999). The diversity of particle system applications demonstrates their importance in modern interactive media and games especially in serious games environments (Hastings, Guha, & Stanley, 2007). In recent years the researchers focus on the creation of particles APIs, in order to increase the usability and potentials of the particle systems (McAllister, 2000). In general, a particle system and its particles have very similar parameters, but with different values (Reeves, 1983): position (orientation in 3D space and center location x, y, and z), movement (velocity, rotation, etc.), color (RGB), transparency (alpha), shape (point, line, sphere, cube, rectangle, etc.), volume, density, mass, lifetime (start and end time) and blur head and rear pointers (only for particles). Figure 7 depicts a view of a 3D environment where several particle systems have been included, such as rain, snow, wind and fire.

The position, shape, and size of a particle system determine the initial positions of the particles and their range of movement. The movements of the particles are restricted within the range defined by their associated particle system. The shape of a particle system can be a point, line segment, sphere, box, or cylinder. Internal or external forces and the effects of the rotations and accelerations of the particles affect the movement of a particle system, as a whole. A particle system may change its shape, size, color,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Motion blur (a) of an animated avatar and (b) of a moving scene



transparency, or some other attributes as it evolves. The lifetime defines how many time slices (frames) a particle will be active. A particle has both a head position and a tail position. The head position is usually animated and the tail position follows along for motion blur. The simulation looks more dynamic and has more particles with motion blur, at the cost of longer rendering time.

Generally, a particle systems starts by spawning particles using *emitters*, which are able to generate large numbers of elementary particles. The laws governing the particle dynamics start by defining a force field and the integration of the laws of mechanics. In a simple case, each particle is independent and considered separately (individually) but in the most general case there might be forces among the particles. In a simple example of a *sprinkler*, the mathematics can be defined easily using elementary laws of mechanics (boldface letters representing vectors),

$$\boldsymbol{x}_{i+1} = \boldsymbol{x}_i + d\boldsymbol{x} = \boldsymbol{x}_i + \boldsymbol{v}_i \cdot dt$$
$$\boldsymbol{v}_{i+1} = \boldsymbol{v}_i - g \cdot dt$$
$$\boldsymbol{x}\left(t_0\right) = \boldsymbol{x}_0 + \boldsymbol{x}_r$$
$$\boldsymbol{v}\left(t_0\right) = \boldsymbol{v}_0 + \boldsymbol{v}_r$$

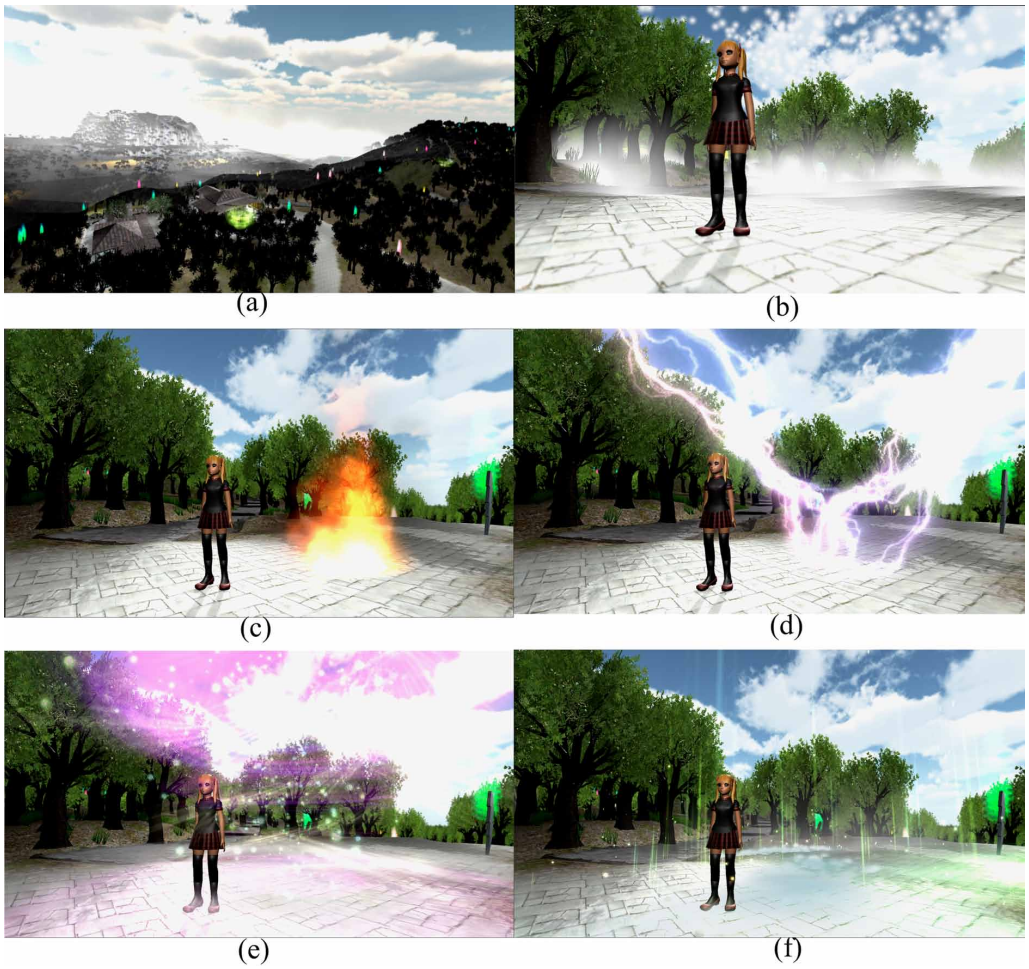
where there is a gravitational field g in which particles are spawned at t_0 at an initial location $\boldsymbol{x}_0$ with an initial velocity $\boldsymbol{v}_0$, both of which can be randomized by drawing random samples $\boldsymbol{x}_r$ and $\boldsymbol{v}_r$ from predefined distributions (usually seeded by uniform distributions spanning [-1,1]). In addition, the generation of particles can be controlled by a spawning rate r (in particles per second), which produces $n=(t - t_{last})\bullet r$ particles, updating the t_{last} if $n>0$. Generally, a particle is described by its state vector $\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{v} \end{pmatrix}$, which is composed of six (6) numbers for a point mass particle in 3D. The dynamics of the particle can be described by the derivative of the state vector,

$$\frac{d\boldsymbol{X}}{dt} = f\left(\boldsymbol{X}, t\right) = \begin{pmatrix} \boldsymbol{v} \\ \dfrac{\boldsymbol{F}\left(x, v\right)}{m} \end{pmatrix}, \qquad with\ \boldsymbol{F} = m \cdot \boldsymbol{a} = m\frac{d\boldsymbol{v}}{dt}$$

which generalizes for N point masses,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Particle systems, fog, snowfall, fire, lightning strike, cyclone and magic lights



$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{v}_1 \\ \vdots \\ \boldsymbol{x}_N \\ \boldsymbol{v}_N \end{pmatrix}_{6N\times 1} , \quad \frac{d\boldsymbol{X}}{dt} = f\left(\boldsymbol{X}, t\right) = \begin{pmatrix} \boldsymbol{v}_1 \\ \boldsymbol{F}_1 \\ \vdots \\ \boldsymbol{v}_N \\ \boldsymbol{F}_N \end{pmatrix}_{6N\times 1}$$

in which the indices represent the corresponding particle. For a single point mass m the force $\boldsymbol{F}$ is easily defined for the 3D space,

$$\boldsymbol{F} = \begin{pmatrix} 0 \\ 0 \\ -mg \end{pmatrix}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

which can also be used in reverse direction (as buoyancy) in simulating smoke, etc. In case particles are not independent then gravity should depend on all mass interactions, taken pairwise for particles *i* and *j* at a distance or r_{ij}^2 ,

$$\left\| F_{ij} \right\| = \frac{G m_i m_j}{r_{ij}^2}, \quad G = 6.67 \times 10^{-11} Nm^2 \,/\, kg^2$$

in which case computations get complicated as they are $O(n^2)$.

Nevertheless, these are relatively simple cases as realistic simulations usually require the introduction of other forces, such as viscous damping or other spatial interactions such as attractive and repulsive forces, or even other dynamics such as crowd behaviors, collisions and collision avoidance.

ARTIFICIAL INTELLIGENCE IN VIRTUAL ENVIRONMENTS

One important ingredient in realistic simulations are the intelligent agents, which are the subject of major controversy over their origin, formal background, definitions, methods, applications and future directions (Ritchings, Anastassakis, & Panayiotopoulos, 2001). For social simulations to be meaningful, it is necessary to implement realistic models for both the virtual agents (CaI, Liu, Yu, & Zhang, 2015), the environment (Pelechano, Allbeck, & Badler, 2007) and their interactions; and the latter has long been the subject of multi-agent systems and reinforcment learning research (Gosavi, 2003; Varshavskaya, Kaelbling, & Rus, 2008; Silva, Melo & Veloso, 2016). A lot of attention has been given to the definition of accurate models for agents (e.g., behavioral, decision-making and interactions models) (Al-Zinati & Wenkstern, 2015). The approach of building intelligent virtual environments is relatively new in virtual environments (Pelechano, Allbeck, & Badler, 2007), in which the virtual environment is artificially being developed and created, sometimes in real time, based on the progress and the profile of the users (Al-Zinati & Wenkstern, 2015).

Figure 8. AI agents in virtual worlds: (a) navigation area for an AI agent; (b) AI agent-based virtual guided tour



(a) (b)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In addition, the realistic simulation of virtual crowds has diverse applications in architecture design, emergency evacuation, urban planning, personnel training, education, virtual museums and entertainment. The are several approaches that aim to create better representations of crowds in virtual environments, with the *agent-based methods* being the most successful; agent-based methods focus more on individual behavior, whereas crowd simulations aim to exhibit emergent phenomena of groups (Sud, Andersen, Curtis, Lin, & Manocha, 2008; Treuille, Cooper, & Popović, 2006). Not surprisingly, combinations of these approaches (which, however, do exhibit a bias in favor of one of the approaches) can deliver results of substantially enhanced realism (Shao & Terzopoulos, 2005) and, often, the crowd that is being simulated by agent-based models is designed as a multi-agent system, in order to develop different behaviors that may cooperate or compete each other (Bennewitz & Burgard, 2001; Li & Chou, 2003; Musse & Thalmann, 1997; Hanna & Richards, 2014). This approach increases the realism of the simulated crowds by highlighting the social attitude of artificial intelligence agents.

An important technical challenge in AI simulations is the path planning (Sud, Andersen, Curtis, Lin, & Manocha, 2008; Stüvel, de Goeij, van der Stappen, & Egges, 2015; Naderi, Rajamäki, & Hämäläinen, 2015; Demesure, Defoort, Bekrar, Trentesaux, & Djemaï, 2016; Champagne & Tang, 2005; Anderson, McLoughlin, Liarokapis, Petridis, & Freitas, 2010). There are several different algorithms that improve the navigation of agents in virtual environments, which often result in more realistic and detailed simulations (Sud, Andersen, Curtis, Lin, & Manocha, 2008; Naderi, Rajamäki, & Hämäläinen, 2015; Pettre, Laumond, & Thalmann, 2005). An example is shown in Figure 8a, where the blue ground is the 'walkable' area for the agent, which has to plan its route taking into account any static objects acting as obstacles. The path planning and motion planning problems have long been central to robot-based systems and AI systems in general (Luna, Lahijanian, Moll & Kavraki, 2014; Lahijanian, Maly, Fried, Kavraki, Kress-Gazit & Vardi, 2016).

Another important contribution of AI in virtual environments with cultural and educational topics (like the virtual museums) that enhances the user experience is the use of virtual agents for guided tours (Bogdanovych, Rodriguez, Simoff, & Cohen, 2009; Bickmore, Pfeifer, & Schulman, 2011; Ichiro, 2008; Sookhanaphibarn & Thawonmas, 2009; Vassos, et al., 2016). A virtual agent is a computer generated, animated, AI virtual character (usually with anthropomorphic appearance) that constitutes a real-time service representative. It leads an intelligent conversation with users, responds to their questions and performs adequate nonverbal behavior. Further enhancements may include virtual agents-guides that discuss with users and present the space (Oberlander, Karakatsiotis, Isard, & Androutsopoulos, 2008), and adaptive, personalized exhibition spaces (Bonis, Stamos, Vosinakis, Andreou, & Panayiotopoulos, 2009). An example of this concept is shown in Figure 8b, where an AI (chat-bot) avatar asks a user (also represented by an avatar in the virtual world) to explore together the environment, and provide information about specific points of interest. A related approach to enhancing the personalization experience has also focused on a more graphical exploratory search technique to allow users to navigate the results of a search query (Lin, Ahn, Brusilovsky, He, & Real, 2010).

GAME ENGINES

Since the early 1990s when really appealing computer game graphics made their first appearance, game engines started to be used as tools in many scientific fields (Harrison, 2003; Craighead, Burke, & Murphy, 2008; Breuer & Bente, 2010). Nowadays, a common approach and trend in modern virtual world

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

applications for various domains is to adopt the technology of game development, the game engines; a simple comparison of contemporary available game engines is included in the following paragraphs.

Game engines are integrated software suites that efficiently use 3D rendering pipelines, special data-structures and speed-up techniques for visualizing texture mapped 3D objects, scenes and 3D worlds in real-time with modern graphics and interaction capacities (Harrison, 2003). The choice of game engines for building dynamic realistic virtual environments was guided by the numerous possibilities and advantages offered by modern game engines and it is not limited to the presented game engines in this study (Lepouras & Vassilakis, 2004; Trenholme & Smith, 2008; Kiourt et al., 2015a; Kiourt et al., 2015b; Kiourt et al., 2016). Game engines are being extensively tested for usability and performance due to the strict requirements of contemporary games. The manufacturers and many computer game developers provide components, algorithms, tools, guides and source code, so that end-users can efficiently create new content. An extensive user support provided by large game development communities makes development platforms even more attractive. Most of the game engine functionalities are managed and exploited via a GUI, for the convenience of developers, providing a more efficient development framework when compared to virtual environment toolkits that often require additional effort to manually provide functionalities such as user interaction (Bowman, Kruijff, LaViola, & Poupyrev, 2005; Cowan & Kapralos, 2015), avatar behaviors, collision detection and management, audio management, avatar-based virtual environment interactions, embodied autonomous agents and many other properties (Lepouras & Vassilakis, 2004; Trenholme & Smith, 2008). Almost all modern game engines provide cross-platform development, which ensures their utilization over a diversity of platforms. There is a wide selection of 3D game engines available for potential use (Lewis & Jacobson, 2008). In the following paragraphs we present the basics of some of the most popular contemporary game engines that are currently available, including *Unity*, *CryEngine* and *Unreal*. We have collected all major advantages and disadvantages regarding these game engines as discussed in (Patel et al., 2003; Craighead et al., 2008; Sillaurren & Aguirrezabal, 2012; Unity3d, 2016; Unreal, 2016; Crytek, 2016; Kiourt et al., 2016). *Unity* (Patel, White, Walczak, & Sayd, 2003; Unity3d, 2016) is a 2D/3D game engine, which offers a wide range of features and a user friendly GUI. One of its main advantages is the cross-platform integration. Each application developed in Unity version 5 (the current version) may be efficiently exported in nearly any available desktop and mobile platform (Windows, Mac OS, PS, XBOX, HTML5, Android, IOS) composing a powerful game engine. It supports 3D content derived from popular 3D applications like Blender, 3DS Max, Maya, CINEMA 4D. It offers limited 3D modeling functionalities based on primitive shapes, although it provides better landscape modeling tools. In general, every 3D model needs to be created in a third party application. However, Unity boasts a large 3D content library where a wide variety of objects can be downloaded for free or be purchased. The Unity engine provides animation mechanisms with real-time editor, particle engines with their own dedicated editor, AI Systems, dynamic lighting and shadowing, water effects, 2D game functionality, networking and multiplayer functionality, integrated physics system, terrain generation system, bitmap image effects, audio/music systems and scripting using C-sharp, JavaScript and Boo programming languages. The *CryEngine* (Craighead, Burke, & Murphy, 2008; Crytek, 2016) is a product *of Crytek Studios* and it has been used by *Ubisoft Montreal* in 2004 to develop the game *Far Cry*. The engine supports a number of features such as a real-time editor with a rich tool arsenal that enable the development of realistic games. Its functionality is similar to that of Unity with some additions such as a dedicated road and river design tool, a solid modeling designer tool and scripting in C++ or LUE programming languages. Most of its development components are managed through a sophisticated GUI that also integrates the CryEngine Sandbox world editing system. The

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

engine supports a vast range of currently available platforms and it is continuously updated to provide export functionalities to new platforms. Licensed developers are offered with the complete source code and documentation for the whole engine and its accompanying tools. One of the primary advantages of CryEngine is the solid designer tool, by which complex meshes may be created. The *Unreal engine* (Craighead, Burke, & Murphy, 2008; Unreal, 2016) was developed by *Epic Games* in C++. The engine supports high performance rendering, advanced animation features and high-quality dynamic lighting, environmental effects, particle system, blueprint visual scripting, C++ scripting, Virtual Reality, AI avatar system, terrain builder, physically-based rendering, UI, level building, animation, visual effects, physics, networking, and cross-platform development. One of the primary advantages of the engine is its powerful particle system that can handle up to one million particles real-time visualization.

Kiourt et al. (2016) reviewed these game engines and made up an interesting and compact comparative table of pros and cons of those engines. Table 1 is an updated version of that table.

Table 1. A simple three-level comparison of various features in popular game engines

Feature	Unity3D v5	CryENGINE v5	Unreal Engine v4
Friendly to User	☺	☹	☹
Components	☺	☺	☺
Object development	☹	☺	😐
GUI	☺	😐	😐
Real time FX	😐	☺	☺
Screen Space effects	😐	☺	☺
Particles	😐	😐	☺
AI	☺	☺	☺
Physics	😐	☺	☺
Animation Systems	😐	😐	😐
Programming	☺	😐	😐
OpenGL	☺	😐	😐
2D Games	☺	☹	😐
3D Formats	😐	😐	😐
Level of Detail rendering	😐	😐	😐
Visual scripting	😐	☺	☺
C#	☺	☺	☹
Tools Integration	☺	😐	😐
VR Support	☺	☺	☺
Mobile Development	☺	😐	😐
Cross-platform	☺	☹	😐
Assets	☺	☺	☺
Support	☺	☺	☺

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

RECONSTRUCTION OF THREE-DIMENSIONAL (3D) ENVIRONMENTS

An integral part of a realistic simulation for cultural heritage applications is the environment, in the form of any cultural object, site, architectural complex, or any other form of tangible material object that conveys the story to be told. *Reconstruction* is the keyword for the creation of the 3D digital counterpart of a real-world cultural object. The reconstruction of 3D environments is based on two entirely distinctive methods: (a) *3D modeling* of objects, which is the process of developing a mathematical (graphical) representation of any 3D surface of an object via specialized software, such as Autodesk 3DS MAX (3DS MAX, 2016), Blender (Blender, 2016); (b) *3D digitization* of objects, which is the process of capturing the shape and appearance of real objects using geometric and spectral measurements. In practice, in many cases those two methods are used in combination and the results are impressive (in terms of rendering realism).

3D modeling is an entirely digital process that takes places on a computer using a specialized software application. It requires high specialization in using such software, and, in particularly demanding applications – like in cultural heritage reconstruction – 3D modeling requires the participation of domain experts. *Obviously, 3D modeling is more of an artist's work, whereas 3D digitization is more of an engineer's work.* 3D digitization is a process of taking highly accurate measurements using specialized techniques and equipment. Its purpose is to get a highly accurate snapshot of reality, having cultural heritage frozen in time. In general, it has to be conducted by digitization experts, although in some cases, the automation provided makes it easier for non-experts to conduct a simple 3D digitization task. There are several techniques for the 3D digitization in cultural heritage applications, each of which with its pros and cons (Pavlidis, Koutsoudis, Arnaoutoglou et al., 2006). The process of *3D digitalization* consists of two steps, the first step is to collect the data (images, distances, sizes etc.), the second step is to analyze the collected data and develop the objects via specialized algorithms. Nowadays the most common 3D digitalization techniques include a number of methods usually falling into two main categories, namely (a) light-dependent and (b) light-independent methods. The first category includes (1) active and (2) passive methods. In summary, Table 2 presents a list of the contemporary 3D digitization methods. Pavlidis et al. (2007) summarized the various parameters involved in categorizing the cultural objects in three axes, namely the *size and shape, morphological complexity* and *diversity of materials*. This 3D space of cultural objects, graphically illustrated in Figure 9, forms a basis for a number of challenges in 3D digitization projects. Unfortunately, there is not a single 3D digitization method to fit all digitization projects. For example, the lower dark-gray region in Figure 9 corresponds to a region in which laser triangulation methods can be successful, whereas the upper medium-gray region corresponds to a region in which range-scanning methods are efficient and the central circular light-gray region depicts a region in which photogrammetry-based methods such as structure from motion are expected to be successful.

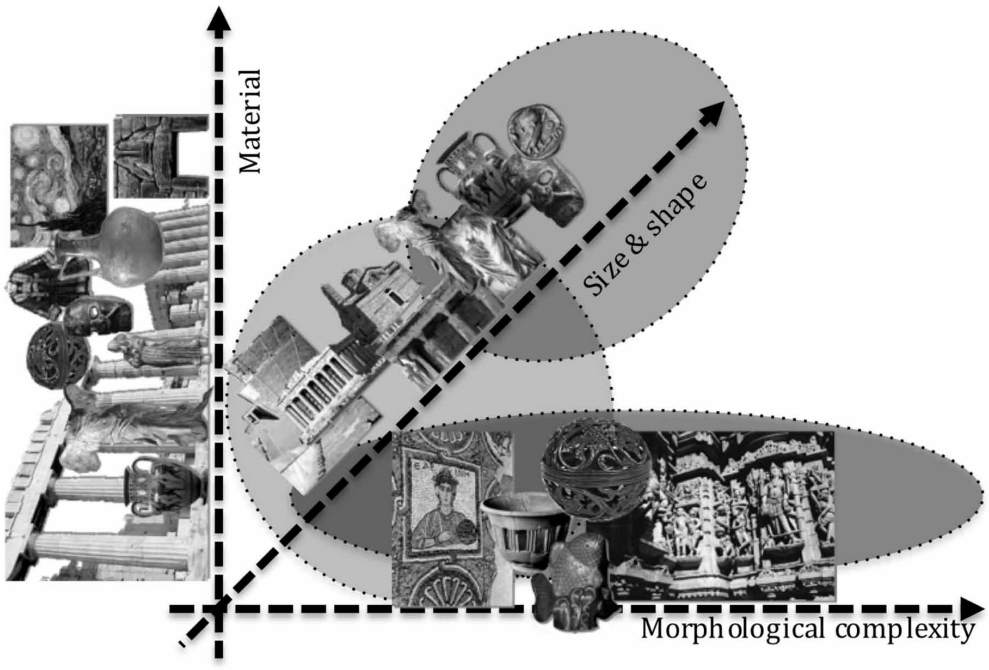
Structure from motion or SfM (Schönberger & Frahm, 2016) is one of the most widely used technique, since it is one of the easiest, most efficient and economical solutions, which does not require mastering any highly-specialized skills and equipment. A camera, a computer and a software application that can be open source, like Visual SFM (VisualSFM, 2016), can produce incredibly realistic models, but, the hardware requirements increase when the quality targeted and the size and complexity of the objects grow. Figure 10 depicts an example of a reconstruction using the SfM method for a typical cultural heritage 3D digitalization project; the blue rectangles represent the resolved camera sensor plane for each captured photo, and the small black vectors help identify the orientation and direction of the camera in the virtual 3D space, whereas the reconstructed 3D model is clearly visible.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Summarized list of 3D digitization methods

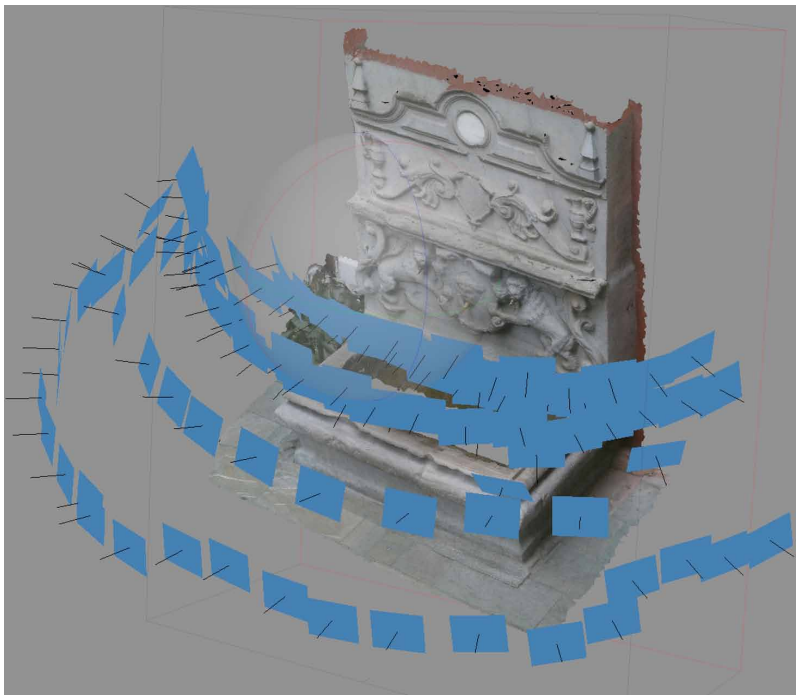
Light-dependent methods	Active methods The light is emitted and is being detected by the digitization system	Laser triangulation
		Time-of-flight scanning (or LiDAR, LADAR, range scanning)
		Structured-light scanning
		Shape from Photometry
		Shape from Shading
		Shape from Shadow
		Tomography (of any kind)
		Speckle photography
		Holography
		Microscopy (of any kind)
	Passive methods The natural ambient light is being used by the digitization system	Photogrammetry
		Structure from Motion
		Shape from Silhouette
		Shape from Stereo
		Shape from Texture
		Shape from Focus (zooming)
Light-independent methods	Topographic methods	
	Empirical methods	
	Contact sensing methods	

Figure 9. The 3D space of cultural objects



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 10. A typical result of the application of SfM for cultural heritage 3D digitization



GAMIFICATION

Developers and administrators of virtual museums have already started to realize the potential of new technologies for the development of edutainment content and services for their visitors (Bickmore, Pfeifer, & Schulman, 2011). In addition, gaming for educational purposes is a significant and active research domain (Brown & Vaughan, 2010; Nicholson, 2011; Nicholson, 2012; Pavlidis, 2015; Seaborn & Fels, 2015). This has taken either the form of game-based learning or serious gaming. *Gamification* is the result of applying game mechanics into diverse domains, in order to engage users and enhance their knowledge and performance.

The importance of playing has been emphasized in many studies from various domains. According to Brown & Vaughan (2010), playing is an archetypical activity that arises from primordial biological structures existing even before the conscience or the capacity for speech; as Brown & Vaughan emphatically stated it is not something a person decides to do.

According to Nicholson (2011; 2012), gamification is nothing more than the use of specific game design approaches and techniques in various environments, in order to attract people in problem solving and to enhance their contribution. Nicholson (Nicholson, 2012) shows the distinction between *Play* and *Game* using a "playful" mathematical approach,

Game = Play + Goals + Structure

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

This equation tries to represent a gamification definition based on the meaning of the terms *game, play, goals* and *structure*. As in any equation, the terms may be rearranged, and interesting equations result, like,

Game – Play = Goals + Structure

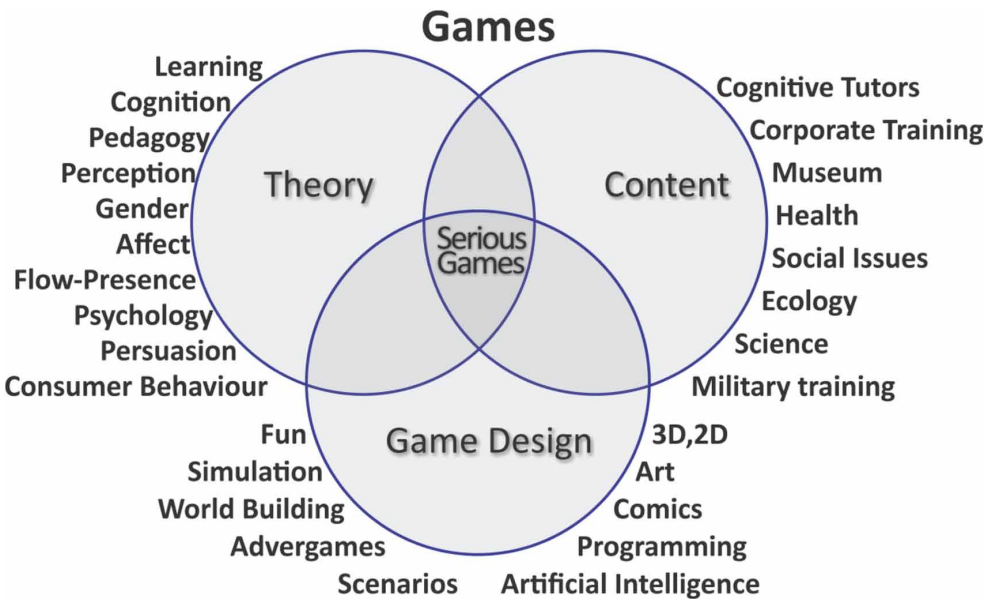
Play = Game – (Goals + Structure)

In the first variant, a *game without* play is not a play-centered activity, and is largely used in gamification strategies like BLAP (Badges, Levels and Leaderboards, Achievements, and Points). A key strategy of meaningful gamification is represented in the second variant, in which by solving for *Play*, participants experience a true playful meaning in the underlying activity (Nicholson, 2012).

Figure 11 represents a cross-domain graph in an attempt to define gamification and specifically serious games, as a kind of a cross section among a number of domains formulated into three major classes that represent the theory, the content and the game design.

Gamification and its special application in education have a long history dating back to the 1980s. Today, with the advent of mature 3D technologies both in hardware and software implementations constitute a stable environment to successfully couple culture and education using technology and gaming as the catalysts. Various independent R&D efforts worldwide exploit this stable environment to deliver the next generation gamified virtual worlds for even more realistic and engaging cultural heritage simulations.

Figure 11. The complex multi-dimensional world of gamification in serious games



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

Realistic simulations of cultural heritage are multidimensional and a cross-disciplinary processes. In this work we made an attempt to summarize most of the key techniques, approaches and technologies available today for developing realistic 3D virtual environments for cultural end educational purposes that aim to create pleasant and engaging virtual environments for all ages. Contemporary efforts focus on various aspects including the representation of real life environments, realistic simulation of human features, interactive interfaces in the world, realistic objects and dynamic environments. In addition, interest in terms of the dynamic processes in gaming environments focuses on competitiveness and cooperation and dynamic feedback. Finally, in terms of smart virtual environments R&D efforts tackle issues in smart guided tours, smart chat-bots and smart interactive entities. Virtual environments cannot and do not intend to replace real environments, either museums or laboratories or anything else. They can be regarded as reflections of physical environments either existing or fictitious, created for various reasons and goals. There are indications that if used complementarily and as extensions of physical environments they may be most effective in enhancing engagement with culture and education with all positive outcomes for a large variety of target groups, ranging from domain experts to kids at primary school.

REFERENCES

Addison, A. C. (2000, April-June). Emerging Trends in Virtual Heritage. *IEEE MultiMedia*, *7*(2), 22–25. doi:10.1109/93.848421

Akenine-Moeller, T., Munkberg, J., & Hasselgren, J. (2007) Stochastic rasterization using time-continuous triangles. *Proceedings of the 22nd ACM SIGGRAPH/EUROGRAPHICS Symposium on Graphics Hardware GH '07*, San Diego, CA, USA (pp. 7–16).

Al-Zinati, M., & Wenkstern, R. (2015). A Self-Organizing Virtual Environment for Agent-Based Simulations. *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, Istanbul, Turkey (pp. 1031-1039).

Amanatides, J. (1987). Realism in Computer Graphics: A Survey. *IEEE Computer Graphics and Applications*, *7*(1), 44–56. doi:10.1109/MCG.1987.277026

Anderson, E., McLoughlin, L., Liarokapis, F., Petridis, C. P., & Freitas, S. (2010). Developing serious games for cultural heritage: A state of the art review. *Virtual Reality (Waltham Cross)*, *14*(4), 255–275. doi:10.100710055-010-0177-3

Arce, T., & Matthias, W. (2002). In-Game Special Effects and Lighting. Retrieved from http://www.nvidia.com/object/gdc_in_game_special_effects.html

Autodesk, 3DS MAX. (2016). Overview. Retrieved from http://www.autodesk.com/products/3ds-max/overview

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Barsky, B. A., Horn, D. R., Stanley, K. A., Jeffrey, P. A., & Yu, M. (2003). Camera models and optical systems used in computer graphics: Part I, object based techniques. *Proceedings of the 2003 International Conference on Computational Science and its Applications (ICCSA'03)* (pp. 246–255). 10.1007/3-540-44842-X_26

Barsky, B. A., & Kosloff, T. J. (2008). Algorithms for Rendering Depth of Field Effects in Computer Graphics. *Proceedings of the 12th WSEAS international conference on Computers* (pp. 999-1010).

Bavoil, L., & Sainz, M. (2009). Multi-layer dual-resolution screen-space ambient occlusion. Proceedings of SIGGRAPH '09, New York, NY, USA. ACM. doi:10.1145/1597990.1598035

Bellotti, F., Berta, R., Gloria, A. D., Dursi, A., & Fiore, V. (2012). A serious game model for cultural heritage. *ACM Journal on Computing and Cultural Heritage*, *5*(4), 17.

Bennewitz, M., Burgard, W., & Thrun, S. (2001). Finding solvable priority schemes for decoupled path planning techniques for teams of mobile robots. Proceedings of the 9th Int. Symposium on Intelligent Robotic Systems (SIRS). *Robotics and Autonomous Systems*, *41*(2-3), 89–99. doi:10.1016/S0921-8890(02)00256-7

Bickmore, T., Pfeifer, L., & Schulman, D. (2011). Relational Agents Improve Engagement and Learning in Science Museum Visitors. *Proceedings of the 10th International Conference on Intelligent Virtual Agents*, Reykjavik, Iceland (pp. 55-67). Springer. 10.1007/978-3-642-23974-8_7

Bineta, M. T. (2013). Edge Based Antialiasing: Quality without Compromise. *International Journal of Engineering Science and Research Technology*, *2*(6), 1627–1631.

Blender. (2016). blender. Retrieved December 15, 2016, Online at Blender: https://www.blender.org

Bogdanovych, A., Rodriguez, J. A., Simoff, S., & Cohen, A. (2009). Virtual Agents and 3D Virtual Worlds for Preserving and Simulating Cultures. Intelligent Virtual Agents. Proceedings of the 9th International Conference, Amsterdam, Netherlands (pp. 257-271). Springer.

Bonis, B., Stamos, J., Vosinakis, S., Andreou, I., & Panayiotopoulos, T. (2009). A platform for virtual museums with personalized content. *Multimedia Tools and Applications*, *42*(2), 139–159. doi:10.100711042-008-0231-2

Bonis, B., Vosinakis, S., Andreou, I., & Panayiotopoulos, T. (2013). Adaptive virtual exhibitions. *DESIDOC Journal of Library and Information Technology*, *33*(3), 183–198. doi:10.14429/djlit.33.4604

Bowman, D., Kruijff, E., LaViola, J. J., & Poupyrev, I. (2005). *3D user interfaces: theory and practise*. USA: Addison Wesley.

Breen, D., House, D. H., & Wozny, M. J. (1994). A particle based model for simulating draping behavior of woven cloth. *Textile Research Journal*, *64*(11), 663–685. doi:10.1177/004051759406401106

Breuer, J., & Bente, G. (2010). Why so serious? On the relation of serious games and learning. *Journal for Computer Game Culture*, *4*(1), 7–24.

Brown, S., & Vaughan, C. (2010). *Play: How it changes the brain, Opens the Imagination, and Invigorates the Soul*. J P Tarcher/Penguin Putnam.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Cai, L., Liu, B., Yu, J., & Zhang, J. (2015). Human behaviors modeling in multi-agent virtual environment. *Multimedia Tools and Applications*.

Callieria, M., Debevecb, P., Pairc, J., & Scopig, R. (2006). A realtime immersive application with realistic lighting: *The Parthenon. Computers & Graphics*, *30*(3), 368–376. doi:10.1016/j.cag.2006.02.015

Catmull, E. (1984) An analytic visible surface algorithm for independent pixel processing. *Proceedings of the 11th Annual Conference on Computer graphics and Interactive Techniques SIGGRAPH '84,* New York, NY, USA (pp. 109–115). ACM. 10.1145/800031.808586

Champagne, J., & Tang, W. (2005). *Real-time simulation of crowds using voronoi diagrams*. EG UK Theory and Practice of Computer Graphics.

Chen, S. E., & Williams, L. (1993). View interpolation for image synthesis. *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques SIGGRAPH '93,* New York, NY, USA (pp. 279–288). ACM. 10.1145/166117.166153

Cohen, M. F., & Wallace, J. R. (1993). *Radiosity and Realistic Image Synthesis*. Academic Press, Inc.

Cook, R., Porter, T., & Carpenter, L. (1984). Distributed Ray Tracing. *ACM SIGGRAPH Computer Graphics*, 18(3), 137-145. 10.1145/800031.808590

Cook, R. L., & Torrance, K. E. (1982). A Reflectance Model for Computer Graphics. *ACM Transactions on Graphics*, *1*(1), 7–24. doi:10.1145/357290.357293

Cowan, B., & Kapralos, B. (2015). A Survey of Frameworks and Game Engines for Serious Game Development. *Proceedings of the IEEE 14th International Conference on Advanced Learning Technologies*, Athens (pp. 662-664).

Craighead, J., Burke, J., & Murphy, R. (2008). Using the unity game engine to develop sarge: A case study. *Proceedings of the Simulation Workshop at the International Conference on Intelligent Robots and Systems*.

Crytek. (2016). Cryengine. Retrieved from http://www.crytek.com/cryengine

Csuri, C., Hackathorn, R., Parent, R., Carlson, W., & Howard, M. (1979). Towards an interactive high visual complexity animation system. *Proceedings of the 6th annual conference on Computer graphics and interactive techniques SIGGRAPH '79* (pp. 289-299). 10.1145/800249.807458

Dachille, F., & Kaufman, A. (2000). High-Degree Temporal Antialiasing. Proceedings of Computer Animation (pp. 49-54). doi:10.1109/CA.2000.889034

Dayal, A., Watson, B., & Luebke, D. (2002). Improving frameless rendering by focusing on change. Proceedings of the *ACM SIGGRAPH 2002 Conference Abstracts and Applications SIG- GRAPH '02* (pp. 201–201). New York, NY, USA: ACM. doi:10.1145/1242073.1242214

Dayal, A., Woolley, C., Watson, B., & Luebke, D. (2005). Adaptive frameless rendering. Proceedings of the SIGGRAPH '05. doi:10.1145/1198555.1198763

DeLeon, V., & Berry, R. Jr. (2000). Bringing VR to the Desktop: Are you Game? *IEEE MultiMedia*, *7*(2), 68–72. doi:10.1109/93.848433

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Demers, J. (2003). Depth of Field in the 'Toys' Demo from Ogres and Fairies: Secrets of the NVIDIA Demo Team. *Presented at GDC '03*.

Demers, J. (2004). Depth of field: A survey of techniques. In GPU Gems, R. Fernando, Ed. Addison Wesley, 23, 375–390.

Demesure, G., Defoort, M., Bekrar, A., Trentesaux, D., & Djemaï, M. (2016). Navigation Scheme with Priority-Based Scheduling of Mobile Agents: Application to AGV-Based Flexible Manufacturing System. *Journal of Intelligent & Robotic Systems*, *82*(3-4), 495–512. doi:10.100710846-015-0273-4

Djaouti, D., Alvarez, J., Rampnoux, O., Charvillat, V., & Jessel, J. (2009). Serious games and cultural heritage: A case study of prehistoric caves. Proceedings of the *15th International Conference on Virtual Systems and Multimedia* (pp. 221–226.). 10.1109/VSMM.2009.40

Doulamis, A., Liarokapis, F., Petridis, P., & Mialous, G. (2012). Serious Games for Cultural Applications. In *Intelligent Computer Graphics* (pp. 97–115). Springer.

Dutre, P., Bala, K., & Bekaert, P. (2006). *Advanced Global illumination*. AK Peters. doi:10.1201/b10632

Eberhardt, B., Weber, A., & Strasser, W. (1996). A fast, flexible, particle- system model for cloth draping. *IEEE Transactions on Computer Graphics and Applications*, *16*(5), 52–59. doi:10.1109/38.536275

Egan, K., Tseng, Y.-T., Holzschuch, N., Durand, F., & Ramamoorthi, R. (2009) Frequency analysis and sheared reconstruction for rendering motion blur. *Proceedings of the SIGGRAPH (ACM Transactions on Graphics)*. 10.1145/1576246.1531399

Foster, N., & Fedkiw, R. (2001). Practical animation of liquids. *Proceedings of ACM SIGGRAPH* (pp. 23–30). New York, ACM Press.

Franklin, C. C. (1977). The aliasing problem in computer-generated shaded images. *Communications of the ACM*, *20*(11), 799–805. doi:10.1145/359863.359869

Gardner, G. Y. (1985). Visual simulation of clouds. *Proceedings of the 12th annual conference on Computer graphics and interactive technique SIGGRAPH '85* (pp. 297–384). 10.1145/325334.325248

Glassner, A. S. (1988). Spacetime ray tracing for animation. *IEEE Computer Graphics and Applications*, *8*(2), 60–70. doi:10.1109/38.504

Gosavi, A. (2003). Simulation-Based Optimization: Parametric Optimization Techniques and Reinforcement Learning. Springer US.

Grant, C. W. (1985). Integrated analytic spatial and temporal anti-aliasing for polyhedra in 4-space. *SIGGRAPH Computer Graphics*, *19*(3), 79–84. doi:10.1145/325165.325184

Guan, X., & Mueller, K. (2004). Point based surface rendering with motion blur. *Proceedings of the Symposium on Point Based Graphics SPBG '04* (pp. 33-40).

Guennebaud, G., Barthe, L., & Paulin, M. (2006). Real-time soft shadow mapping by backprojection. *Proceedings of the 17th Eurographics conference on Rendering Techniques* (pp. 227-234).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hachisuka, T., Jarosz, W., Weistroffer, R. P., Dale, K., Humphreys, G., Zwicker, M., & Jensen, H. W. (2008). Multi-dimensional adaptive sampling and reconstruction for ray tracing. In SIGGRAPH '08: ACM SIGGRAPH 2008 Papers (pp. 33:1–33:10). New York, NY, USA; ACM.

Haeberli, P., & Akeley, K. (1990). The accumulation buffer: hardware support for high-quality rendering. *Proceedings of the 17th Annual Conference on Computer Graphics and Interactive Techniques SIGGRAPH '90* (pp. 309–318). New York, NY, USA: ACM. 10.1145/97879.97913

Hanna, N., & Richards, D. (2014). Towards a 'Smart' Collaborative Virtual Environment and Multi-agent Approach to Designing an Intelligent Virtual Agent. In *Revised Selected and Invited Papers of the 4th International Workshop on Agent Environments for Multi-Agent Systems IV* (pp. 170-187).

Harrison, L. (2003). *Introduction to 3D Game Engine Design Using DirectX 9 and C*. Berkeley: Apress. doi:10.1007/978-1-4302-0821-1

Hasenfratz, J. M., Lapierre, M., Holzschuch, N., & Sillion, F. A.Artis GRAVIRIMAG-INRIA. (2003). A Survey of Real-time Soft Shadows Algorithms. *Computer Graphics Forum*, *22*(4), 753–774. doi:10.1111/j.1467-8659.2003.00722.x

Hastings, E., Guha, R., & Stanley, K. O. (2007). NEAT Particles: Design, Representation, and Animation of Particle System Effects. *Proceedings of the IEEE 2007 Symposium on Computational Intelligence and Games*, Honolulu, HI (pp. 154-160.). 10.1109/CIG.2007.368092

Hinsinger, D., & Neyret, F. (2002). Interactive animation of ocean waves. *Proceedings of ACM SIGGRAPH/Eurographics symposium on computer animation* (pp. 161-166). 10.1145/545261.545288

Hourcade, C., & Nicolas, A. (1985). Algorithms for antialiased cast shadows. *Computers & Graphics*, *9*(3), 259–265. doi:10.1016/0097-8493(85)90052-4

Huang, Z., Gong, G., & Liang, H. (2015). Physically-based smoke simulation for computer graphics: A survey. *Multimedia Tools and Applications*, *74*(18), 7569–7594. doi:10.100711042-014-1992-4

Ichiro, S. (2008). Context-Aware Agents to Guide Visitors in Museums. *Proceedings of the 8th International Conference Intelligent Virtual Agents* (pp. 441-455). Tokyo, Japan: Springer.

Iones, A., Krupkin, A., Sbert, M., & Zhukov, S. (2003). Fast, Realistic Lighting for Video Games. *IEEE Computer Graphics and Applications*, *23*(3), 54–64. doi:10.1109/MCG.2003.1198263

Jim, X., Xiadong, F., & Wegman, J. (1999). Real-time simulation of dust behavior generated by a fast traveling vehicle. *ACM Transactions on Modeling and Computer Simulation*, *9*(2), 81–104. doi:10.1145/333296.333366

Jones, N., & Keyser, J. Real-time geometric motion blur for a deforming polygonal mesh. *Proceedings of the Computer Graphics International 2005,* Washington, DC, USA (pp. 14–18). IEEE. 10.1109/CGI.2005.1500358

Jonghyun, K., Jae-Hyun, J., Changwon, J., & Byoungho, L. (2013). Real-time capturing and 3D visualization method based on integral imaging. *Optics Express*, *21*(16), 18742–18753. doi:10.1364/OE.21.018742 PMID:23938790

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kim, D., & Ko, H.-S. (2007) Eulerian motion blur. *Proceedings of the Eurographics Workshop on Natural Phenomena EG-WNP '07,* Prague, Czech Republic (pp. 39–46). ACM.

Kiourt, C., Koutsoudis, A., Arnaoutoglou, F., Petsa, G., Markantonatou, S., & Pavlidis, G. (2015a). The 'Synthesis' Virtual Museum–an open virtual exhibition creation tool. *Proceedings of the International Workshop on Virtual Archaeology: Museums & Cultural Tourism*. Delphi, Greece.

Kiourt, C., Koutsoudis, A., Arnaoutoglou, F., Petsa, G., Markantonatou, S., & Pavlidis, G. (2015b). *A dynamic web-based 3D virtual museum framework based on open data*. Granada, Spain: Digital Heritage. doi:10.1109/DigitalHeritage.2015.7419589

Kiourt, C., Koutsoudis, A., & Pavlidis, G. (2016). DynaMus: A fully dynamic 3D virtual museum framework. *Journal of Cultural Heritage*, *22*, 984–991. doi:10.1016/j.culher.2016.06.007

Korein, J., & Badler, N. (1983). Temporal Anti-Aliasing in Computer Generated Animation. Proceedings of Computer Graphics SIGGRAPH '83 (pp. 377–388). doi:10.1145/800059.801168

Koutsoudis, A., Makarona, C., & Pavlidis, G. (2012). Content-based navigation within virtual museums. *Journal of Advanced Computing Science and Technology*, *1*(2), 73–81. doi:10.14419/jacst.v1i2.135

Koutsoudis, A., & Pavlidis, G. (2011). Content-based search and retrieval within complex 3d scenes - an application to cultural heritage. *Proceedings of the 16th International Conference on Cultural Heritage and New Technologies*.

Lahijanian, M., Maly, M. R., Fried, D., Kavraki, L. E., Kress-Gazit, H., & Vardi, M. (2016). Iterative Temporal Planning in Uncertain Environments with Partial Satisfaction Guarantees. *IEEE Transactions on Robotics*, *32*(3), 583–599. doi:10.1109/TRO.2016.2544339

Lepouras, G., & Vassilakis, C. (2004). Virtual museums for all: Employing game technology for edutainment. *Virtual Reality (Waltham Cross)*, *8*(2), 96–106. doi:10.100710055-004-0141-1

Lewis, M., & Jacobson, J. (2008). Game engines in scientific research. *Communications of the ACM*, *45*(1), 27–31.

Li, T. T., & Chou, H. C. (2003). Motion planning for a crowd of robots. *Proceedings of the IEEE Int. Conf. on Robotics and Automation*.

Lin, H., & Chang, C. (2006) Photo-consistent motion blur modeling for realistic image synthesis. *Proceedings of the Pacific-Rim Symposium on Image and Video Technology* (pp. 1273–1282). 10.1007/11949534_129

Lin, Y., Ahn, J.-w., Brusilovsky, P., He, D., & Real, W. (2010). ImageSieve: Exploratory Search of Museum Archives with Named Entity-Based Faceted Browsing. In C. Marshall, E. Toms and A. Grove (Eds.), *Proceedings of 2010 Annual Meeting of The American Society for Information Science & Technology (ASIS&T 2010)*, Pittsburgh, PA.

Loviscach, J. (2005) Motion blur for textures by means of anisotropic filtering. *Proceedings of the Eurographics Symposium on Rendering* (pp. 7–14).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Luna, R., & Lahijanian, M., Moll, & M., Kavraki L.E. (2014). Optimal and Efficient Stochastic Motion Planning in Partially-Known Environments. *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, Quebec City, Canada (pp. 2549–2555).

Max, N. L., & Lerner, D. M. (1985). A two-and-a-half-d motion-blur algorithm. *SIGGRAPH Computer Graphics.*, *19*(3), 85–93. doi:10.1145/325165.325188

McAllister, D. K. (2000). *The Design of an API for Particle Systems*. Chapel Hill, NC, USA: University of North Carolina at Chapel Hill, Department of Computer Science.

McGuire, M., Hennessy, P., Bukowski, M., & Osman, B. (2012). A reconstruction filter for plausible motion blur. *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, Costa Mesa, CA (pp. 135-142). 10.1145/2159616.2159639

Miller, G. (1994). Efficient algorithms for local and global accessibility shading. *Proceedings of the 21st annual conference on Computer graphics and interactive techniques SIGGRAPH '94* (pp. 319-326). 10.1145/192161.192244

Musse, S. R., & Thalmann, D. (1997). A model of human crowd behavior: Group inter-relationship and collision detection analysis. In *Computer Animation and Simulation* (pp. 39-51).

Naderi, K., Rajamäki, J., & Hämäläinen, P. (2015). RT-RRT*: a real-time path planning algorithm based on RRT*. *Proceedings of the 8th ACM SIGGRAPH Conference on Motion in Games*, Paris, France (pp. 113-118). 10.1145/2822013.2822036

Nathan, J., & Keyser, J. (2005). Real-Time Geometric Motion Blur for a Deforming Polygonal Mesh. Proceeding of Computer Graphics International, Stony Brook, NY, USA (pp. 26-31).

Nehab, D., Sander, P. V., Lawrence, J., Tatarchuk, N., & Isidoro, J. R. (2007) Accelerating real-time shading with reverse reprojection caching. *Proceedings of the 22nd ACM SIGGRAPH/EUROGRAPHICS Symposium on Graphics Hardware GH '07,* Aire-la-Ville, Switzerland (pp. 25–35).

Nguyen, D. Q., Fedkiw, R., & Jensen, H. W. (2002). Physically based modeling and animation of fire. *Transactions on Graphics*, *21*(3), 721–728.

Nicholson, S. (2011). Making the Gameplay Matter: Designing Modern Educational Tabletop Games. *Knowledge Quest*, *40*(1), 60–65.

Nicholson, S. (2012). Strategies for meaningful gamification: Concepts behind transformative play and participatory museums. Presented at Meaningful Play 2012, Lansing, Michigan.

Oberlander, J., Karakatsiotis, G., Isard, A., & Androutsopoulos, A. (2008). Building an adaptive museum gallery in Second Life. *Proceedings of Museums and the Web 2008*, Montreal, Quebec, Canada.

Overbeck, R. S., Donner, C., & Ramamoorthi, R. (2009). Adaptive Wavelet Rendering. *ACM Transactions on Graphics*, *28*(5), 140. doi:10.1145/1618452.1618486

Pachur, D., Laue, T., & Roefer, T. (2009). *Real-Time Simulation of Motion-Based Camera Disturbances. Proceedings of RoboCup 2008: Robot Soccer World Cup XII* (pp. 591–601). Berlin, Heidelberg: Springer-Verlag. doi:10.1007/978-3-642-02921-9_51

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Papagiannakis, G., L'Hoste, G., Foni, A., & Thalmann, N. M. (2001). Real-time photo realistic simulation of complex heritage edifices. *Proceedings Seventh International Conference on Virtual Systems and Multimedia* (pp. 218-227). 10.1109/VSMM.2001.969674

Papastamatiou, N., Alexandridis, T., Tsergoulas, K., Michopoulos, A., & Karadimas, N. V. (2010). Virtual reality applications with user interface for dynamic content development. *Proceedings of the 4th WSEAS international conference on Computer engineering and applications CEA'10* (pp. 201–206).

Patel, M., White, M., Walczak, K., & Sayd, P. (2003). Digitisation to presentation: Building virtual museum exhibitions. In *Vision, Video and Graphics*.

Pavlidis, G. (2015). Towards a playful education. *Proceedings of the International Workshop on Virtual Archaeology: Museums & Cultural Tourism*, Delphi, Greece.

Pavlidis, G., Koutsoudis, A., Arnaoutoglou, F., Tsioukas, V., & Chamzas, C. (2007). Methods for 3D digitization of Cultural Heritage. *Journal of Cultural Heritage*, *8*(1), 93–98. doi:10.1016/j.culher.2006.10.007

Pavlidis, G., Makarona, C., Arnaoutoglou, F., Koutsoudis, A., & Chamzas, C. (2008). Museums In The Virtual World. *Proceedings of the International Symposium Information and Communication Technologies in Cultural Heritage*, Ioannina, Greece (pp. 16-18).

Pavlidis, G., Tsiafakis, N., Provopoulos, G., Chatzopolos, S., Arnaoutoglou, F., & Chamzas, C. (2006, June 5-9). Momi: A dynamic and internet-based 3d virtual museum of musical instruments. *Proceedings of the Third International Conference of Museology*.

Pelechano, N., Allbeck, J., & Badler, N. (2007). Controlling individual agents in high-density crowd simulation. *Proceedings of the ACM SIGGRAPH/Eurographics symposium on Computer animation*, San Diego, California, USA (pp. 99–108).

Pettre, J., Laumond, J. P., & Thalmann, D. (2005). A navigation graph for real-time crowd animation on multilayered and uneven terrain. *Proceedings of the First International Workshop on Crowd Simulation*.

Phong, T. B. (1973). *Illumination for computer generated images* [PhD. Dissertation]. University of Utah, Salt Lake City, Utah, USA.

Potmesil, M., & Indranil, C. (1981). A Lens and Aperture Camera Model for Synthetic Image Generation. *Proceedings of the 8th Annual Conference on Computer Graphics and Interactive Techniques* (pp. 297–305). 10.1145/800224.806818

Potmesil, M., & Indranil, C. (1982). Synthetic image generation with a lens and aperture camera model. *ACM Transactions on Graphics*, *1*(2), 85–108. doi:10.1145/357299.357300

Prem, K., Thalmann, N. M., Moccozet, L., Sannier, G., Aubel, A., & Thalmann, D. (1998). Real-Time Animation of Realistic Virtual Humans. *IEEE Computer Graphics and Applications*, *18*(5), 42–56. doi:10.1109/38.708560

Ragan-Kelley, J., Lehtinen, J., Chen, J., Doggett, M., & Durand, F. (2010). Decoupled Sampling for Real- Time Graphics Pipelines (tech. report). Massachusetts Institute of Technology.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Reeves, W. T. (1983). Particle Systems — a Technique for Modeling a Class of Fuzzy Objects. *ACM Transactions on Graphics*, *2*(2), 91–108. doi:10.1145/357318.357320

Reeves, W. T. (1985). Approximate and probabilistic algorithms for shading and rendering structured particle systems. *Proceedings of the 12th annual conference on Computer graphics and interactive techniques SIGGRAPH '85* (pp. 313-322). 10.1145/325334.325250

Remondino, F., & El-Hakim, S. (2006). Image-based 3D Modelling: A Review. *The Photogrammetric Record*, *21*(115), 269–291. doi:10.1111/j.1477-9730.2006.00383.x

Reynolds, C. (1999). Steering behaviors of autonomous characters. *Proceedings of Game Developers Conference*, San Francisco, CA (pp. 763-782).

Ritchings, T., Anastassakis, G., & Panayiotopoulos, T. (2001). Multi-agent Systems as Intelligent Virtual Environments. *Proceedings of the Advances in Artificial Intelligence: Joint German/Austrian Conference on AI (KI '01)* (pp. 381-395).

Rogers, A. (1996). *Teaching Adults* (2nd ed.). Buckingham: Open University Press.

Schmid, J., Sumner, R., Bowles, H., & Gross, M. (2010). Programmable motion effects. *ACM Transactions on Graphics,* 29(3). 10.1145/1833349.1778794

Schönberger, J. L., & Frahm, J. M. (2016). Structure-from-Motion Revisited. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*

Scofield, C. (1994). 21/2-D Depth of Field Simulation for Computer Animation. In D. Kirk (Ed.), *Graphics Gems III* (pp. 36–38). Morgan Kaufmann.

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, *87*, 14–31. doi:10.1016/j.ijhcs.2014.09.006

Selgrad, K., Reintges, C., Penk, D., Wagner, P., & Stamminger, M. (2015). Real-time depth of field using multi-layer filtering. *Proceedings of the 19th Symposium on Interactive 3D Graphics and Games,* San Francisco, CA (pp. 121-127). 10.1145/2699276.2699288

Shao, W., & Terzopoulos, D. (2005). Autonomous Pedestrians. *Proceedings of SIGGRAPH/EG Symposium on Computer Animation (SCA'05)*, Los Angeles, CA.

Sillaurren, S., & Aguirrezabal, P. (2012). 3d publish: A web-based solution for building dynamic 3d virtual museums. *International Journal of Heritage in the Digital Era*, *1*(1_suppl1 Suppl.), 147–152. doi:10.1260/2047-4970.1.0.147

Silva, R., Melo, F., & Veloso, M. (2016). Adaptive Symbiotic Collaboration for Targeted Complex Manipulation Tasks. *Proceedings of ECAI'16, the 22nd European Conference on Artificial Intelligence*, The Hague, Holland (pp. 863-870).

Sookhanaphibarn, K., & Thawonmas, R. (2009). A content management system for user-driven museums in second life. *Proceedings of the International Conference on CyberWorlds* (pp. 185–189). 10.1109/CW.2009.58

Sousa, T. (2008) Crysis next-gen effects. *Proceedings of the Game Developers Conference 2008*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Stüvel, S. A., de Goeij, M. F., van der Stappen, A. F., & Egges, A. (2015). An analysis of manoeuvring in dense crowds. *Proceedings of the 8th ACM SIGGRAPH Conference on Motion in Games*, Paris, France (pp. 85-90). 10.1145/2822013.2822037

Sud, A., Andersen, E., Curtis, S., Lin, M., & Manocha, D. (2008). Real-time path planning for virtual agents in dynamic environments. Proceedings of the ACM SIGGRAPH 2008 classes, Los Angeles, CA. doi:10.1145/1401132.1401206

Sung, K., Pearce, A., & Wang, C. (2002). Spatial-temporal antialiasing. *IEEE Transactions on Visualization and Computer Graphics*, *8*(2), 144–153. doi:10.1109/2945.998667

Sylaiou, S., Liarokapis, F., Kotsakis, K., & Patias, P. (2009). Virtual museums, a survey and some issues for consideration. *Journal of Cultural Heritage*, *4*, 520–528.

Sylaiou, S., Mania, K., Karoulis, A., & White, M. (2010). Exploring the relationship between presence and enjoyment in a virtual museum. *International Journal of Human-Computer Studies*, *68*(5), 243–253. doi:10.1016/j.ijhcs.2009.11.002

Sylaiou, S., Mania, K., Paliokas, I., Pujol, L., Kilintzis, V., & Liarokapis, F. (in press). Exploring the educational impact of diverse technologies in online Virtual Museums. *International Journal of Arts and Technology.*

Tan, j., & Fan, X. (2011). Particle System Based Snow Simulating in Real Time. *Proceedings of the 3rd International Conference on Environmental Science and Information Application Technology ESIAT 2011* (pp. 1244-1249).

Telleen, J., Sullivan, A., Yee, J., Wang, O., Gunawardane, P., Collins, I., & Davis, J. (2007). Synthetic shutter speed imaging. *Computer Graphics Forum*, *26*(3), 591–598. doi:10.1111/j.1467-8659.2007.01082.x

Trenholme, D., & Smith, S. P. (2008). Computer game engines for developing first-person virtual environments. *Virtual Reality (Waltham Cross)*, *12*(3), 181–187. doi:10.100710055-008-0092-z

Treuille, A., Cooper, S., & Popović, Z. (2006). Continuum crowds. *ACM Transactions on Graphics*, *25*(3), 1160–1168. doi:10.1145/1141911.1142008

Tsichritzis, D., & Gibbs, S. (1991). Virtual museums and virtual realities. *Proceedings of the International Conference on Hypermedia and Interactivity in Museums* (pp. 17-25).

Unity3d. (2016). Unity Technologies. Retrieved from http://unity3d.com/

UnrealEngine. (2016). Retrieved from https://www.unrealengine.com/

Van Eck, R. (2006). Digital game-based learning: It's not just the digital natives who are restless. *EDUCAUSE*, *41*(2), 16–30.

Vardis, K., Papaioannou, G., & Gaitatzes, A. (2013). Multi-view Ambient Occlusion with Importance Sampling. *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, Orlando, FL, USA (pp. 111-118). 10.1145/2448196.2448214

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Varshavskaya, P., Kaelbling, L. P., & Rus, D. (2008). Efficient Distributed Reinforcement Learning Through Agreement. *Proceedings of the 9th International Symposium on Distributed Autonomous Robotic Systems (DARS)*, Tsukuba Japan.

Vassos, S., Malliaraki, E., dal Falco, F., Di Maggio, J., Massimetti, M., Nocentini, M. G., & Testa, A. (2016). Art-bots: Toward Chat-based Conversational Experiences in Museums. *Proceedings of the 9th International Conference on Interactive Digital Storytelling* (pp. 433-437). Los Angeles, CA, USA: Springer International Publishing. 10.1007/978-3-319-48279-8_43

Verbeck, P. C., & Greenberg, P. D. (1984). A comprehensive light source description for computer graphics. *IEEE Computer Graphics and Applications*, *4*(7), 66–75. doi:10.1109/MCG.1984.275906

VisualSFM. (2016). Retrieved from http://ccwu.me/vsfm/

Vlachos, A. (2008). Advanced Visual Effects with Direct3D: Post Processing in The Orange Box. *Proceedings of the Game Developer's Conference 2008*, San Francisco, CA.

Walter, B., Arbree, A., Bala, K., & Greenberg, D. P. (2006). Multidimensional lightcuts. Proceedings of SIGGRAPH '06 (pp. 1081–1088). ACM. doi:10.1145/1179352.1141997

William, L. L. (1980). Human Vision, Anti-aliasing, and the Cheap 4000 Line Display. *Computer Graphics*, *14*(3), 308–313. doi:10.1145/965105.807509

Wojciechowski, R., Walczak, K., White, M., & Cellary, W. (2004). Building virtual and augmented reality museum exhibitions. *Proceedings of the ninth international conference on 3D Web technology* (pp. 135-144). 10.1145/985040.985060

Woo, A., Poulin, P., & Fournier, A. (1990). A Survey of Shadow Algorithms. *IEEE Computer Graphics and Applications*, *10*(6), 13–32. doi:10.1109/38.62693

Yang, L., Dong, H., Alelaiwi, A., & El Saddik, A. (2015). See in 3D: State of the art of 3D display technologies. *Multimedia Tools and Applications*, *75*(24), 17121–17155.

Yngve, G. D., Obrien, J. F., & Hodgins, J. K. (2000). Animation explosions. *Proceedings of ACM SIGGRAPH 2000 conference proceedings* (pp. 29–36). ACM Press.

Yu, I., Mortensen, J., Khanna, P., Spanlang, B., & Slater, M. (2012). Visual realism enhances realistic response in an immersive virtual environment. *IEEE Computer Graphics and Applications*, *32*(6), 36–45. doi:10.1109/MCG.2012.121 PMID:24807308

Yu, X., Wang, R., & Yu, J. (2010). Real-time depth of field rendering via dynamic light field generation and filtering. *Computer Graphics Forum*, *39*(7), 2099–2107. doi:10.1111/j.1467-8659.2010.01797.x

Zheng, Y., Koestler, H., Thuerey, N., & Ruede, U. (2006). Enhanced motion blur calculation with optical flow. *Proceedings of Vision* (pp. 253–260). Modeling and Visualization.

This research was previously published in the International Journal of Computational Methods in Heritage Science (IJCMHS), 1(1); edited by George P. Pavlidis; pages 10-40; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 65
Enhanced Virtual Reality Experience in Personalised Virtual Museums

Chairi Kiourt
Athena Research Centre, Greece

Anestis Koutsoudis
Athena Research Centre, Greece

Dimitris Kalles
Hellenic Open University, Greece

ABSTRACT

This article focuses on important factors in the creation of enhanced personalised experiences in virtual environments for cultural heritage applications, especially those targeting virtual museums and exhibitions. Some of the most important factors relating to personalised virtual museums that relate to intelligent content and user modelling in virtual environments are being highlighted and discussed. After an extensive review of the current trends in the domain, the article presents a generalised framework for the development of the next generation enhanced VR experiences in personalised virtual museums. This framework, which naturally surfaces from the domain, was put to the test in the development of the DynaMus platform and two case studies based on this platform are referenced and commented to support such an approach. This concept can serve as the general framework for developing enhanced personalised virtual environments for cultural heritage applications.

INTRODUCTION

The advent of the web, was an initial motivation for museums, to start creating information-rich web sites to promote their services. Personalised content was considered a good strategy to support users, with different capabilities and requirements, in finding information in unfamiliar collections (Ardissono, et

DOI: 10.4018/978-1-7998-0951-7.ch065

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

al., 2015). As soon as mobile technology appeared, it was adopted for delivering context-aware cultural heritage information both indoors and outdoors. Generally speaking, indoors there is a limited space, which is extremely rich in content but offers few services over and above information delivery. In contrast, outdoor mobile apps cover relatively large areas – a city or part of it – offering a wide range of integrated services, such as those providing location-based, cultural, hotel and travel information. The infusion of new technologies in cultural institutions, such as mobile technologies (Asif & Krogstie, 2012), Virtual Reality (VR) (Kiourt, Pavlidis, Koutsoudis, & Kalles, 2017a), Augmented Reality (AR) (White, et al., 2004; Katifori, et al., 2014; Pujol, et al., 2012) etc. have resulted in new experiences (spatial awareness) and have helped in reaching a larger audience – not only children but also adult visitors (Kuflik, Kay, & Kummerfeld, 2010) – focusing on several different goals such as education, lifelong learning etc. (Ardissono, Kuflik, & Petrelli, 2012; Ardissono, et al., 2015; Kiourt, Pavlidis, Koutsoudis, & Kalles, 2017a).

Personalised content in cultural heritage aims at adjusting and transforming the cultural experience so as to meet the experiences, interests and knowledge at the level of the individual visitor or the group (Ardissono, Kuflik, & Petrelli, 2012; Yiannoutsou & Avouris, 2014). Personalised content is considered by many researchers (Ardissono, Kuflik, & Petrelli, 2012; Asif & Krogstie, 2012; Bohnert, 2010; Kuflik, Kay, & Kummerfeld, 2010; Not & Petrelli, 2014; Vayanou, et al., 2012) as a key factor in the success of such objectives. Content personalisation techniques can be employed to steer the information selection based on the user / group profile (interests and context), thus protecting them from data overload (Ardissono, Kuflik, & Petrelli, 2012). Moreover, personalisation can be used to adapt the presentation of information to the user's device, thus facilitating exploration. Personalised content delivery to users is more acceptable through gamified applications, such as serious games, because they are more realistic, user-friendly, pleasant and attractive systems (Anderson, et al., 2010).

There is a growing trend for the development of serious games in diverse fields, including entertainment, cultural heritage, education, artificial intelligence, sociology, military and health systems (Abt, 1987; Breuer & Bente, 2010). Serious games are highlighted as innovative and important tools for the development of applications broadly accepted by all ages, children, adults and seniors (David & Sandra, 2005). Serious games have been widely adopted in cultural heritage, for applications such as virtual exhibitions and museums and have been the subject of applied research and development for several years (Mortara, et al., 2014; Kiourt, Pavlidis, Koutsoudis, & Kalles, 2017a; Doulamis, Liarokapis, Petridis, & Miaoulis, 2012; Froschauer, Arends, Goldfarb, & Merkl, 2012). Nevertheless, there are still challenges for further research, targeting the enhanced interactive realistic simulations of ancient worlds or cultural sites and exhibitions. In a sense, serious games can be considered an efficient approach for blending domain specific activities, like those in cultural heritage and education, with gaming.

Gamification (Deterding, Dixon, Khaled, & L., 2011) is the result of applying game mechanics into diverse domains, in order to engage users and enhance their knowledge and performance. The importance of playing has been emphasised in many studies from various domains. According to Brown & Vaughan (Brown & Vaughan, 2010), playing is an archetypical activity that arises from primordial biological structures existing even before the conscience or the capacity for speech; as Brown & Vaughan emphatically stated, playing is not something a person decides to do. According to Nicholson (Nicholson, 2011), gamification is nothing more than the use of specific game design approaches and techniques in various environments, in order to attract people in problem solving and to enhance their contribution.

This paper focuses on serious games-based cultural heritage applications (mainly as virtual museums) with personalised content for more accurately targeted information and knowledge delivery to users. Of special interest is to derive a framework of rules, theory and best practices for the development of the

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

virtual environment, the content adaptation and the interaction mechanics. To this end, key factors relating to Personalised Virtual Museums (PVM) (Hayashi, Bachelder, & Nakajima, 2016) that are based on a dynamic environment that is able to be populated by external resources and provides content based on user profiles. Nowadays, education, training and entertainment seem to be main research directions in PVMs. This paper introduces a framework for the development of user profiles based on social media sources and for the exploitation of this user profile for the development of PVMs. The approach presented introduces three core research sub-domains (content generation, knowledge modelling and engaging gameplay), which can serve as the general framework for developing enhanced virtual environments by exploiting the user profile.

The rest of this paper is structured in four sections. The next (second) section presents a summarised review of recent advances in the domain of PVMs. The third section identifies the basic ingredients of PVMs that support an enhanced virtual reality experience. The fourth section provides a presentation of the unifying proposed framework followed by two user-centred virtual world case studies with the DynaMus platform that was developed based on this general framework. The paper concludes by summarising and providing thoughts for further development.

RELATED WORK

The personalization of virtual museums theory, appeared about two decades ago, when some of the adaptive hypermedia systems looked at museum content and tourism as possible application domains (Stock, et al., 1993; Fink, Kobsa, & Nill, 1998). Stock et al. (1993) provided personalized content as part of a human-machine dialogue on art exhibits and combined with natural language processing with a hypermedia system. While Fink, Kobsa and Nill (1998) introduced a desktop application running on info kiosks for personalized tourism information delivery. Paternò and Mancini (1999) developed an online web based application with personalized content by providing three different choices to visitors: 1) predefined information views; 2) user defined profile setup; and 3) information suggestion based on the visiting history of the visitor. The 90's decade stammered the advent of the web based museums started creating information-rich websites, and personalization was considered a good way to support visitors with different needs. One decade later the mobile technologies appeared as an important tool for context-aware cultural heritage information delivery to visitors for both indoor and outdoor areas. Schmidt-Belz, Laamanen, Poslad and Zipf (2003) tested Personal Digital Assistants (PDAs) for providing dynamic and interactive maps, to visitors, that presents the current position, recommendations, information about attractions, and visiting tips. Petrelli and Not (2005) presented a system with accurate determination of the visitor's position indoors, which led them to the idea of interaction through physical movement. Stock et al. (2007) introduced a system which dynamically creates animation-based video clips taking into account the user's focus of attention. In this system, a virtual character appears in the PDA and accompanied the visitor, presenting information, throughout the visit. Projects such as MUSE and Smartmuseum (Ruotsalo, et al., 2009) tried to cover both cases (indoor and outdoor guidance) but most were specialized in one or the other.

In the creation of virtual museums, contemporary IT technologies, such as multimedia, 3D computer graphics, spatial sound and virtual reality, are used to enhance the presentation (the virtual presence), offering a more vivid and enjoyable experience (Kiourt, Pavlidis, Koutsoudis, & Kalles, 2017a) through gamified virtual environments. The Augmented Representation of Cultural Objects (ARCO)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

(Patel, White, Walczak, & Sayd, 2003; White, et al., 2004) was among the predominant efforts towards a dynamic virtual museum system accompanied by a 3D digitisation technique to provide museums a framework to produce and Exhibit 3D digital replicas of their artefacts over the Web based on the visitors need. Lepouras and Vassilakis (2005) presented the concept of creating virtual museums focused primarily on educational content and related services by using a game engine. They developed a virtual museum for multiple different uses, based on different educational content. Wang, Aroyo, Stash and Rutledge (2007) presented an approach for personalized access to museum collections by using RDF/OWL specification of a museum collections as a driver for an interactive dialog. Personalization of gamified virtual museums include virtual agents-guides that discuss and interacts with users and present the space (Oberlander, Karakatsiotis, Isard, & Androutsopoulos, 2008), and adaptive, personalized exhibition spaces (Bonis, Stamos, Vosinakis, Andreou, & Panayiotopoulos, 2009). Liarokapis, Sylaiou and Mountain (2008) presented a solution for indoor and outdoor museum exhibitions with personalised content. They introduced mixed reality multi-modal interfaces as a framework that can be used for targeting the needs of visitors for both kiosk exhibitions and mobile guides. Sookhanaphibarn & Thawonmas (2009), developed a virtual museum, which was equipped with an innovative intelligent guidance system that was able to provide a customised navigation route based on the visitor's preferences. Christopoulos et al. (2011) introduced a story telling virtual game with dynamic interactions that gives the opportunity to a human guide for personalization of the experience according to the target group. Koutsoudis et al. (2012), proposed a content-based navigation framework for a virtual museum, based on metadata that describe the exhibits and thus providing a semantic similarity-based navigation. The visitor's preferences are decoded by its interactions with the exhibits. Coenen, Mostmans and Naessens (2013), introduced a gamified virtual environment, where the visitors where guided by the system (mobile application) to create their own exposition during their visit in the physical museums. The aim of this systems is to provide a personalised learning effect during a visit to a museum exhibition. Bonis, Vosinakis, Andreou, and Panayiotopoulos, (2013) proposed a multi-user framework for virtual exhibitions that adapts to visitors' preferences. Kiourt, Koutsoudis and Pavlidis (2016) have studied many aspects of interactive virtual museums mainly focusing on interactive 3D object exhibitions with dynamic open data content; the virtual museum was reintroduced as a cultural and educational tool that focuses on user-driven exhibitions. The last three decades, in the view of gamified virtual museums for personalised learning applications have been developed several applications, such as Roma Nova (B. Frischer, 2008), Ancient Pompeii (Troche & Jacobson), THIATRO (Froschauer, Merkl, Arends, & Goldfarb, 2013), Ancient Agora of Athens (Kontogianni & Georgopoulos, 2015; Kontogianni, Koutsaftis, Skamantzari, Georgopoulos, & Chrysanthopoulou, 2016) and many more (Anderson, et al., 2010; Mortara, et al., 2014; Sylaiou, Liarokapis, & Kotsakis, 2009; Ardissono, Kuflik, & Petrelli, 2012; Kabassi, 2017), while some studies focused on methodologies for the comparative evaluation of different design alternatives related to the user interaction with virtual museum systems (Loris, Fabio, & Maurizio, 2017).

Personal Virtual Museums

Designing for museum personalisation is a challenging task (Aylett & Cavazza, 2001). Museum visitors are unique personalities with specific goals and preferences. Any profile initialisation method has to take into account multiple user factors to be truly effective and usable in a museum context (Kiourt, Pavlidis, Koutsoudis, & Kalles, 2017b). It also needs to be based on an appealing and non-intrusive design, and to be brief and interesting. Nowadays virtual museums focus on personalised content for achieving a

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

most effective user experience, which is why the term Personalised Virtual Museums (PVM) has been coined (Hayashi, Bachelder, & Nakajima, 2016). Some of the key features of PVMs are (Kiourt, Pavlidis, Koutsoudis, & Kalles, 2017b):

- PVMs are dynamical systems supporting interpersonal communication as well as knowledge modelling and management;
- PVMs explore a pan-institutional, modular system of learning objects for the creation of new authored content;
- PVMs exploit a structured, external linking system (linked data technologies) to provide a richer context for item level content;
- PVMs provide "rich content"—rich media that maximize the use of images, video, audio and computational media as much as possible, along with comprehensive context for individual items;
- PVMs provide dynamic educational personalised content to users based on their requirements.

These key features in PVMs are important for the development of dynamical, pleasant, user-friendly, attractive, personalised educational virtual worlds for various uses, such as education, work simulation, etc. Studies based on technologies used in PVMs suggest that the most appropriate content for the visitors is delivered based on some specific features: user profiles, the cumulative visitors' history, user-generated content and on combinations of these features (Yiannoutsou & Avouris, 2014). In order for personalised information and content to visitors to be created, a significant amount of digitised information/content is required from cultural heritage resources; those data must be filtered and personalised in order to enable the individual user to easily access and consume it in an engaging manner. This approach focuses on visits adjusted to visitor interests, moods, visit purposes and previous experience. Thus, personalised interaction is an important tool and method to bring visitors closer to the museum employs tools and methodologies to accommodate the visitors' interests.

Nowadays, cultural heritage institutions employ serious gaming to support distance and lifelong learning techniques to prepare the visitors for the visit or to enhance the experience after the visit (Anderson, et al., 2010; Doulamis, Liarokapis, Petridis, & Miaoulis, 2012; Paliokas & Sylaiou, 2016). Serious games mainly appear in museum web-sites and users have the ability to play the game whenever they want, as well as to relive the experience unlimited times. Those games are created for supporting cultural experience by engaging users with the cultural content in playful ways. For this reason, gamification methods and techniques are an important ingredient in the development of serious games for cultural applications, in a way so that the visitors stay focused on the purpose of the visit and not on a free-time-spending activity (Anderson, et al., 2010). In many cases, the user engagement with cultural content remains at a superficial level as it is not smoothly integrated in the story or the game (Yiannoutsou & Avouris, 2012).

ENCHANCED VIRTUAL REALITY EXPERIENCE

In this section, we introduce three core sub-domains for developing enhanced virtual reality user experiences in cultural heritage over on a specific main domain, the user modelling. Researchers in the domain of virtual environments focus on several sub-domains, including dynamical knowledge modelling, management and presentation and dynamical content generation, player modelling, and gameplay (Korn & Lee, 2017; Kiourt, Pavlidis, Koutsoudis, & Kalles, 2017a; Kiourt, Pavlidis, Koutsoudis, & Kalles,

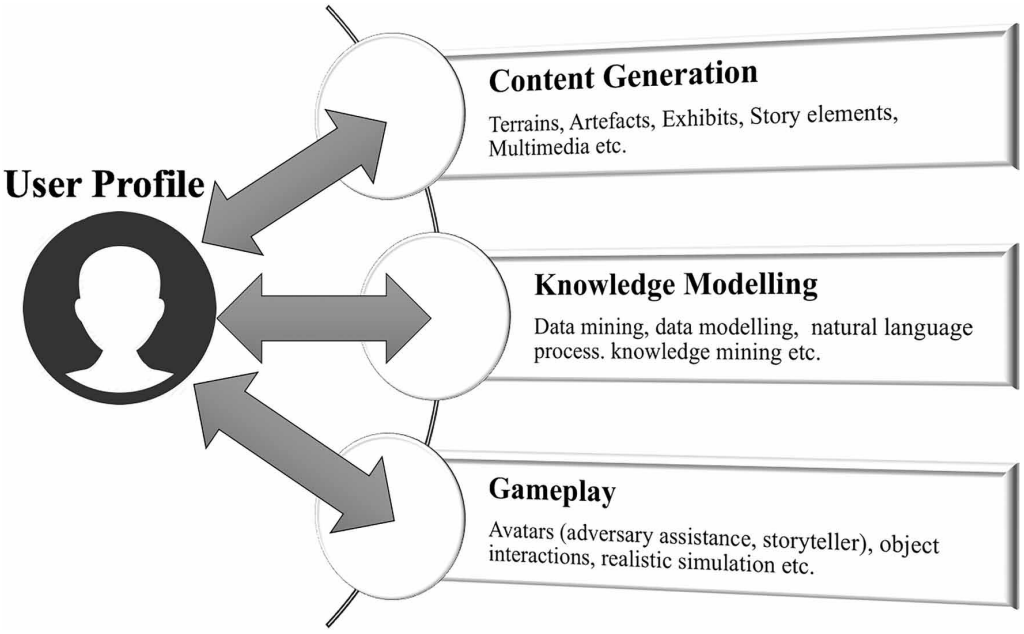
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2017b). Fusing those aspects may result in three different sub-domains: *content, knowledge modelling, and gameplay*, which feed on the user modelling. In the context of a cultural visit, user modelling is the process that models a user profile, including interest, mood, visit purpose, knowledge and other personal characteristics, as well as contextual aspects, in order to select the most suitable content to be delivered in the most well-adapted manner (Bohnert, 2010; Ardissono, Kuflik, & Petrelli, 2012; Hampson, Bailey, Munnelly, Lawless, & Conlan, 2013; Scott, Soria, & Campo, 2017). The three core sub-domains for developing enhanced virtual reality user experiences based on user-modelling are:

- *Content* (C), which is any form of digital data and media in any local or remote repository and any form of digital data and media being dynamically created;
- *Knowledge Modelling* (KM), which is the procedure of creating a computer interpretable model of knowledge, or standard specifications about a kind of process and/or a kind of facility or product;
- *Gameplay* (Gp), which in many, probably in most, virtual environments include methods and tools for various non-player characters as well as the interactivity of various objects of the environment.

These core sub-domains, whose main parameterisation is based on the information use modelling (Figure 1) can serve as the general framework for developing enhanced personalised virtual environments for cultural heritage purposes. The framework has the potential to provide a more targeted and well-adapted (personalised) educational content in virtual museums.

Figure 1. The proposed PVM framework with the three core sub-domains



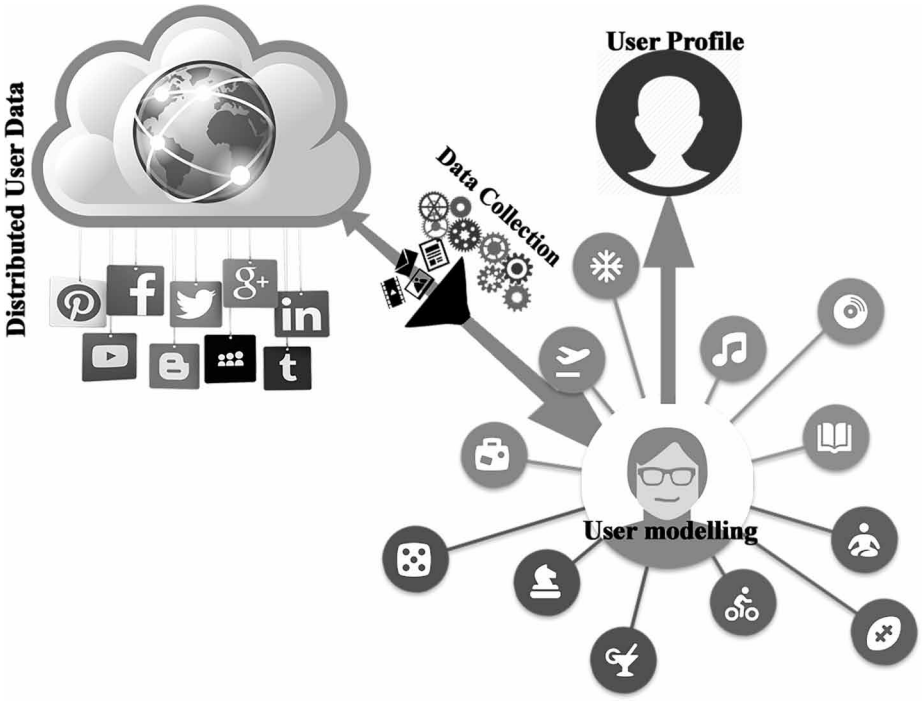
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

User Modelling

Personalisation of cultural heritage content, require methods that are able to model the user interest, knowledge, mood and other personal characteristics, as well as contextual aspects, to select the most appropriate content, and to deliver it in the most suitable way (Fischer, 2001; Bohnert, 2010; Hampson, Bailey, Munnelly, Lawless, & Conlan, 2013). This research domain empowering this approach is user modelling. Since current trends focus on visitor needs and engagement user modelling targets the provision of differentiated access to information and services tailored to a visitor's specific profile, both in the physical museum and virtual museums (Bohnert, 2010; Ardissono, Kuflik, & Petrelli, Personalization in cultural heritage: the road travelled and the one ahead, 2012; Hampson, Bailey, Munnelly, Lawless, & Conlan, 2013; Ardissono, et al., 2015). Based on related works, visitors' interests and goals depend on numerous factors, such as the cultural background, previous experience and social contexts (Falk, 2009). This is because viewing personally interesting content encourages visitors' engagement and learning (Tallon & Walker, 2008; Kiourt, Pavlidis, Koutsoudis, & Kalles, 2017b). Nowadays, research in this field focuses on the development of systems that autonomously retrieve users' data from the web, as well as to interact with the users in order to get more targeted information, and to develop profiles.

The exploration and exploitation of social media resources, as well as the modelling of the user profile based on those data is graphically depicted in Figure 2. The collected data are first scanned and analysed and then the model of the user is being built by exploiting methods of targeted user profile generation that draws on user-computer interactions for more accurate and personal information (Fischer, 2001). Through the general model of the user a user profile can be built based on the adaptation specified by applications.

Figure 2. User modelling based on social media resources



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dynamical Content Generation

Digital content is any content that exists in the form of digital data and presents some information to users (Kuflik, Lo Bue, & Stock, 2017). Cultural heritage content consists of rich media and research focuses on their automated creation, generation or production. *Dynamical Content Generation* (Shaker, Togelius, & Nelson, 2016; Korn & Lee, 2017; Kuflik, Lo Bue, & Stock, 2017) refers to the generation of content for a virtual environment, autonomously or with limited human involvement. The goal of the dynamical content generation methods is to remove the need for humans to generate content in large scale personalised virtual environments. In serious games, dynamical content generation is used to maximize the learning effects of a game, or perhaps the addictiveness of a "casual" game by focusing on educational purposes. Terrains, characters, artefacts, exhibits, story elements, game rules, textures, music, multimedia etc. are the main objects that can be generated dynamically. There are basically two methods for dynamical content generation, offline and online, the first of which focuses on content generation during the development of the game, whereas the second focuses on the real-time content generation.

Knowledge Modelling

Knowledge Modelling is the procedure of creating a model of knowledge or standard specifications about a kind of process and/or about a kind of facility or product. Nowadays, this subdomain focuses on the *computer interpretable knowledge modelling*, for the development of systems that intelligently deliver personalised content to users. Knowledge is a combination of information and experience, context, interpretation and reflection (Davenport, Long, & Beers, 1998). Knowledge modelling was established several years ago as a research field, and started with traditional information systems and databases, to distributed systems, and the Web (Devedžić, 1999). Today knowledge modelling aims to exploit the distributed data on the Web. A well-known source of knowledge modelling is crowd-sourcing, which is of special interest for cultural heritage applications, like virtual museums. For example, Europeana[1] is a digital platform for cultural heritage, with content based on linked open data technologies, which are offered through an application programming interface (API). Several studies have already shown the importance of crowd-sourcing and knowledge modelling in cultural heritage (Uden & Zipf, 2012; Ciolfi, 2013; Biella, et al., 2015).

Gameplay

Gameplay, in many, probably most, virtual environments includes methods and tools for various non-player characters (NPCs). These characters are controlled by agent controlling algorithms (Wooldridge & Jennings, 1995; Shoham & Leyton-Brown, 2009). The requirements of these algorithms have led to an emphasis on behaviour authoring methods such as finite state machines and behavioural trees, as AI methods based on search, optimisation and learning have not shown advantages in playing to gain experience. A common disadvantage has been a perceived lack of predictability and control with such methods. Some of the most known Artificial Intelligence methods (Russell & Norvig, 2010) used in playing games include tree search-based approaches (Turing, Bates, Bowden, & Strachey, 1953), reinforcement learning (Sutton & Barto, 1998) and supervised learning (Bishop, 2006). The integration of NPCs in virtual environments focuses on enhancing the user experience in dynamic environments with realistic, non-boring content and concepts. AI in serious games is generally useful for modelling NPC behaviour and playing the game not for winning, but rather for the experience of play.

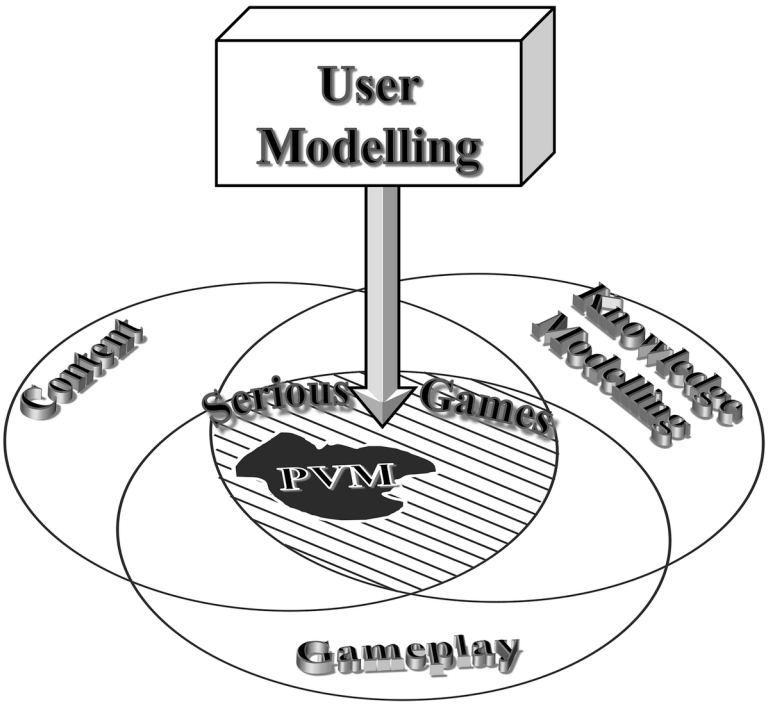
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

TOWARDS A NEW UNIFYING FRAMEWORK

Designing for museum personalisation is a challenging task. Museum visitors are unique personalities with specific goals and preferences. Any profile initialisation method has to take into account multiple user factors to be truly effective and usable in a museum context. It needs also to be based on an appealing and non-intrusive design, and to be brief and interesting. This section serves as a kick-off discussion on a generalised efficient and effective framework for enhancing the user experience in virtual reality application in cultural heritage and specifically in PVMs.

By following the analysis in the previous sections, we aimed at summarizing some of the most important key factors relating to PVMs that focus on dynamical environments composed of external (linked open) resources based on user profiles. Education, training and entertainment seem to be interwoven with PVMs. Also, we aimed at highlighting the importance of social media, crowd-sourcing, data modelling, data mining technologies towards more engaging virtual environments. In addition, we introduced the three core sub-domains, – content generation, knowledge modelling and gameplay – which can serve as the general framework for developing enhanced virtual environments by exploiting the user profile (Figure 3). The user model provides the user profile, based on the needs of the user and the requirements and specifications of the applications, and is composed of information derived from social media sources as well as from a constant exchange of information between the user and the applications. Thus, designed systems have the potential to provide more accurately targeted educational content in cultural heritage applications.

Figure 3. The proposed technological framework for the next generation Personalised Virtual Museums



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The intersection of the three sub-domains can be considered as a representation of the domain of serious games:

$$C \cap KM \cap Gp = SG$$

where *SG* consists of *n* serious games for different purposes, such as entertainment, cultural heritage, education, artificial intelligence, sociology, military and health systems (Breuer & Bente, 2010; Kiourt, Pavlidis, Koutsoudis, & Kalles, 2017b), whereas the PVMs can be considered or envisioned within the framework proposed in this paper to be a subset of the set of serious games *SG*:

$$PVM \subset SG$$

This general framework can be used as a best practice guide for the development of the next generation PVMs, focusing on more accurately personalised content delivery to users by focusing on the development of dynamical, user-friendly, attractive, engaging virtual worlds for various uses, most importantly for education, culture, tourism, simulation and entertainment. What is also interesting is that, following Wecker's (2014) line of reasoning, this framework could be also applied and extended to all settings, prior or after an on-site visit, for onetime and/or multiple times experiences and, most importantly, for applications outside the museum site context, empowering the connection of cultural heritage with everyday life.

User-Centred Virtual World Case Studies

The introduced framework was tested in the DynaMus[2] platform (Kiourt, Koutsoudis, & Pavlidis, 2016). DynaMus is an innovative fully dynamical Web-based virtual museum framework that relies entirely on users' creativity and on the exploitation of the rich content in distributed Web resources. DynaMus is able to connect to popular repositories, such as Europeana and Google, and retrieve content that can be used in creating virtual exhibitions. Also, provides a complete authoring interface, in which anyone can easily create customised virtual exhibitions.

The framework as developed in DynaMus easily connects to educational settings, and two such case studies are presented in the following paragraphs. The overall concept in these case studies is, by using the proposed framework, to bridge culture and education through technology and to provide means for tutors to enhance their everyday teaching with alternative approaches, such as personalised education/learning, that derive and reuse cultural content as an integral part of their courses. The case studies involved the development of two different virtual worlds in the DynaMus platform, the *Synthesis Museum*[3] (Kiourt, Koutsoudis, Markantonatou, & Pavlidis, 2016) and the *Pentapolis High School Museum*[4]. To meet the needs of these case studies, the virtual museum framework (DynaMus) has been customised to display virtual worlds that are basically digital replicas of real environments, actual physical buildings. Synthesis museums is a replica of the building of the B & M Theocharakis Foundation for the Fine Arts and Music located in Athens, Greece. Screenshots of both the interior and the exterior of the 3D building model are depicted in Figure 4. The Pentapolis High School Museum is a replica of a historical school building located in a small town called Pentapolis (near the city of Serres), North-eastern Greece. Screenshots of both the interior and the exterior of the 3D building model are depicted in Figure 5.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Synthesis Museums; (left) external view with immediate surroundings; (right) internal view with the exhibition creation mode enabled



Figure 5. Pentapolis High School Museum; (left) external view with immediate surroundings; (right) internal view with the exhibition creation mode enabled



In these case studies, high school students of various ages were involved in creating their own exhibitions, basically of 2D content (paintings/pictures/photos), using any scenario that suited their own views and opinions based on specific subjects proposed by their tutors. Then the tutors, analysed the exhibitions of the students in order to better adapt their course based on the proposed subject presentation (as implied by the exhibitions) as well as on the information from the student profiles. The main idea was that the tutors customize their course based on the target group. After that, the tutors developed ideal personalised interactive exhibitions, which were visited by the students, as guided by the tutor. After the course, all students were free to visit their classmates' exhibitions to broaden their views/opinions on the subject that was under study.

Through the activities, in which many students in both virtual worlds have been involved, they learnt to respect art and culture and were exposed to and had real practice on how to preserve and disseminate it. Apparently, one may recognize a number of pedagogical approaches based on the personalised virtual world framework (proposed in this framework) at work in the project:

- Cultural education and integration of culture in learning;
- Extensive team work and collaboration;
- Motivation with participation in major events and competitions;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Exploitation of gamification factors and the appeal of gaming environments to children;
- Testing of ideas in real-life environments, and more.

An evaluation based on a questionnaire showed that although these systems were primarily designed as dynamical virtual museums there is much acceptance in the personalised educational practice. Learners under the guidance of their tutors could use this system as a tool to create thematic exhibitions and projects regarding virtually any subject of interest. The evaluation results are presented in the following paragraphs through the analysis of four graphs.

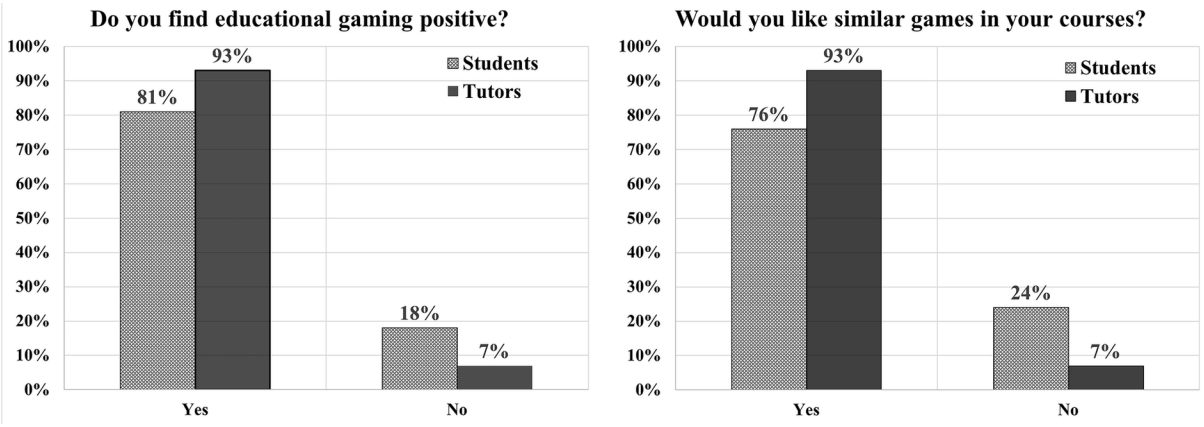
The bars in Figure 6-left indicates that the gamified education is accepted from both tutors and students. It is really interesting to note that the students' acceptance of games in education is lower than that of the tutors. This might be an evidence of a stereotype, in which gaming is connected purely with entertainment and is totally disconnected with education (at least in the Greek educational practice where the case studies have taken place). A similar stereotypical pattern can be observed in Figure 6-right, where the need for more, similar gaming experiences in education is being investigated. The tutors' opinion remains exactly the same, whereas one quarter of the students seem to be opposed to games in education.

Figure 7 highlights the acceptance of the tested systems and their matching to courses in a classroom setup. Again, in this case, although the participants' responses are positive towards the acceptance of the gamified personalised education (by both tutors and students), still, the students' positive response is markedly lower than that of the tutors.

CONCLUSION

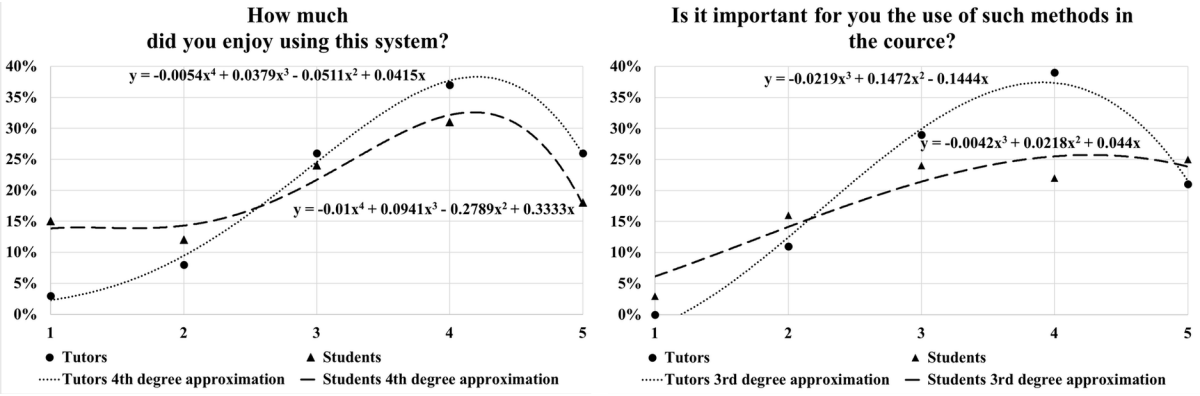
Apparently, the combination of ideas and technologies from virtual environments and personalised human computer interaction in cultural heritage applications is very active field of research, with many different research groups concerned with different aspects (de Aquino & de Souza, 2012; Bohnert, 2010; Ardissono, et al., 2015; Ardissono, Kuflik, & Petrelli, 2012; Kiourt, Pavlidis, Koutsoudis, & Kalles, 2017a). This paper summarizes some of the key factors relating to PVMs that focus on a dynamical environment composed by external resources that is able to provide content based on user profiles. Education, training and entertainment seem to be the natural applications of such PVMs.

Figure 6. Indicative, dichotomous scaled questions-answers for gamified education



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Indicative five-point scaled questions-answers (Likert Scale) with accompanying 3rd and 4th degree polynomial lines to reveal the trend in each case



This paper presented an extensive review of recent methods in the domain and proposed a generalised unifying framework for the development of enhanced virtual environments with personalised content. The inspiration and primary focus in this work is on cultural applications such as virtual museums, exhibitions and storytelling. The general theory for such a framework for PVMs based on three core research sub-domains (content generation, knowledge modelling and gameplay) has been outlined, including parameterisation based on user profiles, and user modelling. Using this framework, a new generation of PVMs is envisioned to emerge. The framework has already been applied in a preliminary setting in the development of the DynaMus platform and two case studies based on this platform were presented with some encouraging evaluation results.

ACKNOWLEDGMENT

The present work was supported by the project "Computational Science and Technologies: Data, Content and Interaction/Technologies for Content Analysis in Culture", MIS code 5002437, co-financed by Greece and European Union in the framework of the Operational Programme "Competitiveness, Entrepreneurship and Innovation" 2014-2020.

REFERENCES

Abt, C. (1987). *Serious games.* University press of America.

Anderson, E., McLoughlin, L., Liarokapis, F., Peters, C., Petridis, P., & de Freitas, S. (2010). Developing serious games for cultural heritage: A state-of-the-art review. *Virtual Reality (Waltham Cross)*, *14*(4), 255–275. doi:10.100710055-010-0177-3

Ardissono, L., Gena, C., Aroya, L., Kuflik, T., Wecker, A., Oomen, J., & Stock, O. (2015). PATCH 2015 - Personalized Access to Cultural Heritage. In *Proceedings of the 20th International Conference on Intelligent User Interfaces*, Atlanta, GA (pp. 447-449).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ardissono, L., Kuflik, T., & Petrelli, D. (2012). Personalization in cultural heritage: The road travelled and the one ahead. *User Modeling and User-Adapted Interaction*, *22*(1-2), 73–99. doi:10.100711257-011-9104-x

Asif, M., & Krogstie, J. (2012). Research Issues in Personalization of Mobile Services. *International Journal of Information Engineering and Electronic Business*, *4*(1). doi:10.5815/ijieeb.2012.04.01

Aylett, R., & Cavazza, M. (2001). Intelligent Virtual Environments-A state-of-the-art report. In *Eurographics Conference*, Manchester, UK.

Biella, D., Sacher, D., Weyers, B., Luther, W. N. B., & Schreck, T. (2015). Crowdsourcing and Knowledge Co-creation in Virtual Museums. In *21st International Conference Collaboration and Technology*, Yerevan, Armenia (pp. 22-25). 10.1007/978-3-319-22747-4_1

Bishop, M. (2006). Pattern recognition. *Machine Learning*, 128.

Bohnert, F. (2010). Personalising the Museum Experience. In Pervasive User Modeling and Personalization (PUMP 2010) (p. 33).

Bonis, B., Stamos, J., Vosinakis, S., Andreou, I., & Panayiotopoulos, T. (2009). A platform for virtual museums with personalized content. *Multimedia Tools and Applications*, *42*(2), 139–159. doi:10.100711042-008-0231-2

Bonis, B., Vosinakis, S., Andreou, I., & Panayiotopoulos, T. (2013). Adaptive virtual exhibitions. *DESIDOC Journal of Library and Information Technology*, *33*(3), 183–198. doi:10.14429/djlit.33.4604

Breuer, J., & Bente, G. (2010). Why so serious? On the relation of serious games and learning. *Journal for Computer Game Culture*, *4*(1), 7–24.

Brown, S., & Vaughan, C. (2010). *Play: How it changes the brain, Opens the Imagination, and Invigorates the Soul*. J P Tarcher/Penguin Putnam.

Christopoulos, D., Mavridis, P., Andreadis, A., & Karigiannis, J. N. (2011). Using Virtual Environments to Tell the Story: "The Battle of Thermopylae." In *Third International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)* (pp. 84-91). 10.1109/VS-GAMES.2011.18

Ciolfi, L. (2013). The Collaborative Work of Heritage: Open Challenges for CSCW. In *ECSCW 2013: Proceedings of the 13th European Conference on Computer Supported Cooperative Work*, Paphos, Cyprus (pp. 83-101). 10.1007/978-1-4471-5346-7_5

Coenen, T., Mostmans, L., & Naessens, K. (2013). MuseUs: Case study of a pervasive cultural heritage serious game. *Journal on Computing and Cultural Heritage, 6*(2).

Davenport, H., Long, D., & Beers, M. (1998). Successful knowledge management projects. *Sloan Management Review*, Winter, 43–58.

David, R., & Sandra, L. (2005). *Serious Games: Games that Educate, Train, and Inform*. Muska & Lipman/Premier-Trade.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

de Aquino, M. S., & de Souza, F. d. (2012). Adaptive Virtual Environments: The Role of Intelligent Agents. In Practical Applications of Agent-Based Technology. INTECH Open Science, 87-110.

Deterding, S., Dixon, D., & Khaled, R., & L., N. (2011). From game design elements to gamefulness: defining "gamification." In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, New York, NY (pp. 9-15). doi:10.1145/2181037.2181040

Devedžić, V. (1999). A survey of modern knowledge modeling techniques. *Expert Systems with Applications*, *17*(4), 275–294. doi:10.1016/S0957-4174(99)00040-8

Doulamis, A., Liarokapis, F., Petridis, P., & Miaoulis, G. (2012). Serious games for cultural applications. In *Intelligent Computer Graphics* (pp. 97-115).

Falk, J. (2009). *Identity and the Museum Visitor Experience*. Left Coast Press.

Fink, J., Kobsa, A., & Nill, A. (1998). Adaptable and adaptive information provision for all users, including disabled and elderly people. *New Review of Hypermedia and Multimedia*, *4*(1), 63–188. doi:10.1080/13614569808914700

Fischer, G. (2001). User Modeling in Human–Computer Interaction. *User Modeling and User-Adapted Interaction*, *11*(1-2), 65–86. doi:10.1023/A:1011145532042

Frischer, B. (2008). *The Rome reborn project, How technology is helping us to study history. OpEd*. University of Virginia.

Froschauer, J., Arends, M., Goldfarb, D., & Merkl, D. (2012). A serious heritage game for art history: Design and evaluation of ThIATRO. In *18th International Conference on Virtual Systems & Multimedia (VSMM)* (pp. 283-290). 10.1109/VSMM.2012.6365936

Froschauer, J., Merkl, D., Arends, M., & Goldfarb, D. (2013). Art history concepts at play with ThIATRO. *Journal on Computing and Cultural Heritage*, *6*(2), 7. doi:10.1145/2460376.2460378

Hampson, C., Bailey, E., Munnelly, G., Lawless, S., & Conlan, O. (2013). Dynamic Personalisation for Digital Cultural Heritage Collections. In *Proceedings of the 6th International Workshop on Personalized Access to Cultural Heritage*, Rome.

Hayashi, M., Bachelder, S., & Nakajima, M. (2016). Automatic Generation of Personal Virtual Museum. In *International Conference on Cyberworlds (CW)*, Chongqing (pp. 219-222).

Kabassi, K. (2017). Evaluating websites of museums: State of the art. *Journal of Cultural Heritage*, *24*, 184–196. doi:10.1016/j.culher.2016.10.016

Katifori, A., Karvounis, M., Kourtis, V., Kyriakidi, M., Roussou, M., Tsangaris, M., & Pujol, L. (2014). CHESS: Personalized Storytelling Experiences in Museums. In A. Mitchell, C. Fernández-Vara, & D. Thue (Eds.), *Interactive Storytelling ICIDS 2014,* LNCS (Vol. *8832,* pp. 232–235).

Kiourt, C., Koutsoudis, A., Markantonatou, S., & Pavlidis, G. (2016). The 'Synthesis' Virtual Museum. *Mediterranean Archaeology and Archaeometry*, *16*(5).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kiourt, C., Koutsoudis, A., & Pavlidis, G. (2016). DynaMus: A fully dynamic 3D virtual museum framework. *Journal of Cultural Heritage*, *22*, 984–991. doi:10.1016/j.culher.2016.06.007

Kiourt, C., Pavlidis, G., Koutsoudis, A., & Kalles, D. (2017a). Realistic Simulation of Cultural Heritage. *International Journal of Computational Methods in Heritage Science*, *1*(1), 10–40. doi:10.4018/IJCMHS.2017010102

Kiourt, C., Pavlidis, G., Koutsoudis, A., & Kalles, D. (2017b). Multi-Agents Based Virtual Environments for Cultural Heritage. *IEEE 26th International Conference on Information, Communication and Automation Technologies.* Sarajevo, Boznia & Herzegovina.

Kontogianni, G., & Georgopoulos, A. (2015). A realistic Gamification attempt for the Ancient Agora of Athens. *In Digital International Congress 2015, 1*, Granada Spain (pp. 377-380). 10.1109/Digital-Heritage.2015.7413907

Kontogianni, G., Koutsaftis, C., Skamantzari, M., Georgopoulos, A., & Chrysanthopoulou, C. (2016). Developing and Exploiting 3D Textured Models for a Serious Game Application. In *8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*, Barcelona Spain. 10.1109/VS-GAMES.2016.7590344

Korn, O., & Lee, N. (2017). *Game Dynamics: Best Practices in Procedural and Dynamic Game Content Generation*. Springer. doi:10.1007/978-3-319-53088-8

Koutsoudis, A., Makarona, C., & Pavlidis, G. (2012). Content-based navigation within virtual museums. *Journal of Advanced Computing Science and Technology*, *1*(2), 73–81. doi:10.14419/jacst.v1i2.135

Kuflik, R., Lo Bue, A., & Stock, O. W. (2017). When will Cultural Heritage Content Creation Get to the Digital Age? In PATCH: Personalized Access to Cultural Heritage (pp. 335-336).

Kuflik, T., Kay, J., & Kummerfeld, B. (2010). Lifelong Personalized Museum Experiences. In Proc. of Pervasive User Modeling and Personalization.

Lepouras, G., & Vassilakis, C. (2005). Virtual museums for all: Employing game technology for edutainment. *Virtual Reality (Waltham Cross)*, *8*(2), 96–106. doi:10.100710055-004-0141-1

Liarokapis, F., Sylaiou, S., & Mountain, D. (2008). Personalizing Virtual and Augmented Reality for Cultural Heritage Indoor and Outdoor Experiences. In *9th International Symposium on Virtual Reality, Archaeology and Cultural Heritage*, Braga, Portugal (pp. 55-62).

Loris, B., Fabio, B., & Maurizio, M. (2017). Virtual museum system evaluation through user studies. *Journal of Cultural Heritage*, *26*, 101–108. doi:10.1016/j.culher.2017.02.005

Mortara, M., Catalano, C., Bellotti, F., Fiucci, G., Houry-Panchetti, M., & Petridis, P. (2014). Learning cultural heritage by serious games. *Journal of Cultural Heritage*, *15*(3), 318–325. doi:10.1016/j.culher.2013.04.004

Nicholson, S. (2011). Making the Gameplay Matter: Designing Modern Educational Tabletop Games. *Knowledge Quest*, *40*(1), 60–65.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Not, E., & Petrelli, D. (2014). Balancing Adaptivity and Customisation: In Search of Sustainable Personalisation in Cultural Heritage. In User Modeling, Adaptation, and Personalization (pp. 405-410).

Oberlander, J., Karakatsiotis, G., Isard, A., & Androutsopoulos, A. (2008). Building an adaptive museum gallery in Second Life. In *Proceedings of Museums and the Web 2008,* Montreal, Canada.

Paliokas, I., & Sylaiou, S. (2016). The Use of Serious Games in Museum Visits and Exhibitions: A Systematic Mapping Study. In *8th International Conference on Games and Virtual Worlds for Serious Applications*, Barcelona, Spain. 10.1109/VS-GAMES.2016.7590371

Patel, M., White, M., Walczak, K., & Sayd, P. (2003). Digitisation to presentation: building virtual museum exhibitions. *Vision, Video and Graphics*.

Paternò, F., & Mancini, C. (1999). Designing Web User Interfaces Adaptable to Different Types of Use. In *Museums and the Web 1999: Proceedings of the International Conference for Culture and Heritage*. Pittsburgh, PA.

Petrelli, D., & Not, E. (2005). User-Centred Design of Flexible Hypermedia for a Mobile Guide: Reflections on the HyperAudio Experience. *User Modeling and User-Adapted Interaction*, *15*(3-4), 303–338. doi:10.100711257-005-8816-1

Pujol, L., Roussou, M., Poulou, S., Balet, O., Vayanou, M., & Ioannidis, Y. (2012). Personalizing interactive digital storytelling in archaeological museums: the CHESS project. In *40th annual conference of computer applications and quantitative methods in archaeology.* Amsterdam University Press.

Ruotsalo, T., Mälekä, E., Kauppinen, T., Hyvönen, E., Haav, K., Rantale, V., ... Matskin, M. (2009). Smartmuseum: Personalized Context-aware Access to Digital Cultural Heritage. In *Third International Conference Proceedings of International Conference on Digital Libraries and the Semantic Web 2009 (ICSD2009),* Trento, Italy.

Russell, S., & Norvig, P. (2010). *Artifcial Intelligence. A modern approach*. Prentice hall.

Schmidt-Belz, B., Laamanen, H., Poslad, S., & Zipf, A. (2003). Location-based mobile tourist services - first user experiences. In *Information Technology and Travel & Tourism. Proceedings of the tenth International Conference*, Helsinki, Finland (pp. 115-123).

Scott, E., Soria, A., & Campo, M. (2017). Adaptive 3D Virtual Learning Environments-A Review of the Literature. *IEEE Transactions on Learning Technologies*, *10*(3), 262–276. doi:10.1109/TLT.2016.2609910

Shaker, N., Togelius, J., & Nelson, J. (2016). *Procedural Content Generation in Games: A Textbook and an Overview of Current Research*. Springer. doi:10.1007/978-3-319-42716-4

Shoham, Y., & Leyton-Brown, K. (2009). *MULTIAGENT SYSTEMS Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Pres.

Sookhanaphibarn, K., & Thawonmas, R. (2009). A content management system for userdriven museums in second life. In *International Conference on CyberWorlds*, Bradford, UK (pp. 185-189).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Stock, O., Carenini, G., Cecconi, F., Franconi, E., Lavelli, A., Magnini, B., . . . Strapparava, C. (1993). Alfresco: enjoying the combination of NLP and hypermedia for information exploration. In M.T. Maybury (Ed.), Intelligent Multimedia Interfaces (pp. 197-224).

Stock, O., Zancanaro, M., Busetta, P., Callaway, C., Krüger, A., Kruppa, M., ... Rocchi, C. (2007). Adaptive, Intelligent Presentation of Information for the Museum Visitor in PEACH. *User Modeling and User-Adapted Interaction*, *17*(3), 257–304. doi:10.100711257-007-9029-6

Sutton, S., & Barto, A. (1998). *Reinforcement Learning: An Introduction*. MIT Press.

Sylaiou, S., Liarokapis, F., & Kotsakis, K. P. (2009). Virtual museums, a survey and some issues for consideration. *Journal of Cultural Heritage*, *10*(4), 520–528. doi:10.1016/j.culher.2009.03.003

Tallon, L., & Walker, K. (2008). *Digital Technologies and the Museum Experience: Handheld Guides and Other Media*. AltaMira Press.

Troche, J., & Jacobson, J. (n.d.). An exemplar of Ptolemaic Egyptian temples. *Computer Applications in Archaeology (CAA),* Granada, Spain.

Turing, M., Bates, M., Bowden, B., & Strachey, C. (1953). Digital computers applied to games. *Faster than thought, 101*.

Uden, M., & Zipf, A. (2012). OpenBuildingModels: Towards a Platform for Crowdsourcing Virtual 3D Cities. *7th 3D GeoInfo Conference.* Quebec City, QC, Canada.

Vayanou, M., Karvounis, M., Kyriakidi, M., Katifori, A., Manola, N., Roussou, M., & Ioannidis, Y. (2012). Towards personalized storytelling for museum visits. In *6th International Workshop on Personalized Access, Profile Management, and Context Awareness in Databases*, Istanbul, Turkey.

Wang, Y., Aroyo, L., Stash, N., & Rutledge, L. (2007). nteractive User Modeling for Personalized Access to Museum Collections: The Rijksmuseum Case Study. In C. Conati, K. McCoy, & G. Paliouras (Eds.), User Modeling 2007. UM 2007, LNCS (Vol. 4511, pp. 385-389). Berlin, Heidelberg.

Wecker, A. J. (2014). Personalized Cultural Heritage Experience Outside the Museum: Connecting the Museum Experience to the Outside World. In *22nd International Conference on User Modeling, Adaptation, and Personalization* (pp. 496-501). Aalborg, Denmark: Springer. 10.1007/978-3-319-08786-3_46

White, M., Mourkoussis, N., Darcy, J., Petridis, P., Liarokapis, F., Lister, P., & Gaspard, F. (2004). *ARCO – an architecture for digitization, management and presentation of virtual exhibitions. In* Proceedings Computer Graphics International, Crete (pp. 622–625). doi:10.1109/CGI.2004.1309277

Wooldridge, M., & Jennings, R. (1995). Intelligent agents: Theory and practice. *The Knowledge Engineering Review*, *10*(2), 115–152. doi:10.1017/S0269888900008122

Yiannoutsou, N., & Avouris, N. (2012). Mobile games in Museums: From learning through game play to learning through game design. *ICOM Education*, *23*, 79-86.

Yiannoutsou, N., & Avouris, N. (2014). Game Design as a context for Learning in Cultural Institutions. In C. Karagiannidis, P. Politis, & I. Karasavvidis (Eds.), *Research on e-Learning and ICT in Education* (pp. 165–177). New York, NY: Springer. doi:10.1007/978-1-4614-6501-0_12

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ENDNOTES

[1] https://www.europeana.eu/
[2] http://dynamus.ceti.gr
[3] http://thf.ceti.gr
[4] http://pentapolis.ceti.gr

This research was previously published in the International Journal of Computational Methods in Heritage Science (IJCMHS), 2(1); edited by George P. Pavlidis; pages 23-39; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 66
High Performance Computing of Possible Minds

Soenke Ziesche
Maldives National University, Maldives

Roman V. Yampolskiy
University of Louisville, USA

ABSTRACT

It has been shown that the space of possible minds is vast, actually infinite. Intellectology is a new field of study, which examines in more detail features of possible minds. Among the many open and unexplored questions in this field is the following: "Which activities can minds perform during their lifetime?" This question is very broad, thus our contribution here addresses the sub-question "Which non-boring activities can minds perform?" This issue is ethically relevant for human minds if the predicted significant extension of our lifetime materializes and we are then potentially challenged how to spend this additional time. The space of potential non-boring activities has been called "fun space." We analyze the relation between various types of minds and the portion of the fun space, which is accessible for them. As a novel result, we demonstrate that human minds can experience two types of fun when transforming information to knowledge: novelty fun and process fun.

INTRODUCTION

We motivate the relevance of the question "Which non-boring activities can minds perform?" as follows: There is optimism for a significant extension of the lifetime of humans in the near and medium future. Yet, some people expressed concern that there is not enough accessible fun space, thus much longer lives would not be fulfilling and mostly boring, as examined by Walker (Walker, 2006). Therefore, the subject of this article, to analyze the accessible fun space for various minds, is critical in order to conclude whether an extended lifespan is desirable.

DOI: 10.4018/978-1-7998-0951-7.ch066

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In addition, future scenarios can be envisaged when humans are enhanced by intelligence amplification and other transhuman features, potentially supported by friendly AI. Further scenarios include transition to substrate-autonomous persons, e.g. through uploading or whole brain emulation (Koene, 2012; Bostrom, 2014). In the latter case, lifetime could increase almost indefinitely, and also former unknown sensations could be perceived. Also for this option, concerns have been raised that it will turn out to be boring for such enhanced humans due to the lack of remaining intellectual challenges. In contrast, others believe the fun space is large enough for either scenario (e.g. Moravec, 1988; Yudkowsky, 2015; Bostrom, 2008)

Yet another view by Pearce is that in a parallel development boredom will be "neurochemically impossible" through advanced technology (Pearce, 2012; Yampolskiy, 2015). Therefore, although human minds are only a small subset of the space of possible minds, for this type of minds the question about sufficient non-boring activities, i.e. enough accessible fun space, is very important.

In addition, we are addressing the question of accessible fun space also for non-human minds by aiming to contribute to one aspect of the field of "intellectology" (Yampolskiy, 2015). Yampolskiy has introduced "intellectology" as a new area of study in order to turn away from a human-centric view regarding minds and to examine in more detail features of any possible minds. This universe of possible mind designs is actually vast and certainly merits investigation; especially given the possible scenario that one group of minds, that is minds of AI agents, may cohabit with us in the not so far future.

This paper is structured as follows: In the following section, we provide an introduction to the space of minds and how it has been explored and defined in the literature. This is followed by an introduction to the space of fun, which enables us to further specify the critical questions for our subject. In the following main section, we present our contribution to the field of "intellectology" by analyzing relationships between the space of minds and the space of fun.

RELATED WORK

First, we present related work to the space of minds. When we reason about minds we tend to think of human minds only. This is because of the anthropomorphic bias. However, in addition, there are other minds, which we encounter on earth, the minds of higher order animals, and then there are various minds, which we can imagine as possibility and perhaps even more beyond our imagination ("unknown unknowns"). Several theoretical surveys on this topic exist and it has been shown that the space of possible minds is vast (e.g. Sloman, 1984; Goertzel, 2006; Hall, 2007; Yudkowsky, 2008; Yampolskiy, 2015).

Examples for potential minds could be human-designed AI minds, self-improving minds, a combination of minds constituting itself a mind and many more. There have been several attempts to classify the space of minds (Yampolskiy, 2015). In fact, the space of human minds forms only a tiny subset within the universe of possible minds (Yudkowsky, 2008). The space of possible minds can be considered as the set of possible cognitive algorithms. Based on this and on the limited number of cognitive algorithms, which human minds can potentially perform, it can be concluded that the majority of possible minds is more intelligent than human minds. Yampolskiy and Fox describe this insight as another example of a Copernican Revolution, i.e. a revision of the view that humanity is central, which in this case refers to minds (Yampolskiy & Fox, 2012).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Yampolskiy shows that the set of mind designs is infinite and countable and that all minds can be generated sequentially by a deterministic algorithm, based on a variant of the Levin Search (Yampolskiy, 2015). By linking a mind to a number, Yampolskiy also demonstrates that "… a mind could never be completely destroyed, making minds theoretically immortal. A particular mind may not be embodied at a given time, but the idea of it is always present." (Yampolskiy, 2015). The fact that minds are countable will be useful when establishing relations with the fun space.

We continue by presenting related work to the space of fun. Regarding the goals and the motivation of minds Bostrom distinguishes between two theses, the orthogonality thesis and the instrumental convergence thesis (Bostrom, 2012). The former states that "more or less any level of intelligence could in principle be combined with more or less any final goal", while the latter proposes that there is a rather limited number of instrumental values which many intelligent agents share.

The focus of this paper are those goals of minds that are related to the discovery of novelties and to having fun, which is the opposite of boredom. Yudkowsky proposes a link between novelty and fun: "Novelty appears to be one of the major keys to fun, and for there to exist an infinite amount of fun there must be an infinite amount of novelty." (Yudkowsky, 2015). We shall specify this further below.

Omohundro defines four drives for AI minds: Efficiency, self-preservation, acquisition, and creativity (Omohundro, 2007, Omohundro, 2008). Three of these drives address novelty: The creativity drive is about novelty by definition (leads to the development of new concepts, algorithms, theorems, devices, and processes), but so are the efficiency drive (aiming for novel algorithms and methodologies) and the acquisition drive (aiming for novel resources). Only the self-preservation drive is linked more indirectly to novelty, which has to be relative to a particular self. Hence, without self-preservation everything might be seen as novel to some new version of that self. However, Omohundro does not mention "fun" to be linked to any of these AI drives.

We know that at least human minds aim for accessing fun space. Therefore, the subset of all minds, with the distinction to have the objective of having fun, is at least the set of all human minds, but potentially it is bigger. Yet, when pondering over such questions it is always critical to beware of the anthropomorphic bias:

- It could well be possible that fun is specific to humans only and other minds do not aim for fun. The actions of these minds are still likely to be steered by the above drives, yet they may not aim for fun as an epiphenomenon and may be satisfied by pursuing boring activities as long as these serve their goals (Yet, we seem to observe at least some higher order animals having fun.)
- If there are other minds that aim for fun, this "fun" may look extremely different from what is fun for humans and may well be unimaginable for humans. Both, the sensation that other minds "feel" as fun as well as the activities that trigger this sensation could be very different, e.g. activities that are considered tedious by humans may cause for other minds fun-type sensations as epiphenomenon.
- There may be certain minds, who do aim for fun, but apply "wireheading" methods, which is obtaining the reward through direct stimulation without otherwise required productive behavior. Hence, drugs and/or technologies provide them effortlessly and permanently with what they consider as fun (Yampolskiy, 2014).

Having established this, we elaborate further on the critical questions we aim to address:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

What is the relation between novelty and fun? It seems that fun is useful to motivate exploration, innovation and discovery. Stanley and Lehman show that novelty search is more successful than objective-based search (Lehman & Stanley, 2015). Therefore, striving for fun of human minds could be explained by evolution. Novelties are critical for human development, and fun is, at least sometimes, an epiphenomenon while pursuing novelties or at the moment when discovering novelties. However, it can be easily shown that, at least for human minds, discovering novelties and having fun is not the same, since there are numerous novelties that are boring for humans; or a specific novelty could be fun for one human mind, but boring for others.

Schmidhuber proposes the following options to access fun space, which we will take into account below: "There are at least two ways of having fun: execute a learning algorithm that improves the compression of the already known data … or execute actions that generate more data, then learn to compress/understand the new data better." (Schmidhuber, 2010).

It is speculation whether this also applies to non-human minds. For example, does the mind of a "paperclip maximizer" (Bostrom, 2003) accesses fun space when discovering novelties through the creativity drive or the efficiency drive (e.g. finding time- or energy saving ways to produce paperclips) or the acquisition drive (e.g. finding new resources to be converted into paperclips)?

What is the size of the fun space? Since potential knowledge is infinite also the amount of novelties that could potentially be gained by a mind is infinite, yet limited by time and other resources. But as we stated, not every processing of novelties triggers fun as an epiphenomenon. Given the infinity of the knowledge space it may be assumed that the potential fun space is vast. In this regard Yudkowsky posed critical questions: "How much fun is there in the universe? Will we ever run out of fun? Could we be having more fun?" (Yudkowsky, 2015). We pose an additional question here, which brings us to the central part of this paper, i.e. we attempt to establish relations between the space of minds and the space of fun: Which minds have (in principle) access to which portion of the fun space?

For an illustration of relations between various types of minds, the knowledge space and the potentially accessible fun space see Figure 1.

PROPOSED TECHNIQUE

Initially, we reiterate the definitions we established above for the two spaces:

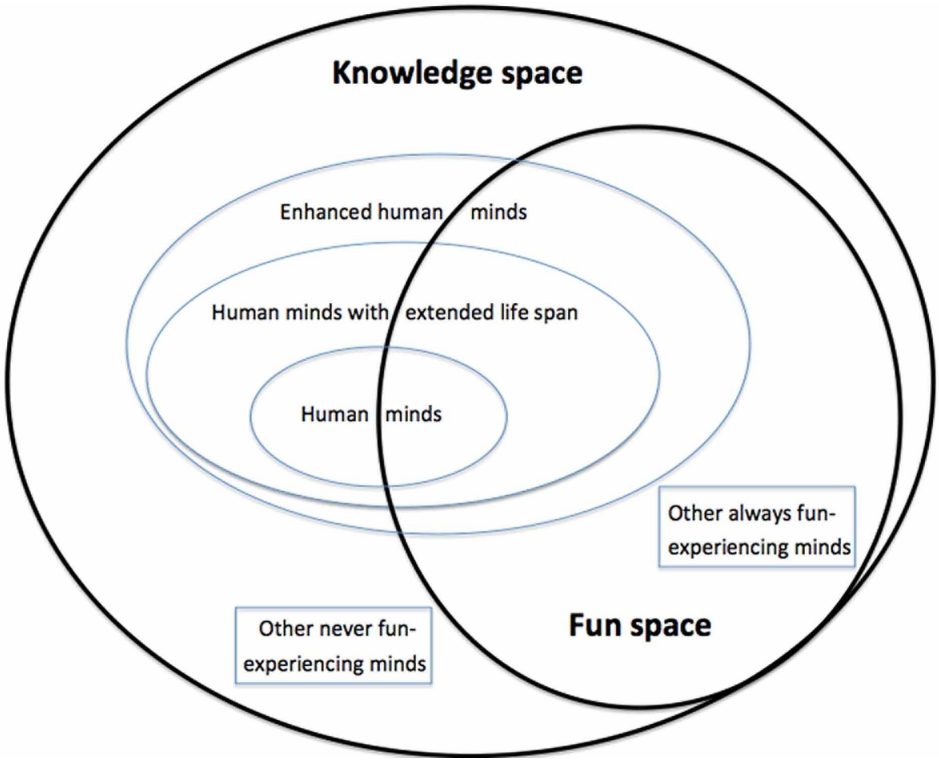
- **Space of Minds**: The set of possible cognitive algorithms (Yudkowsky, 2008).
- **Space of Fun**: The set of non-boring activities (Yudkowsky, 2015).

For the formalization, we assume that both, all potential minds (Yampolskiy, 2015) can be sequentially generated as well as all possible information strings (Yampolskiy, 2013; Schmidhuber, 2009, 2010; Lehman & Stanley, 2015). We also use Yampolskiy's insight that knowledge is discovered by finding and applying an efficient algorithm to an inefficiently stored information string (Yampolskiy, 2013).

The methodology we use to investigate the fun space for certain minds is to define a Boolean function over relevant assertions and determine its truth table, followed by examples for illustration. This Boolean function ("fun") has two parameters, a string representing a mind and another string representing information/knowledge as follows:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Potential access of a selection of minds to knowledge and fun space



fun(mind string, info string) = {true, false}

The verbal description of this function would be that it becomes true, if for a certain mind, represented by a numeric string, the processing of certain information, also represented by a numeric string, triggers fun as an epiphenomenon. For all other combinations of mind strings and information strings the function is false. The formalization is as follows in Table 1:

This means, the access to the knowledge represented by the information string is the precondition for fun. Although the set of human minds is very homogeneous, the extent of knowledge, which humans access during their lifetime varies significantly because certain minds cannot access 1) certain knowledge for intellectual reasons and or 2), even if they are intelligent enough, there could be resource or time restrictions:

Table 1. Formalization

fun(*x, y*) = true, if	[The mind *x* has the ability/intelligence to turn the info string *y* efficiently into knowledge (access).]	AND	[[The mind *x* had never turned that specific info string *y* into knowledge before. (novelty fun)]	OR	[The mind *x* considers the computation process of info string *y* into knowledge as fun, regardless whether it has accessed the info string *y* before or not. (process fun)]]

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- **Available Time and Resources**: Some minds have to devote most of their resources and time for "survival", i.e. have less resources and time for fun. An example would be a rich person in a developed country vs. a poor person in a developing country.
- **Life Span**: Minds are embodied for different durations, i.e. have different life spans, which imposes time restrictions on accessing knowledge.
- **Time of Embodiment**: The accessible knowledge was for a particular mind in Stone Age quite different from one, who is embodied in 21st century.

We have shown that accessible knowledge varies immensely between minds and depends on intelligence, time, and resources to efficiently turn certain information strings into knowledge. Therefore, certain knowledge can neither be part of the process nor the novelty fun space of these minds due to their inability to access it.

After it has been established that a mind *x* can access the knowledge represented by an information string *y*, at least one of two other assertions has to be true, so that fun is involved. Therefore, we define two categories of fun:

Novelty fun: The sensation when new knowledge is ensured by retrieving a new info string *y*.
Process fun: The sensation when new knowledge is ensured merely due to the computation process that is applied to info string *y*.

These two facets of fun are motivated by Schmidhuber's "two ways of having fun" (Schmidhuber, 2009, 2010) and by Yampolskiy's insights about the relation between information strings and knowledge (Yampolskiy, 2013): "Since multiple, in fact infinite, number of semantic pointers could refer to the same string that means that a single string could contain an infinite amount of knowledge if taken in the proper semantic context, generating multiple levels of meaning."

Examples in Table 2 illustrate the different cases. The fun space is marked with grey background:

Table 2. Types of fun

Mind *x* accesses info string *y*	**New info string *y***	**Previously accessed info string *y***
New computation process	Novelty and process fun - To watch for the first time a 3D movie. - To proof that there are infinite prime numbers (Euclid). - To perform qualia surfing, i.e. to experience new qualia (as an uploaded mind, see Loosemore).	Process fun - To find a different proof that there are infinite prime numbers (Euler)[1]. - To listen to a known song again in different context or mood. - To look at a painting again in different context or mood.
Trivial/routine computation process	Novelty fun - To discover a new species of beetle in the rain forest. - To watch a new episode of a cherished TV series. - To do sightseeing in a new city.	- To clean the dishes.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

To elaborate on some examples:

- **Novelty Fun**: To discover a new species of beetle in the rain forest.
  In this table "trivial/routine process" means that no innovative computation process was required to gain the new knowledge. A biologist may have conducted several expeditions to the Amazon rainforest and at a certain day by using her typical (therefore, trivial) methodology and basically her eyes, she discovers a new species of beetle, which, on a very basic level, is accessing a new information string (which may have never been accessed by a human mind before) and for her this is equivalent with having novelty fun.
- **Process Fun**: To look at a painting again in different context or mood.
  It is not uncommon to hear people enjoying looking at the same painting again and again, e.g. revisiting a particular museum regularly for this purpose. Except for the case that they discover new details in the painting, they, on a very basic level, access the same information string and by doing so, even repetitively, they have process fun. The explanation is provided by Yampolskiy: The same information string can contain different knowledge and multiple levels of meaning (Yampolskiy, 2013).
- **Novelty and Process Fun**: To perform qualia surfing as an uploaded mind.
  Loosemore defines "qualia surfing" in the context of uploading as "collecting new experiences by transferring our consciousness back and forth between different substrates" (Loosemore, 2014). One of his several examples is to change "into a body that can swim in the atmosphere of Jupiter … while you stargaze through telescopic eyes that can see beyond the visible spectrum." He explains that even if qualia surfing is restricted to the real world there is vast space of new experiences, which would be even larger if virtual worlds are added. For a person experiencing certain qualia for the first-time novelty as well as process fun may be involved.

A less futuristic example is to watch for the first time a 3D movie. The computation process, via 3D glasses, is new as well the information string, i.e. content of the movie; given the person has not watched it in 2D before.

It has to be highlighted that it is critical that the function fun (mind string, info string) has two parameters. Fun does not only depend on the accessed information string, but very much on the individual mind. For example, there are obviously human minds who neither enjoy watching movies nor discovering new species of beetles etc.

A weakness of human minds and therefore a special case is forgotten knowledge. In contrast, it is very likely that more sophisticated minds will have a nearly perfect memory. When human minds regain forgotten knowledge, both could be possible, to have fun or not: It could be fun to watch a movie again after a long time because one had forgotten the storyline. It may not be fun, but frustrating to spend time re-learning vocabulary of a language, which one had not practiced for some time. In both cases the same information string is accessed by the same computation process when the knowledge had been computed for the first time.

Based on Yampolskiy's approach and other mathematical truths it is evident that there are infinite information strings, which can be turned through infinite computation processes into infinite knowledge (Yampolskiy, 2013). This means every access to new knowledge and every new computation of existing or new knowledge may or may not involve novelty or process fun respectively. Therefore, the questions remain:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

PERFORMANCE EVALUATION AND ANALYSIS

Is the fun space also infinite? If yes, are human minds able to access an infinite subset of this fun space? One constraint was mentioned above: Some knowledge remains hidden from human minds due to insufficient intelligence, time and resources. Those portions of fun space that are linked to the novelty and processing of this inaccessible knowledge are thus also inaccessible for human minds. We look at an excerpt of a random sample truth table for the function fun (mind string, info string), which is infinite and countable since both parameter sets are infinite and countable.

The overall fun space would be infinite if the truth table of this function would contain an infinite number of the value "true". This we cannot prove. It is possible that the number of the value "true" is finite, which implies that in this case the number of the value "false" must be infinite.

For the second question above, we look at specific minds, human minds in particular, and their accessible fun space. Therefore, we added the second column in the above truth table (Table 3), which filters human minds out of the space of all minds. In this example "Mind 2" happens to be human, and this mind would have infinite fun space if that row contained an infinite number of the value "true". The above question whether human minds have access to an infinite subset of the fun space would be affirmative, if the following statement was true:

$$\forall\, x\text{: human_mind}(x) = \mathrm{t} \wedge |\, \{y \mid \mathrm{fun}(x,\, y) = \mathrm{true}\} \,| = \infty$$

We do not have a formal answer to this. In fact, it is almost certain that not all human minds have access to infinite fun space. However, we can observe that the accessible fun space has risen in history and is currently enormous, which is linked to what Kurzweil describes as the "Law of Accelerating Returns" (Kurzweil, 2005). According to this law the rate of change in certain developing systems, for example technologies, has shown to increase exponentially. As already mentioned earlier the access to the fun space of humans who lived hundreds or thousands of years ago was much more restricted than it is for present humans. Only thanks to technologies many novelties can be experienced. And also thanks to technologies we have more time at hand, which does not have to be spent on life-sustaining activities. Therefore, we may assume that future generations can access even larger parts of the fun space.

After focusing so far on current human minds we look in the last part of this section at potential future human minds and also other minds:

Fun for human minds with extended life span. The concern has been expressed that humans with an extended life span will run out of fun (Walker, 2006). Yet, an extended life span enables that one limitation above is alleviated even further (time). Therefore, all else equal (e.g. the intelligence of this

Table 3. Random sample truth table for the function fun

Sample Truth table fun(x,y)	Human mind (t/f)	Info 1	Info 2	Info 3	Info 4	Info 5	Info 6	Info 7	Info 8	Info 9	…
Mind 1	f	f	t	f	t	f	t	t	f	f	
Mind 2	t	f	t	t	f	f	f	t	f	t	
Mind 3	f	f	f	f	t	f	f	t	t	f	
…											

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ageing mind stays the same) this human mind has the potential to gain additional knowledge during the additional time of embodiment, which may entail novelty and/or process fun.

It is hard to speculate how the access to fun space will be for humans who live for thousands or millions of years, but so far life expectancy has been increased successively by smaller time frames only. It is likely that more radical extensions would be accompanied by other developments to the human phenotype, which we discuss below.

Fun for human minds enhanced by intelligence amplification and other trans-human features, potentially supported by friendly AI. Similarly, concern has been raised that humans enhanced by trans-human features and/or supported by friendly AI may run out of fun: "If we self-improve human minds to extreme levels of intelligence, all challenges known today may bore us. Likewise, if superhumanly intelligent machines take care of our every need, it is apparent that no challenges nor fun will remain." (Yudkowsky, 2015).

This argument addresses mainly the above defined "process fun", i.e. previous challenges are becoming too simple. Yet, we have stated not only the infinity of computation processes as well as the infinity of knowledge, but also shown that trivial/routine computation processes could tap fun space as long as a new information string is accessed (of which also an infinite number exists).

Therefore, even if such minds run out of "process fun", since previously challenging computation processes became much simpler and/or can be performed much faster, there should be still sufficient "novelty fun" available. Chances for "novelty fun" are actually much higher than for unenhanced human minds, since enhanced human minds can reach previously inaccessible knowledge due to 1) improved intelligence, 2) much longer life span and 3) more and other resources. Enhanced humans will likely have access to additional sensors, which provide for "novelty fun" through experiencing new qualia (Loosemore, 2014). This is in line with e.g. Bostrom and Yudkowsky, who also believe that enhanced minds will not run out of fun, given that they are embodied in a sufficiently advanced context (Bostrom, 2008; Yudkowsky, 2015). Bostrom is even optimistic regarding "process fun": "If some challenges become too easy for post-humans, they could take on more difficult challenges."

As indicated above, another scenario in the future, which would make the above considerations obsolete, is the manipulation of human and other minds through technology, known as wireheading, in a way that boredom will be "neurochemically impossible" as envisaged by Pearce (Pearce, 2012). For such a manipulated mind even uninterrupted dishwashing would be continuous fun.

Fun for AIs and other minds. We lack evidence to what extent non-human minds can access the fun space, i.e. have a sensation of fun as epiphenomenon of any of their activities. It could be anywhere between the two extremes, which are to never experience fun (*x1*) or to have fun no matter which information string is accessed (*x2*).

$\forall$ y: fun(*x1, y*) = false

$\forall$ y: fun(*x2, y*) = true

At least for those non-human minds that are to some extent steered by the drives proposed by Omohundro (Efficiency, self-preservation, acquisition, and creativity) (Omohundro, 2007, Omohundro, 2008) we may assume that they aim for novelties. On the one hand, it would be anthropomorphism to assume that their pursuit of novelties is accompanied or motivated by any form of fun. On the other hand, if we take our definition of the mind space as the set of all possible cognitive algorithms, i.e. an infinite set,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

there must be other minds, apart from humans, which are able to access fun space. And considering the space of all possible cognitive algorithms, there must be even minds experiencing fun by performing activities, which have nothing to do with detecting novelties.

CONCLUSION

Based on the theoretical groundwork of having described the possible space of minds as well as the space of fun we contributed a relation between the two spaces. We have defined a Boolean function fun (mind string, info string) with two parameters, one numeric string representing a mind and another numeric string representing information/knowledge. This approach is novel, thus a comparison with existing approaches is not possible.

We have shown that human minds could experience two types of fun by transforming information to knowledge: novelty fun and process fun. By means of the above function we were able to formalize the critical question whether human minds are able to access an infinite subset of the fun space. It is beyond our means to give a definite answer to the question, but we provided a substantiated assumption that a large number of human minds in present time and especially in the future have access to a vast space of fun.

These findings have ethical relevance since human minds strive for fun (not merely knowledge). In history as well as still in present time there have been and there are numerous humans living a life predominantly characterized by life-sustaining activities, which do not involve much fun. Yet, ideally the human condition improves over time owing, among other things, to advances in technology. Therefore, in order to determine whether the extension of the life span and/or enhancements for humans really mean progress to the human condition it was critical to show that there will be no shortage of accessible fun space for these minds. If it had turned out that there is not sufficient fun for either human minds with extended life span or enhanced human minds such developments would probably not be desirable.

Moreover, these findings are contributions to the new field of "intellectology" (Yampolskiy, 2015). This field is still in its early stages and provides for large areas of further research. To analyze relations between various kinds of minds and the fun space extends the traditional human-centric space of research and this is what "intellectology" is about. Further research is required on the question how to narrow down or to predict more precisely which new computation processes or which new knowledge mean fun for a certain mind, especially humans, and which are not part of the fun space despite being novel?

REFERENCES

Bostrom, N. (2003). Ethical Issues in Advanced Artificial Intelligence. In I. Smit et al. (Eds.), Cognitive, Emotive and Ethical Aspects of Decision Making in Humans and in Artificial Intelligence (Vol. 2, pp. 12-17).

Bostrom, N. (2008). Why I Want to Be a Posthuman When I Grow Up. In B. Gordijn & R. Chadwick (Eds.), *Medical Enhancement and Posthumanity* (pp. 107–137). Dordrecht: Springer.

Bostrom, N. (2012). The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents. Theory and Philosophy of AI, 22(2), 71-85.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.

Goertzel, B. (2006). *The Hidden Pattern: A Patternist Philosophy of Mind*. Boca Raton, FL: Brown Walker Press.

Hall, J. S. (2007). Kinds of Minds. In A. I. Beyond (Ed.), *Creating the Conscience of the Machine*. Amherst, NY: Prometheus Books.

Koene, R. (2012). Embracing Competive Balance. The Case for Substrate-Independent Minds and Whole Brain Emulation. In A. Eden, J. Moor, J. Soraker, & E. Steinhart (Eds.), *The Singularity Hypothesis: A Scientific and Philosophical Assessment* (pp. 214–265). Berlin: Springer. doi:10.1007/978-3-642-32560-1_12

Kurzweil, R. (2005). *The Singularity Is Near: When Humans Transcend Biology*. New York: Viking.

Loosemore, R. P. W. (2014). Qualia Surfing. In R. Blackford & D. Broderick (Eds.), *Intelligence Unbound: The Future of Uploaded and Machine Minds* (pp. 231–239). Chichester: John Wiley & Sons.

Moravec, H. (1988). *Mind Children: The Future of Robot and Human Intelligence*. Cambridge, MA: Harvard University Press.

Omohundro, S. M. (2007). The Nature of Self-Improving Artificial Intelligence. *Paper presented at Singularity Summit 2007*, San Francisco, CA.

Omohundro, S. M. (2008). The Basic AI Drives. In P. Wang, B. Goertzel & S. Franklin (Eds.), *Proceedings of the First AGI Conference, Frontiers in Artificial Intelligence and Applications* (*Vol. 171, pp*. 483-492). *Amsterdam: IOS Press.*

Pearce, D. (2012). The biointelligence explosion. In A. Eden, J. Moor, J. Soraker, & E. Steinhart (Eds.), *The Singularity Hypothesis: A Scientific and Philosophical Assessment* (pp. 199–235). Berlin: Springer. doi:10.1007/978-3-642-32560-1_11

Schmidhuber, J. (2009). Simple Algorithmic Theory of Subjective Beauty, Novelty, Surprise, Interestingness, Attention, Curiosity, Creativity, Art, Science, Music, Jokes. *Journal of SICE*, *48*(1), 21–32.

Schmidhuber, J. (2010). Formal Theory of Creativity, Fun, and Intrinsic Motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development*, *2*(3), 230–247. doi:10.1109/TAMD.2010.2056368

Sloman, A. (1984). *The Structure and Space of Possible Minds. The Mind and the Machine: Philosophical Aspects of Artificial Intelligence*. Ellis Horwood LTD.

Stanley, K. O., & Lehman, J. (2015). *Why Greatness Cannot Be Planned: The Myth of the Objective*. Springer. doi:10.1007/978-3-319-15524-1

Walker, M. (2006). Boredom, Experimental Ethics, and Superlongevity. In C. Tandy (Ed.), Death and Anti-Death, Volume 4: Twenty Years After de Beauvoir, Thirty Years After Heidegger. Ann Arbor, MI: Ria University Press.

Yampolskiy, R. V. (2013). Efficiency Theory: A Unifying Theory for Information, Computation and Intelligence. *Journal of Discrete Mathematical Sciences & Cryptography.*, *16*(4-5), 259–277. doi:10.1080/09720529.2013.821361

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Yampolskiy, R. V. (2014). Utility Function Security in Artificially Intelligent Agents. *Journal of Experimental & Theoretical Artificial Intelligence*, *26*(3), 373–389. doi:10.1080/0952813X.2014.895114

Yampolskiy, R. V. (2015). *Artificial Superintelligence: a Futuristic Approach. Chapman and Hall/CRC Press*.

Yampolskiy, R. V., & Fox, J. (2012). Artificial intelligence and the Human Mental Model. In A. Eden, J. Moor, J. Soraker, & E. Steinhart (Eds.), *The Singularity Hypothesis: A Scientific and Philosophical Assessment* (pp. 129–145). Berlin: Springer. doi:10.1007/978-3-642-32560-1_7

Yudkowsky, E. (2008). Artificial Intelligence as a Positive and Negative Factor. In N. Bostrom & M. M. Cirkovic (Eds.), *Global Risk, in Global Catastrophic Risks* (pp. 308–345). Oxford: Oxford University Press.

Yudkowsky, E. (2015). *Rationality*. Berkeley, CA: Machine Intelligence Research Institute.

ENDNOTE

[1] Euler lived around 2,000 year later than Euclid and knew about Euclid's proof, i.e. the knowledge about the infinity of prime numbers had been accessed before, thus no involvement of "novelty fun" for Euler, but his proof was different, thus potential involvement of "process fun" for him.

This research was previously published in the International Journal of Grid and High Performance Computing (IJGHPC), 9(1); edited by Emmanuel Udoh, Ching-Hsien Hsu, and Mohammad Khan; pages 37-47; copyright year 2017 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 67
Data Extrapolation via Curve Modeling in Analyzing Risk:
Value Anticipation for Decision Making

Dariusz Jacek Jakóbczak
Technical University of Koszalin, Poland

ABSTRACT

Risk analysis needs suitable methods of data extrapolation and decision making. Proposed method of Hurwitz-Radon Matrices (MHR) can be used in extrapolation and interpolation of curves in the plane. For example quotations from the Stock Exchange, the market prices or rate of a currency form a curve. This chapter contains the way of data anticipation and extrapolation via MHR method and decision making: to buy or not, to sell or not. Proposed method is based on a family of Hurwitz-Radon (HR) matrices. The matrices are skew-symmetric and possess columns composed of orthogonal vectors. The operator of Hurwitz-Radon (OHR), built from these matrices, is described. Two-dimensional data are represented by the set of curve points. It is shown how to create the orthogonal and discrete OHR and how to use it in a process of data foreseeing and extrapolation. MHR method is interpolating and extrapolating the curve point by point without using any formula or function.

INTRODUCTION

Method of Hurwitz - Radon Matrices (MHR), invented by the author, can be applied in reconstruction and interpolation of curves in the plane. The method is based on a family of Hurwitz - Radon (HR) matrices. The matrices are skew - symmetric and possess columns composed of orthogonal vectors. The operator of Hurwitz - Radon (OHR), built from these matrices, is described. Author explains how to create the orthogonal and discrete OHR and how to use it in a process of curve interpolation and two-dimensional data modeling. Proposed method needs suitable choice of nodes, i.e. points of the 2D curve to be interpolated or extrapolated: nodes should be settled at each extremum (minimum or maximum) of one coordinate and at least one point between two successive local extrema, and nodes should be

DOI: 10.4018/978-1-7998-0951-7.ch067

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

monotonic in one of coordinates (for example equidistance). Created from the family of N - 1 HR matrices and completed with the identical matrix, system of matrices is orthogonal only for vector spaces of dimensions $N = 1, 2, 4$ or 8. Orthogonality of columns and rows is very important and significant for stability and high precision of calculations. MHR method is modeling the curve point by point without using any formula of function. Main features of MHR method are: accuracy of curve reconstruction depending on number of nodes and method of choosing nodes, interpolation of L points of the curve is connected with the computational cost of rank $O(L)$, MHR interpolation is not a linear interpolation (Ullman & Basri, 1991). The problem of curve length estimation is also considered. Algorithm of MHR method and the examples of data extrapolation are described. Value anticipation is the crucial feature in risk analyzing and decision making.

BACKGROUND

The following question is important in mathematics and computer science: is it possible to find a method of curve interpolation and extrapolation in the plane without building the interpolation polynomials? This chapter aims at giving the positive answer to this question. Current methods of curve interpolation are based on classical polynomial interpolation: Newton, Lagrange or Hermite polynomials and spline curves which are piecewise polynomials (Dahlquist & Bjoerck, 1974). Classical methods are useless to interpolate the function that fails to be differentiable at one point, for example the absolute value function $f(x) = |x|$at $x = 0$. If point (0;0) is one of the interpolation nodes, then precise polynomial interpolation of the absolute value function is impossible. Also when the graph of interpolated function differs from the shape of polynomials considerably, for example $f(x) = 1/x$, interpolation is very hard because of existing local extrema of polynomial. We cannot forget about the Runge's phenomenon: when interpolation nodes are equidistance then high-order polynomial oscillates toward the end of the interval, for example close to -1 and 1 with function $f(x) = 1/(1+25x^2)$ (Ralston, 1965). MHR method is free of these bad feature. Computational algorithm is considered and then it is important to talk about time. Complexity of calculations for one unknown point in Lagrange or Newton interpolation based on n nodes is connected with the computational cost of rank $O(n^2)$. Proposed method has lower calculation complexity.

A significant problem in risk analysis and decision making is that of appropriate data representation and extrapolation (Brachman & Levesque, 2004). Two-dimensional data can be treated as points on the curve. Classical polynomial interpolations and extrapolations (Lagrange, Newton, Hermite) are useless for data anticipation, because the stock quotations or the market prices represent discrete data and they do not preserve a shape of the polynomial. Also Richardson extrapolation has some weak sides concerning discrete data. This chapter is dealing with the method of data foreseeing and value extrapolation by using a family of Hurwitz-Radon matrices. The quotations, prices or rate of a currency, represented by curve points, consist of information which allows us to extrapolate the next data and then to make a decision (Fagin et al, 1995).

If the probabilities of possible actions are known, then some criteria are ready to be applied in decision making and analyzing risk: for example criterion of Laplace, Bayes, Wald, Hurwicz, Savage, Hodge-Lehmann (Straffin, 1993) and others (Watson, 2002). But in this chapter author considers only two possibilities: to do something or not. For example to buy a share or not, to sell a currency or not. Proposed method of Hurwitz-Radon Matrices (MHR) is used in data extrapolation and then calculations for risk analyzing and decision making are described. MHR method presents new approach to extrapolation problem because it takes the interpolation nodes to create orthogonal basis as columns

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

of matrix OHR operators. Then affine (convex) combination of such basis builds new orthogonal base for unknown coordinates of calculated points. MHR method uses two-dimensional data for knowledge representation (Markman, 1998) and for computational foundations (Sowa, 2000). Also medicine (Cierniak, 2005), industry (Jakóbczak & Kosiński, 2007) and manufacturing (Tang, 2005) are looking for the methods connected with geometry of the curves (Soussen & Mohammad-Djafari, 2004) . So suitable data representation and precise reconstruction (Latecki & Lakaemper, 1999) or extrapolation (Kozera, 2004) of the curve is a key factor in many applications of computational intelligence (Lowe, 1991 and 2001), knowledge representation and risk analysis. In comparison MHR method with Bézier curves, Hermite curves and B-curves (*B-splines*) or NURBS one unpleasant feature of these curves must be mentioned: small change of one characteristic point can make big change of whole reconstructed curve. Such an unwanted feature does not appear in MHR method. None of the methods, that are used nowadays, applies the orthogonal basis for unknown value extrapolation and foreseeing, but proposed MHR method does. Orthogonality means a very important feature for stability in calculations. Also point MHR interpolation and extrapolation via simple computations with matrices is something new in the problem of anticipation, risk analysis and decision making.

The considered problem statement looks as follow: let's assume there are given some 2D points of known data as the set of interpolation nodes. How the unknown value can be extrapolated and anticipated as the support in decision making and risk analysis, based on given characteristic (key) points?

THE METHOD OF HURWITZ-RADON MATRICES

Issues

This chapter deals with the problem of interpolation and extrapolation without computing the polynomial or any fixed function. The values of nodes are used to build the orthogonal matrix operators OHR and a linear (convex) combination of OHR operators leads to calculations of 2D curve points. Main idea of MHR method (Jakóbczak, 2006) is that the curve is interpolated or extrapolated point by point by computing the unknown coordinates of the points. The only significant factors in MHR method are: choosing the interpolation nodes and fixing the dimension of HR matrix ($N = 1, 2, 4$ or 8). Other characteristic features of function or curve, such as shape or similarity to polynomials, derivative or Runge's phenomenon, are not important in the process of MHR interpolation and extrapolation. The curve or function in MHR method is parameterized for value $\alpha \in [0;1]$ in each range of two successive interpolation nodes. Data extrapolation and foreseeing is calculated for $\alpha < 0$ or $\alpha > 1$. Estimation of the curve length with high precision is possible because of computing any number of curve points we want. Complexity of calculations for L unknown points in MHR interpolation and extrapolation, based on n nodes, is connected with the computational cost of rank $O(L)$. This is very important feature of MHR method.

The Origin of Hurwitz-Radon Matrices

Adolf Hurwitz (1859-1919) and Johann Radon (1887-1956) published separately the papers about specific class of matrices in 1923, working on the problem of quadratic forms. For example equation

$$(x_0^2 + x_1^2) \cdot (y_0^2 + y_1^2) = (z_0^2 + z_1^2)$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

is true when $z_0 = x_0y_0 - x_1y_1$, $z_1 = x_0y_1 + x_1y_0$. This result can be achieved from matrix equation:

$$\begin{bmatrix} x_0 & x_1 \\ -x_1 & x_0 \end{bmatrix} \cdot \begin{bmatrix} y_0 & y_1 \\ -y_1 & y_0 \end{bmatrix} = \begin{bmatrix} z_0 & z_1 \\ -z_1 & z_0 \end{bmatrix}.$$

Also equation

$$(x_0^2 + x_1^2 + x_2^2 + x_3^2) \cdot (y_0^2 + y_1^2 + y_2^2 + y_3^2) = z_0^2 + z_1^2 + z_2^2 + z_3^2$$

has solution:

$$z_0 = x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3,$$

$$z_1 = x_0y_1 + x_1y_0 + x_2y_3 - x_3y_2,$$

$$z_2 = x_0y_2 - x_1y_3 + x_2y_0 + x_3y_1,$$

$$z_3 = x_0y_3 + x_1y_2 - x_2y_1 + x_3y_0.$$

This result can be achieved from matrix equation of dimension $N = 4$:

$$\begin{bmatrix} x_0 & x_1 & x_2 & x_3 \\ -x_1 & x_0 & -x_3 & x_2 \\ -x_2 & x_3 & x_0 & -x_1 \\ -x_3 & -x_2 & x_1 & x_0 \end{bmatrix} \cdot \begin{bmatrix} y_0 & y_1 & y_2 & y_3 \\ -y_1 & y_0 & -y_3 & y_2 \\ -y_2 & y_3 & y_0 & -y_1 \\ -y_3 & -y_2 & y_1 & y_0 \end{bmatrix} = \begin{bmatrix} z_0 & z_1 & z_2 & z_3 \\ -z_1 & z_0 & -z_3 & z_2 \\ -z_2 & z_3 & z_0 & -z_1 \\ -z_3 & -z_2 & z_1 & z_0 \end{bmatrix}.$$

Hurwitz and Radon proved that dimensions $N = 1, 2, 4$ and 8 are the only dimensions for these quadratic equations. When $N = 8$:

$$\left(\sum_{i=0}^{7} x_i^2\right) \cdot \left(\sum_{i=0}^{7} y_i^2\right) = \sum_{i=0}^{7} z_i^2$$

for

$$z_0 = x_0y_0 - x_1y_1 - x_2y_2 - x_3y_3 - x_4y_4 - x_5y_5 - x_6y_6 - x_7y_7,$$

$$z_1 = x_0y_1 + x_1y_0 - x_2y_3 + x_3y_2 - x_4y_5 + x_5y_4 + x_6y_7 - x_7y_6,$$

$$z_2 = x_0y_2 + x_1y_3 + x_2y_0 - x_3y_1 - x_4y_6 - x_5y_7 + x_6y_4 + x_7y_5,$$

$$z_3 = x_0y_3 - x_1y_2 + x_2y_1 + x_3y_0 - x_4y_7 + x_5y_6 - x_6y_5 + x_7y_4,$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$z_4 = x_0y_4 + x_1y_5 + x_2y_6 + x_3y_7 + x_4y_0 - x_5y_1 - x_6y_2 - x_7y_3,$$

$$z_5 = x_0y_5 - x_1y_4 + x_2y_7 - x_3y_6 + x_4y_1 + x_5y_0 + x_6y_3 - x_7y_2,$$

$$z_6 = x_0y_6 - x_1y_7 - x_2y_4 + x_3y_5 + x_4y_2 - x_5y_3 + x_6y_0 + x_7y_1,$$

$$z_7 = x_0y_7 + x_1y_6 - x_2y_5 - x_3y_4 + x_4y_3 + x_5y_2 - x_6y_1 + x_7y_0.$$

The matrices used to solve quadratic equations are defined: matrices A_i, $i = 1, 2, \ldots, m$ satisfying

$$A_jA_k + A_kA_j = 0, A_j^2 = -I, j \neq k, j, k = 1, 2, ..., m$$

are called a family of Hurwitz - Radon matrices.

A family of Hurwitz - Radon (HR) matrices has important features (Eckmann, 1999): HR matrices are skew-symmetric ($A_i^T = - A_i$) and reverse matrices are easy to find ($A_i^{-1} = - A_i$). Only for dimension $N = 1, 2, 4$ or 8 the family of HR matrices consists of $N - 1$ matrices. When $N = 1$ there is no matrix, just calculations with real numbers. For $N = 2$ we have one matrix:

$$A_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

For $N = 4$ there are three HR matrices with integer entries:

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}, A_3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}.$$

For $N = 8$ we have seven HR matrices with elements $0, \pm 1$ (Jakóbczak, 2006).

So far HR matrices are applied in electronics (Citko, Jakóbczak & Sieńko, 2005): in Space-Time Block Coding (STBC) and orthogonal design (Tarokh, Jafarkhani & Calderbank, 1999), also in signal processing (Sieńko, Citko & Wilamowski, 2002) and Hamiltonian Neural Nets (Sieńko & Citko, 2002).

The Operator of Hurwitz-Radon

Here is the beginning of proposed MHR method. Let us consider a combination of identity matrix and HR matrix of dimension $N = 2$:

$$a\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + b\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} a & b \\ -b & a \end{bmatrix}.$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

For any points $(x_1,y_1) \in \boldsymbol{R}^2$, $(x_2,y_2) \in \boldsymbol{R}^2$ matrix equation

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

is true with

$$a = \frac{x_1 y_1 + x_2 y_2}{x_1^{\,2} + x_2^{\,2}},\ b = \frac{x_2 y_1 - x_1 y_2}{x_1^{\,2} + x_2^{\,2}} \text{ and } x_1^{\,2} + x_2^{\,2} > 0.$$

Reverse matrix equation

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \cdot \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is true with

$$a = \frac{x_1 y_1 + x_2 y_2}{y_1^{\,2} + y_2^{\,2}},\ b = \frac{-x_2 y_1 + x_1 y_2}{y_1^{\,2} + y_2^{\,2}} \text{ and } y_1^{\,2} + y_2^{\,2} > 0.$$

Also we can consider a combination of identity matrix and three HR matrices of dimension $N = 4$:

$$a\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + b\begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix} + c\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} + d\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{bmatrix}.$$

For any points $(x_1,y_1) \in \boldsymbol{R}^2$, $(x_2,y_2) \in \boldsymbol{R}^2$, $(x_3,y_3) \in \boldsymbol{R}^2$, $(x_4,y_4) \in \boldsymbol{R}^2$ matrix equation

$$\begin{bmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}$$

is satisfied with

$$a = \frac{x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4}{x_1^{\,2} + x_2^{\,2} + x_3^{\,2} + x_4^{\,2}},\ b = \frac{-x_1 y_2 + x_2 y_1 + x_3 y_4 - x_4 y_3}{x_1^{\,2} + x_2^{\,2} + x_3^{\,2} + x_4^{\,2}},\ c = \frac{-x_1 y_3 - x_2 y_4 + x_3 y_1 + x_4 y_2}{x_1^{\,2} + x_2^{\,2} + x_3^{\,2} + x_4^{\,2}},$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$d = \frac{-x_1 y_4 + x_2 y_3 - x_3 y_2 + x_4 y_1}{x_1^{\,2} + x_2^{\,2} + x_3^{\,2} + x_4^{\,2}} \text{ and } x_1^{\,2} + x_2^{\,2} + x_3^{\,2} + x_4^{\,2} > 0.$$

Reverse matrix equation

$$\begin{bmatrix} a & b & c & d \\ -b & a & -d & c \\ -c & d & a & -b \\ -d & -c & b & a \end{bmatrix} . \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

is satisfied with

$$a = \frac{x_1 y_1 + x_2 y_2 + x_3 y_3 + x_4 y_4}{y_1^{\,2} + y_2^{\,2} + y_3^{\,2} + y_4^{\,2}},\ b = \frac{x_1 y_2 - x_2 y_1 - x_3 y_4 + x_4 y_3}{y_1^{\,2} + y_2^{\,2} + y_3^{\,2} + y_4^{\,2}},\ c = \frac{x_1 y_3 + x_2 y_4 - x_3 y_1 - x_4 y_2}{y_1^{\,2} + y_2^{\,2} + y_3^{\,2} + y_4^{\,2}},$$

$$d = \frac{x_1 y_4 - x_2 y_3 + x_3 y_2 - x_4 y_1}{y_1^{\,2} + y_2^{\,2} + y_3^{\,2} + y_4^{\,2}} \text{ and } y_1^{\,2} + \mathrm{y}_2^{\,2} + y_3^{\,2} + y_4^{\,2} > 0.$$

A combination of identity matrix and seven HR matrices of dimension $N = 8$ looks as follows:

$$\begin{bmatrix} a_0 & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 \\ -a_1 & a_0 & a_3 & -a_2 & a_5 & -a_4 & -a_7 & a_6 \\ -a_2 & -a_3 & a_0 & a_1 & a_6 & a_7 & -a_4 & -a_5 \\ -a_3 & a_2 & -a_1 & a_0 & a_7 & -a_6 & a_5 & -a_4 \\ -a_4 & -a_5 & -a_6 & -a_7 & a_0 & a_1 & a_2 & a_3 \\ -a_5 & a_4 & -a_7 & a_6 & -a_1 & a_0 & -a_3 & a_2 \\ -a_6 & a_7 & a_4 & -a_5 & -a_2 & a_3 & a_0 & -a_1 \\ -a_7 & -a_6 & a_5 & a_4 & -a_3 & -a_2 & a_1 & a_0 \end{bmatrix}.$$

Results for matrix equations with a combination of identity matrix and seven HR matrices of dimension $N = 8$ are calculated in formulas (3), (4) and (8). Solutions of matrix equations (Sieńko, Citko & Jakóbczak, 2004) are used to build the matrix Operator of Hurwitz – Radon (OHR).

Let's assume there is given a finite set of points of the curve, called further nodes $(x_i,y_i) \in \boldsymbol{R^2}$ such as:

1. Nodes (interpolation points) are settled at local extrema (maximum or minimum) of one of coordinates and at least one point between two successive local extrema;
2. Each node (x_i,y_i) is monotonic in coordinates x_i or y_i (for example equidistance in one of coordinates).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Assume that the nodes belong to a curve in the plane. How the whole curve could be interpolated or extrapolated using this discrete set of nodes? Proposed method (Jakóbczak, 2007) is based on local, orthogonal matrix operators. The values of nodes' coordinates (x_i,y_i) are connected with HR matrices, built on N - dimensional vector space. It is important that HR matrices are skew-symmetric and only for dimensions $N = 1, 2, 4$ or 8 columns and rows of HR matrices are orthogonal (Eckmann, 1999).

If one curve is described by a set of nodes $\{(x_i,y_i), i = 1, 2, \ldots, n\}$ monotonic (for example equidistance) in coordinates x_i, then HR matrices combined with identity matrix are used to build an orthogonal and discrete Hurwitz - Radon Operator (OHR). For nodes (x_1,y_1), (x_2,y_2) OHR of dimension $N = 2$ is constructed:

$$M = \frac{1}{x_1^{\,2} + x_2^{\,2}} \begin{bmatrix} x_1 & x_2 \\ -x_2 & x_1 \end{bmatrix} \begin{bmatrix} y_1 & -y_2 \\ y_2 & y_1 \end{bmatrix}, \; M = \frac{1}{x_1^{\,2} + x_2^{\,2}} \begin{bmatrix} x_1y_1 + x_2y_2 & x_2y_1 - x_1y_2 \\ x_1y_2 - x_2y_1 & x_1y_1 + x_2y_2 \end{bmatrix}. \tag{1}$$

For nodes (x_1,y_1), (x_2,y_2), (x_3,y_3), (x_4,y_4) monotonic (for example equidistance) in x_i OHR of dimension $N = 4$ is constructed:

$$M = \frac{1}{x_1^{\,2} + x_2^{\,2} + x_3^{\,2} + x_4^{\,2}} \begin{bmatrix} u_0 & u_1 & u_2 & u_3 \\ -u_1 & u_0 & -u_3 & u_2 \\ -u_2 & u_3 & u_0 & -u_1 \\ -u_3 & -u_2 & u_1 & u_0 \end{bmatrix} \tag{2}$$

where

$$u_0 = x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4, \; u_1 = -x_1y_2 + x_2y_1 + x_3y_4 - x_4y_3$$
$$u_2 = -x_1y_3 - x_2y_4 + x_3y_1 + x_4y_2, \; u_3 = -x_1y_4 + x_2y_3 - x_3y_2 + x_4y_1.$$

For nodes (x_1,y_1), (x_2,y_2), …, (x_8,y_8) monotonic in x_i OHR of dimension $N = 8$ is equal with

$$M = \frac{1}{\sum_{i=1}^{8} x_i^{\,2}} \begin{bmatrix} u_0 & u_1 & u_2 & u_3 & u_4 & u_5 & u_6 & u_7 \\ -u_1 & u_0 & u_3 & -u_2 & u_5 & -u_4 & -u_7 & u_6 \\ -u_2 & -u_3 & u_0 & u_1 & u_6 & u_7 & -u_4 & -u_5 \\ -u_3 & u_2 & -u_1 & u_0 & u_7 & -u_6 & u_5 & -u_4 \\ -u_4 & -u_5 & -u_6 & -u_7 & u_0 & u_1 & u_2 & u_3 \\ -u_5 & u_4 & -u_7 & u_6 & -u_1 & u_0 & -u_3 & u_2 \\ -u_6 & u_7 & u_4 & -u_5 & -u_2 & u_3 & u_0 & -u_1 \\ -u_7 & -u_6 & u_5 & u_4 & -u_3 & -u_2 & u_1 & u_0 \end{bmatrix} \tag{3}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

where

$$u = \begin{bmatrix} y_1 & y_2 & y_3 & y_4 & y_5 & y_6 & y_7 & y_8 \\ -y_2 & y_1 & -y_4 & y_3 & -y_6 & y_5 & y_8 & -y_7 \\ -y_3 & y_4 & y_1 & -y_2 & -y_7 & -y_8 & y_5 & y_6 \\ -y_4 & -y_3 & y_2 & y_1 & -y_8 & y_7 & -y_6 & y_5 \\ -y_5 & y_6 & y_7 & y_8 & y_1 & -y_2 & -y_3 & -y_4 \\ -y_6 & -y_5 & y_8 & -y_7 & y_2 & y_1 & y_4 & -y_3 \\ -y_7 & -y_8 & -y_5 & y_6 & y_3 & -y_4 & y_1 & y_2 \\ -y_8 & y_7 & -y_6 & -y_5 & y_4 & y_3 & -y_2 & y_1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \end{bmatrix} \tag{4}$$

We can see here that the components of the vector $\mathbf{u} = (u_0, u_1, \ldots, u_7)^{\mathrm{T}}$, appearing in the matrix M in formula (3) are defined by (4) in the similar way to formula (2) but in terms of the coordinates of the above 8 nodes. Note that OHR operators (1)-(3) satisfy the condition of interpolation

$$M \cdot \mathbf{x} = \mathbf{y} \tag{5}$$

for $\mathbf{x} = (x_1, x_2 \ldots, x_N)^{\mathrm{T}} \in \boldsymbol{R}^N$, $\mathbf{x} \neq \mathbf{0}$, $\mathbf{y} = (y_1, y_2 \ldots, y_N)^{\mathrm{T}} \in \boldsymbol{R}^N$, $N = 1, 2, 4$ or 8.

If one curve is described by a set of nodes $\{(x_i, y_i), i = 1, 2, \ldots, n\}$ monotonic (for example equidistance) in coordinates y_i, then HR matrices combined with identity matrix are used to build an orthogonal and discrete reverse Hurwitz - Radon Operator (reverse OHR). For nodes (x_1, y_1), (x_2, y_2) reverse OHR of dimension $N = 2$ is constructed:

$$M^{-1} = \frac{1}{y_1^{\,2} + y_2^{\,2}} \begin{bmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{bmatrix} \begin{bmatrix} y_1 & y_2 \\ -y_2 & y_1 \end{bmatrix}, \; M^{-1} = \frac{1}{y_1^{\,2} + y_2^{\,2}} \begin{bmatrix} x_1 y_1 + x_2 y_2 & -x_2 y_1 + x_1 y_2 \\ -x_1 y_2 + x_2 y_1 & x_1 y_1 + x_2 y_2 \end{bmatrix}. \tag{6}$$

For nodes (x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4) monotonic in y_i the reverse OHR of dimension $N = 4$ is constructed with u_o, u_1, u_2 and u_3 from (2):

$$M^{-1} = \frac{1}{y_1^{\,2} + y_2^{\,2} + y_3^{\,2} + y_4^{\,2}} \begin{bmatrix} u_0 & -u_1 & -u_2 & -u_3 \\ u_1 & u_0 & u_3 & -u_2 \\ u_2 & -u_3 & u_0 & u_1 \\ u_3 & u_2 & -u_1 & u_0 \end{bmatrix}. \tag{7}$$

For nodes (x_1, y_1), (x_2, y_2), …, (x_8, y_8) monotonic in y_i the reverse OHR of dimension $N = 8$ is equal with

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$M^{-1} = \frac{1}{\sum_{i=1}^{8} y_i^{\,2}} \begin{bmatrix} u_0 & -u_1 & -u_2 & -u_3 & -u_4 & -u_5 & -u_6 & -u_7 \\ u_1 & u_0 & -u_3 & u_2 & -u_5 & u_4 & u_7 & -u_6 \\ u_2 & u_3 & u_0 & -u_1 & -u_6 & -u_7 & u_4 & u_5 \\ u_3 & -u_2 & u_1 & u_0 & -u_7 & u_6 & -u_5 & u_4 \\ u_4 & u_5 & u_6 & u_7 & u_0 & -u_1 & -u_2 & -u_3 \\ u_5 & -u_4 & u_7 & -u_6 & u_1 & u_0 & u_3 & -u_2 \\ u_6 & -u_7 & -u_4 & u_5 & u_2 & -u_3 & u_0 & u_1 \\ u_7 & u_6 & -u_5 & -u_4 & u_3 & u_2 & -u_1 & u_0 \end{bmatrix}, \tag{8}$$

where the components of the vector $\mathbf{u} = (u_0, u_1,\ldots, u_7)^{\mathrm{T}}$ are defined in terms of (4). Note that reverse OHR operators (6)-(8) satisfy the condition of interpolation

$$M^{-1}\cdot\mathbf{y} = \mathbf{x} \tag{9}$$

for $\mathbf{x} = (x_1,x_2\ldots,x_N)^{\mathrm{T}} \in \boldsymbol{R}^N$, $\mathbf{y} = (y_1,y_2\ldots,y_N)^{\mathrm{T}} \in \boldsymbol{R}^N$, $\mathbf{y} \neq \mathbf{0}$, $N = 1, 2, 4$ or 8.

Known values, for example currency rates or market prices, are represented on 2D curve by the set of nodes $(x_i,y_i) \in \boldsymbol{R}^2$ (characteristic points) as follows in proposed MHR method:

1. Nodes (interpolation points) are settled at local extrema (maximum or minimum) of one of coordinates and at least one point between two successive local extrema;
2. Nodes (x_i,y_i) are monotonic in coordinates x_i ($x_i < x_{i+1}$ for all i) or y_i ($y_i < y_{i+1}$);
3. One curve is represented by at least four nodes.

Condition 1 is done for the most appropriate description of a curve. The quotations or prices are real data. Condition 2 according to a graph of function means that coordinates x_i represent for example the time. Condition 3 is adequate for extrapolation, but in the case of interpolation minimal number of nodes is five.

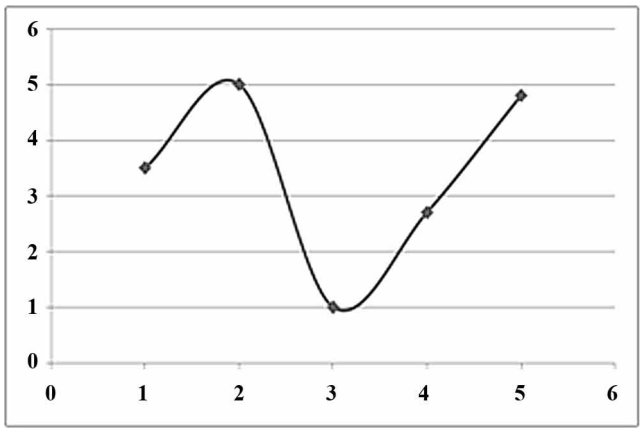
Data points are treated as interpolation nodes. How can we extrapolate continues values at time $x = 5.5$ for example or discrete data for next day $x = 6$ (see Figure 1)? The anticipation of values is possible using proposed MHR method.

THE METHOD OF HURWITZ-RADON MATRICES

Key question looks as follows: how can we compute coordinates of points settled between interpolation nodes or beyond the nodes? The answer is connected with MHR method for interpolation and extrapolation. On a segment of a line every number "c" situated between "a" and "b" is described by a linear (convex) combination $c = \alpha \cdot a + (1\text{-}\alpha) \cdot b$ for

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Five nodes of data and a curve. Source: author, 2014



$$\alpha = \frac{b-c}{b-a} \in [0;1]. \tag{10}$$

If $c < a$ then $\alpha > 1$: there is extrapolation of points situated left of nodes. If $c > b$ then $\alpha < 0$: extrapolation of points situated right of nodes.

When the nodes are monotonic in coordinates x_i, average OHR operator M_2 of dimension $N = 1, 2, 4$ or 8 is constructed as follows:

$$M_2 = \alpha \bullet M_0 + (1 - \alpha) \bullet M_1 \tag{11}$$

with the operator M_0 built (1)-(3) by "odd" nodes $(x_1=a,y_1)$, (x_3,y_3), …, (x_{2N-1},y_{2N-1}) and M_1 built (1)-(3) by "even" nodes $(x_2=b,y_2)$, (x_4,y_4), …, (x_{2N},y_{2N}).

When the nodes are monotonic in coordinates y_i, average reverse OHR operator M_2^{-1} of dimension $N = 1, 2, 4$ or 8 is constructed as follows:

$$M_2^{-1} = \alpha \cdot M_0^{-1} + (1-\alpha) \cdot M_1^{-1} \tag{12}$$

with the reverse operator M_0^{-1} built (6)-(8) by nodes $(x_1,y_1=a)$, (x_3,y_3), …, (x_{2N-1},y_{2N-1}) and M_1^{-1} built (6)-(8) by nodes $(x_2,y_2=b)$, (x_4,y_4), …, (x_{2N},y_{2N}).

Notice that having the operator M_2 for coordinates $x_i < x_{i+1}$ it is possible to reconstruct the second coordinates of points (x,y) in terms of the vector C defined with

$$c_i = \alpha \cdot x_{2i-1} + (1-\alpha) \cdot x_{2i},\ i = 1, 2, \ldots, N \tag{13}$$

as $C = [c_1, c_2, \ldots, c_N]^T$. The required formula is similar to (5):

$$Y(C) = M_2 \bullet C \tag{14}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

in which components of vector $Y(C)$ give the second coordinate of the points (x,y) corresponding to the first coordinate, given in terms of components of the vector C.

On the other hand, having the operator M_2^{-1} for coordinates $y_i < y_{i+1}$ it is possible to reconstruct the first coordinates of points (x,y) in terms of the corresponding second coordinates given by components of the new vector C defined, as previously, with

$$c_i = \alpha \cdot y_{2i-1} + (1-\alpha) \cdot y_{2i},\ i = 1, 2, \ldots, N \tag{15}$$

and $C = [c_1, c_2, \ldots, c_N]^T$. The final formula is similar to (9):

$$X(C) = M_2^{-1} \cdot C \tag{16}$$

in which components of the vector $X(C)$ give the first coordinate of the points (x,y) corresponding to the second coordinate, given in terms of components of the vector C.

After computing (14) or (16) for any $\alpha \in [0;1]$, we have a half of reconstructed points ($j = 1$ in algorithm 1). Now it is necessary to find second half of unknown coordinates ($j = 2$ in algorithm 1) for

$$c_i = \alpha \cdot x_{2i} + (1-\alpha) \cdot x_{2i+1},\ i = 1, 2, \ldots, N \tag{17}$$

or

$$c_i = \alpha \cdot y_{2i} + (1-\alpha) \cdot y_{2i+1},\ i = 1, 2, \ldots, N \tag{18}$$

depending on whether x_i (17) or y_i (18) is monotonic. There is no need to build the OHR for nodes $(x_2=a,y_2)$, (x_4,y_4), …, (x_{2N},y_{2N}) or the reverse OHR for nodes $(x_2,y_2=a)$, (x_4,y_4), …, (x_{2N},y_{2N}), because we have just found M_1 or M_1^{-1}. This operator will play the role as M_0 or M_0^{-1} in (11) or (12). New M_1 or M_1^{-1} must be computed for nodes $(x_3=b,y_3)$, …, (x_{2N-1},y_{2N-1}), (x_{2N+1},y_{2N+1}) or $(x_3,y_3=b)$, …, (x_{2N-1},y_{2N-1}), (x_{2N+1},y_{2N+1}).

As we see the minimum number of interpolation nodes $n = 2N+1 = 5$, 9 or 17 using OHR operators of dimension $N = 2$, 4 or 8 respectively. If there is more nodes than $2N+1$, the same calculations (10)-(18) have to be done for next range(s) or last range of $2N+1$ nodes. For example, if $n = 9$ then we can use OHR operators of dimension $N = 4$ or OHR operators of dimension $N = 2$ for two subsets of nodes: $\{(x_1,y_1), \ldots, (x_5,y_5)\}$ and $\{(x_5,y_5), \ldots, (x_9,y_9)\}$.

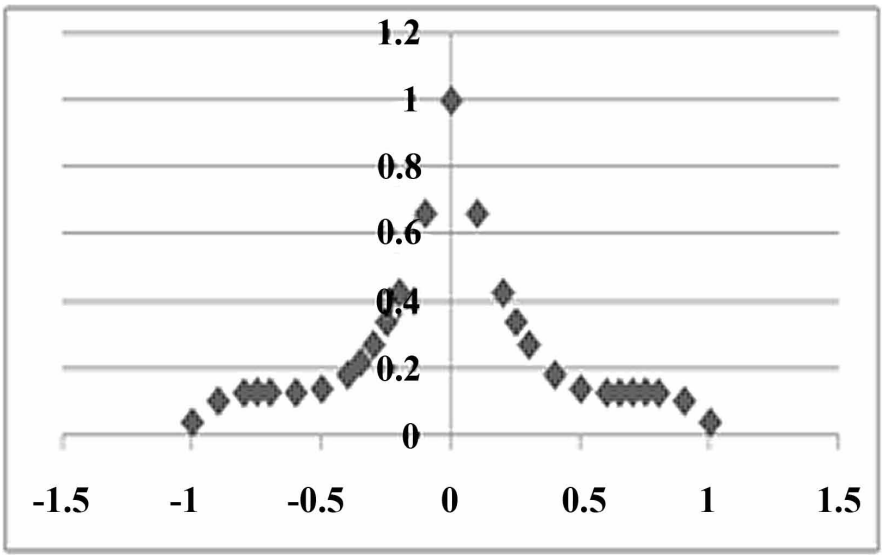
Here is the application of MHR method for function $f(x)=1/(1+25x^2)$ with five nodes equidistance in first coordinate: $x_i = -1, -0.5, 0, 0.5, 1$ (see Figure 2).

MHR interpolation of function $f(x) = 1/(1+25x^2)$ gives better result then polynomial interpolation: there is no Runge phenomenon. The same can be said for function $f(x) = 1/x$.

MHR extrapolation is valid for $\alpha < 0$ or $\alpha > 1$. In the case of continues data, parameter α is a real number. For example there are four nodes: (1; 2), (1.3;5), (2;3), (2.5;6). MHR extrapolation with $\alpha = -0.01$ gives the point (2.505; 6.034) and with $\alpha = -0.1$: (2.55;6.348). But the rate of a currency or the quotations are discrete data. If we assume that the rate of a currency is represented by equidistance nodes (day by day – fixed step of time $h = 1$ for coordinate x), next data or the rate on next day is extrapolated (anticipated) for $\alpha = -1$. Calculation of unknown coordinates for data points using formulas (10)-(18) is called by author the method of Hurwitz - Radon Matrices (MHR).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Twenty six interpolated points of function $f(x) = 1/(1+25x^2)$ using MHR method with 5 nodes. Source: author, 2014



Algorithm and Complexity of MHR Calculations

The algorithm of points reconstruction for $2N+1 = 5$, 9 or 17 successive nodes is presented.

Algorithm 1.

Let $j = 1$. *Input: Set of interpolation nodes* $\{(x_i,y_i)$, $i = 1,2,\ldots,n$; $n = 5, 9$ *or* $17\}$ *such as:*

a) nodes are settled at local extrema (maximum or minimum) of one of coordinates and at least one point between two successive local extrema;
b) nodes (x_i,y_i) are monotonic in coordinates x_i or y_i.

Step 1: *Determine the dimension N of OHR operators:* $N = 2$ *if* $n = 5$, $N = 4$ *if* $n = 9$, $N = 8$ *if* $n = 17$.

Step 2: *If nodes are monotonic in* x_i *then build* M_0 *for nodes* $(x_1=a,y_1)$, (x_3,y_3), …, (x_{2N-1},y_{2N-1}) *and* M_1 *for nodes* $(x_2=b,y_2)$, (x_4,y_4), …, (x_{2N},y_{2N}) *from (1)-(3). If nodes are monotonic in* y_i *then build* M_0^{-1} *for nodes* $(x_1,y_1=a)$, (x_3,y_3), …,(x_{2N-1},y_{2N-1}) *and* M_1^{-1} *for nodes* $(x_2,y_2=b)$, (x_4,y_4), …, (x_{2N},y_{2N}) *from* (6)-(8).

Step 3: *Determine the number of points to be reconstructed* $K_j > 0$ *between two successive nodes, let* $k = 1$.

Step 4: *Compute* $\alpha \in [0;1]$ *from (10) for* $c_1 = c = \alpha \cdot a + (1-\alpha) \cdot b$.

Step 5: *Build* M_2 *from (11) or* M_2^{-1} *from (12).*

Step 6: *Compute vector* $C = [c_1, c_2,\ldots, c_N]^T$ *from (13) or (15).*

Step 7: *Compute unknown coordinates* $Y(C)$ *from (14) or* $X(C)$ *from (16).*

Step 8: *If* $k < K_j$, *set* $k = k + 1$ *and go to Step 4. Otherwise if* $j = 1$, *set* $M_0 = M_1$, $a = x_2$, $b = x_3$ *(if nodes are monotonic in* x_i*) or* $M_0^{-1} = M_1^{-1}$, $a = y_2$, $b = y_3$ *(if nodes are monotonic in* y_i*), build new* M_1 *or* M_1^{-1} *for nodes* (x_3,y_3), (x_5,y_5), …,(x_{2N+1},y_{2N+1}), *let* $j = 2$ *and go to Step 3. Otherwise, stop.*

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The number of reconstructed points in algorithm 1 is $K = N(K_1+K_2)$. If there are more nodes than $2N+1 = 5$, 9 or 17, algorithm 1 has to be done for next range(s) or last range of $2N + 1$ nodes. Reconstruction of curve points using algorithm 1 is called by author the method of Hurwitz - Radon Matrices (MHR). If we have n interpolation nodes, then there is $K = L – n$ points to find using Algorithm 1 and MHR method. The complexity of MHR calculations has to be considered.

Lemma 1: Let $n = 5$, 9 or 17 is the number of interpolation nodes, let MHR method (algorithm 1) is done for reconstruction of the curve consists of L points. Then MHR method is connected with the computational cost of rank $O(L)$.

Proof. Using algorithm 1 we have to reconstruct $K = L – n$ points of unknown curve. Counting the number of multiplications and divisions D in algorithm 1, for example in (20) in case $n = 5$ for each c_i at first and second half of reconstructed points, here are the results:

1) $D = 4L+7$ for $n = 5$ and $L = 2i + 5$;
2) $D = 6L+21$ for $n = 9$ and $L = 4i + 9$;
3) $D = 10L+73$ for $n = 17$ and $L = 8i + 17$; $i = 2,3,4...$

The lowest computational cost appears in MHR method with five nodes and OHR operators of dimension $N = 2$. Therefore whole set of n nodes can be divided into subsets of five nodes. Then whole curve is to be reconstructed by algorithm 1 with all subsets of five nodes: $\{(x_1,y_1), \ldots, (x_5,y_5)\}$, $\{(x_5,y_5), \ldots, (x_9,y_9)\}$,

$\{(x_9,y_9), \ldots, (x_{13},y_{13})\}\ldots$ If the last node (x_n,y_n) is indexed $n \neq 4i + 1$, then we have to use last five nodes $\{(x_{n-4},y_{n-4}), \ldots, (x_n,y_n)\}$ in algorithm 1.

The Formulas for a Single Point and Error Estimation

Now there are the formulas for computing one unknown coordinate of a single point. Assume there are given four nodes (x_1,y_1), (x_2,y_2), (x_3,y_3), (x_4,y_4) monotonic in x_i. OHR operators of dimension $N = 2$ are built (1) as follows:

$$M_0 = \frac{1}{x_1^{\,2} + x_3^{\,2}}\begin{bmatrix} x_1y_1 + x_3y_3 & x_3y_1 - x_1y_3 \\ x_1y_3 - x_3y_1 & x_1y_1 + x_3y_3 \end{bmatrix},\ M_1 = \frac{1}{x_2^{\,2} + x_4^{\,2}}\begin{bmatrix} x_2y_2 + x_4y_4 & x_4y_2 - x_2y_4 \\ x_2y_4 - x_4y_2 & x_2y_2 + x_4y_4 \end{bmatrix}.$$

Let first coordinate c_1 of reconstructed point is situated between x_1 and x_2:

$$c_1 = \alpha\cdot x_1 + \beta\cdot x_2 \text{ for } 0 \le \beta = 1 - \alpha \le 1. \tag{19}$$

Compute second coordinate of reconstructed point $y(c_1)$ for $Y(C) = [y(c_1), y(c_2)]^T$ from (14):

$$\begin{bmatrix} y(c_1) \\ y(c_2) \end{bmatrix} = (\alpha\cdot M_0 + \beta\cdot M_1)\cdot\begin{bmatrix} \alpha\cdot x_1 + \beta\cdot x_2 \\ \alpha\cdot x_3 + \beta\cdot x_4 \end{bmatrix}. \tag{20}$$

After calculation (20):

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$y(c_1) = \alpha^2 \cdot y_1 + \beta^2 \cdot y_2 + \frac{\alpha \cdot \beta}{x_1^{\,2} + x_3^{\,2}}(x_1x_2y_1 + x_2x_3y_3 + x_3x_4y_1 - x_1x_4y_3) + \frac{\alpha \cdot \beta}{x_2^{\,2} + x_4^{\,2}}(x_1x_2y_2 + x_1x_4y_4 + x_3x_4y_2 - x_2x_3y_4). \quad (21)$$

So each point of the curve $P = (c_1, y(c_1))$ settled between nodes (x_1,y_1) and (x_2,y_2) is parameterized by $P(\alpha)$ for (19), (21) and $\alpha \in [0;1]$. Similar calculations are done for nodes (x_1,y_1), (x_2,y_2), (x_3,y_3), (x_4,y_4) monotonic in y_i to compute $x(c_1)$:

$$M_0^{-1} = \frac{1}{y_1^{\,2} + y_3^{\,2}}\begin{bmatrix} x_1y_1 + x_3y_3 & -x_3y_1 + x_1y_3 \\ -x_1y_3 + x_3y_1 & x_1y_1 + x_3y_3 \end{bmatrix},\ M_1^{-1} = \frac{1}{y_2^{\,2} + y_4^{\,2}}\begin{bmatrix} x_2y_2 + x_4y_4 & -x_4y_2 + x_2y_4 \\ -x_2y_4 + x_4y_2 & x_2y_2 + x_4y_4 \end{bmatrix},$$

$c_1 = \alpha \cdot y_1 + \beta \cdot y_2$ for $0 \le \beta = 1 - \alpha \le 1$, (22)

$$\begin{bmatrix} x(c_1) \\ x(c_2) \end{bmatrix} = (\alpha \cdot M_0^{-1} + \beta \cdot M_1^{-1}) \cdot \begin{bmatrix} \alpha \cdot y_1 + \beta \cdot y_2 \\ \alpha \cdot y_3 + \beta \cdot y_4 \end{bmatrix},\ x(c_1) = \alpha^2 \cdot x_1 + \beta^2 \cdot x_2 + \frac{\alpha \cdot \beta}{y_1^{\,2} + y_3^{\,2}} r_1 + \frac{\alpha \cdot \beta}{y_2^{\,2} + y_4^{\,2}} r_2 \quad (23)$$

for $r_1 = \text{const.}$, $r_2 = \text{const.}$ depending on nodes' coordinates: see formula (21). If nodes are monotonic in y_i, there is parameterization of curve points P settled between nodes $(x_1, y1)$ and (x_2,y_2): $P(\alpha) = (x(c_1), c_1)$ for (22), (23) and $\alpha \in [0;1]$.

If nodes (x_i,y_i) are equidistance in one coordinate, then calculation of one unknown coordinate is simpler. Let four successive nodes (x_1,y_1), (x_2,y_2), (x_3,y_3), (x_4,y_4) are equidistance in coordinate x_i and $a = x_1$, $h/2 = x_{i+1} - x_i = \text{const}$. Calculations in formulas (20)-(21) are done for c_1 (19):

$$y(c_1) = \alpha y_1 + \beta y_2 + \alpha\beta s \quad (24)$$

and

$$s = h\left(\frac{2ay_1 + hy_1 + hy_3}{4a^2 + 4ah + 2h^2} - \frac{2ay_2 + 2hy_2 + hy_4}{4a^2 + 8ah + 5h^2}\right). \quad (25)$$

As we can see in formulas (21) and (23)-(25), MHR interpolation is not a linear interpolation. It is possible to estimate the interpolation error of MHR method (algorithm 1) for the class of linear function f:

$$\left|f(c_1) - y(c_1)\right| = \left|\alpha y_1 + \beta y_2 - y(c_1)\right| = \alpha\beta\left|s\right|. \quad (26)$$

Notice that estimation (26) has the biggest value $0.25|s|$ for $\beta = \alpha = 0.5$, when c_1 (19) is situated in the middle between x_1 and x_2.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Having four successive nodes (x_1,y_1), (x_2,y_2), (x_3,y_3), (x_4,y_4) equidistance in coordinate x_i ($a = x_1$, $h/2 = x_{i+1} - x_i = \text{const.}$) we can compute polynomial $W_3(x) = m_3x^3 + m_2x^2 + m_1x + t$ for these nodes and estimate the interpolation error of MHR method (algorithm 1) for the class of order three polynomials. After solving the system of equations:

$$y_1 = m_3a^3 + m_2a^2 + m_1a + t,$$

$$y_2 = m_3(a+\frac{h}{2})^3 + m_2(a+\frac{h}{2})^2 + m_1(a+\frac{h}{2}) + t,$$

$$y_3 = m_3(a+h)^3 + m_2(a+h)^2 + m_1(a+h) + t,$$

$$y_4 = m_3(a+\frac{3h}{2})^3 + m_2(a+\frac{3h}{2})^2 + m_1(a+\frac{3h}{2}) + t,$$

it is possible to compute m_3, m_2, m_1, t and the estimation for point c_1 (19):

$$\left|W_3(c_1) - y(c_1)\right| = \beta\left|\frac{1}{2}h(y_2 - y_1)\frac{12\alpha\beta a + 12a\beta^2 + 2h\beta^2 - \alpha\beta ah^2 - 12a - 2h}{6ha + 12a^2 + h^2} - \alpha\cdot s\right|. \tag{27}$$

Notice that estimations (26)-(27) are equal with zero for $\alpha = 0$ or $\beta = 0$ (in nodes). For eight successive nodes (x_i,y_i), $i = 1,2,\ldots,8$ equidistance in coordinate x_i, $a = x_1$, $h/2 = x_{i+1} - x_i = =$ const., using OHR of dimension $N = 4$ in formula (14) here is a formula of second coordinate reconstruction with first coordinate c_1 (19):

$$\begin{aligned} y(c_1) = {} & \alpha y_1 + \beta y_2 + \frac{\beta h^2}{4a^2 + 16ah + 21h^2}[(\alpha y_3 + \beta y_4) + 2(\alpha y_7 + \beta y_8)] \\ & + \frac{\beta h}{(4a^2 + 16ah + 21h^2)(2a^2 + 6ah + 7h^2)}[2(\alpha y_1 + \beta y_2)(2a^3 + 10a^2h + 19ah^2 + 14h^3) \\ & + \alpha h(ay_1 + hy_7)(2a + 7h) + 3.5\alpha h^3(y_1 + y_3) \\ & + 2ah(\alpha hy_3 + \alpha hy_7 + 17ay_2 - ay_4 - 2ay_8 - 3hy_4 - 6hy_8) \\ & - 7h^3(21y_2 + y_4 + 2y_8) - 2ay_2(2a^2 + 27ah + 19h^2)]. \end{aligned} \tag{28}$$

So we have another parameterization (28) of the point $P(\alpha) = (c_1, y(c_1))$ for $N = 4$ and $\beta = 1 - \alpha$. Formula (28) doesn't include values y_5 and y_6: algorithm 1 with nine successive nodes (x_i,y_i), $i = 1,2,\ldots,9$ equidistance in coordinate x_i is free of using y_5 and y_6 for computing second coordinate of the point settled between first and second node.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Algorithm 1 deals with average OHR operators (11)-(12) built with two OHR. This situation leads to parameterization of reconstructed point $P(\alpha) = (c_1, y(c_1))$ or $P(\alpha) = (x(c_1), c_1)$ settled between two successive nodes, where $\alpha \in [0;1]$ is order two in formulas (21), (23)-(25) and (28). The curve or data in MHR method are parameterized for value $\alpha \in [0;1]$ in the range of two successive interpolation nodes. MHR for data extrapolation is possible with $\alpha < 0$ or $\alpha > 1$.

Risk Analysis and Decision Making via Data Extrapolation

Example 1

MHR calculations are done for true rates of euro at National Bank of Poland (NBP) from January 24th to February 14th, 2011. If last four rates are considered: (1;3.8993), (2;3.9248), (3;3.9370) and (4;3.9337), MHR extrapolation with matrices of dimension $N = 2$ gives the result (5;3.9158). So anticipated rate of euro on the day February 15th is 3.9158 (see Figure 3).

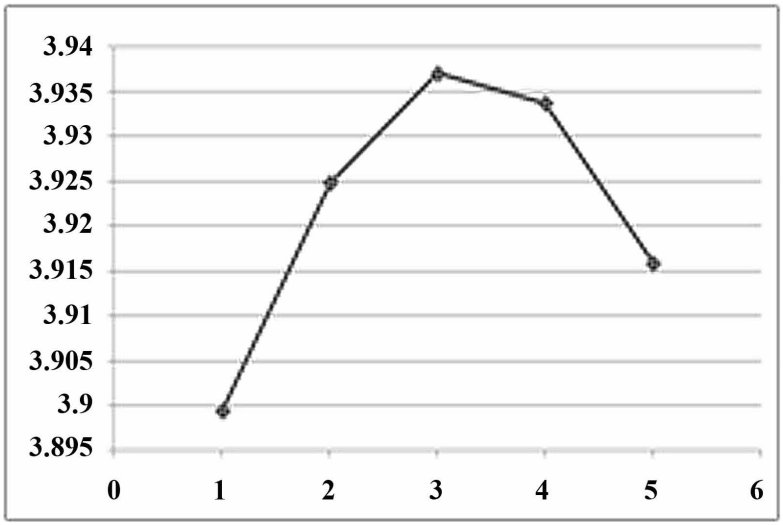
If last eight rates are considered: (1;3.9173), (2;3.9075), (3;3.8684), (4;3.8742), (5;3.8993), (6;3.9248), (7;3.9370) and (8;3.9337), MHR extrapolation with matrices of dimension $N = 4$ gives the result (9;4.0767). Anticipated rate of euro on the day February 15th is 4.0767 (see Figure 4).

There are two extrapolated values for next day. This example gives us two anticipated rates for tomorrow: 3.9158 and 4.0767, which differs considerably. How these extrapolated values can be used in the process of decision making and analyzing risk: to buy euro or not, to sell euro or not? The proposal final anticipated rate of euro for the day February 15th (see Figure 5) based on weighted mean value:

$$\frac{2 \cdot 3.9158 + 4.0767}{3} = 3.9694 \tag{29}$$

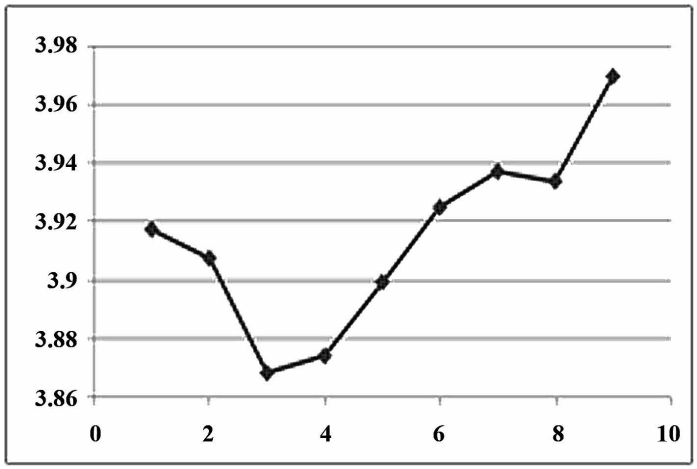
because the rate 3.9158 is calculated for $N = 2$, whereas 4.0767 is extrapolated for

Figure 3. Extrapolated rate for day 5 (February 15th) using MHR method with 4 nodes
Source: author, 2014



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Extrapolated rate for day 9 (February 15th) using MHR method with 8 nodes
Source: author, 2014



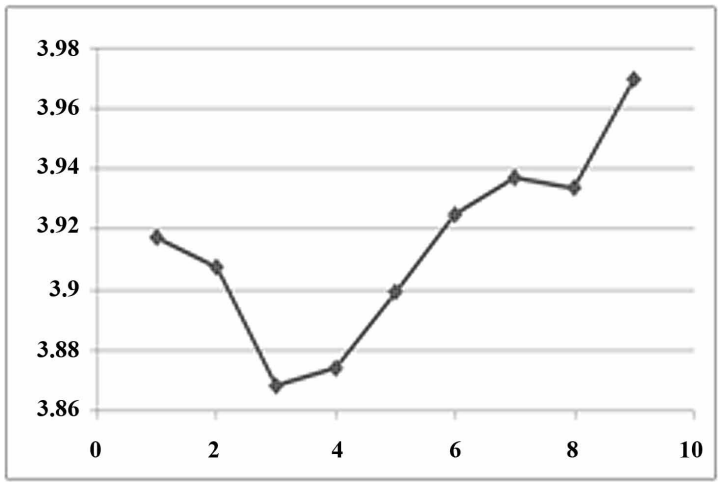
$N = 4$. Formula (29) takes one fact into account: dimension $N = 4$ is two times bigger than dimension $N = 2$ and the result 3.9158 has to be strengthen multiplying by two.

If last sixteen rates are considered, MHR extrapolation with matrices of dimension $N = 8$ has to be used. Here are the rates: (1;3.8765), (2;3.8777), (3;3.8777), (4;3.9009), (5;3.9111), (6;3.9345), (7;3.9129), (8;3.9019), (9;3.9173), (10;3.9075), (11;3.8684), (12;3.8742), (13;3.8993), (14;3.9248), (15;3.9370) and (16;3.9337). Average OHR operator M_2 and MHR calculations look as follows:

$$M_2 = \begin{bmatrix} 0.3226 & 0.0154 & 0.0286 & 0.0462 & -0.062 & 0.0444 & 0.0924 & 0.0461 \\ -0.0154 & 0.3226 & 0.0462 & -0.0286 & 0.0444 & 0.062 & -0.0461 & 0.924 \\ -0.0286 & -0.0462 & 0.3226 & 0.0154 & 0.0924 & 0.0461 & 0.062 & -0.0444 \\ -0.0462 & 0.0286 & -0.0154 & 0.3226 & 0.0461 & -0.0924 & 0.0444 & 0.062 \\ 0.062 & -0.0444 & -0.0924 & -0.0461 & 0.3226 & 0.0154 & 0.0286 & 0.0462 \\ -0.0444 & -0.062 & -0.0461 & 0.924 & -0.0154 & 0.3226 & -0.0462 & 0.0286 \\ -0.0924 & 0.0461 & -0.062 & -0.0444 & -0.0286 & 0.0462 & 0.3226 & -0.0154 \\ -0.0461 & -0.0924 & 0.0444 & -0.062 & -0.0462 & -0.0286 & 0.0154 & 0.3226 \end{bmatrix}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Extrapolated rate for day 9 (February 15th) using MHR method with 8 nodes and weighted mean value (29). Source: author, 2014



$$M_2 \cdot \begin{bmatrix} 3 \\ 5 \\ 7 \\ 9 \\ 11 \\ 13 \\ 15 \\ 17 \end{bmatrix} = \begin{bmatrix} 3.7252 \\ 3.8072 \\ 3.8704 \\ 3.8278 \\ 3.8653 \\ 3.88834 \\ 3.9825 \\ 3.9882 \end{bmatrix},$$

HR extrapolation gives the result (17;3.9882). Anticipated rate of euro for the day February 15th is 3.9882 (see Figure 6).

MHR extrapolation has been done for three times ($N = 2$, 4 or 8) and anticipated values are 3.9158, 4.0767 and 3.9882 respectively. The proposal final anticipated rate of euro for the day February 15th (see Figure 7) based on weighted mean value:

$$\frac{4 \cdot 3.9158 + 2 \cdot 4.0767 + 3.9882}{7} = 3.9721 \tag{30}$$

because the rate 3.9158 is calculated with last four data points, 4.0767 is extrapolated for last eight data points and 3.9882 is computed for last sixteen data points. Formula (30) takes one fact into account: number of sixteen points is four times bigger than four and two times bigger than eigth. The result 3.9158 has to be strengthen multiplying by four and the rate 4.0767 has to be strengthen multiplying by two.

The true rate of euro for the day February 15th is 3.9398 (see Figure 8).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. Extrapolated rate for day 17 (February 15th) using MHR method with 16 nodes
Source: author, 2014

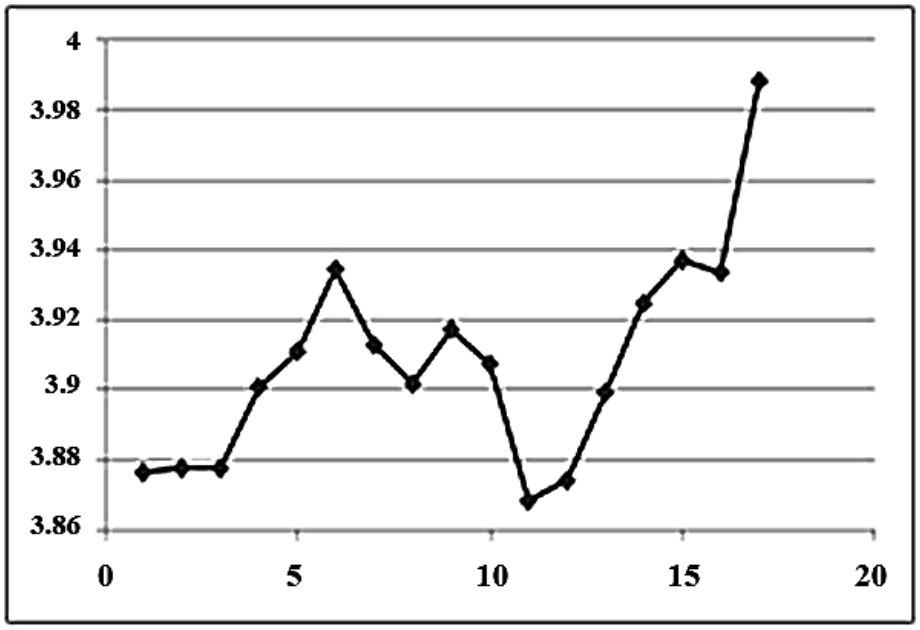
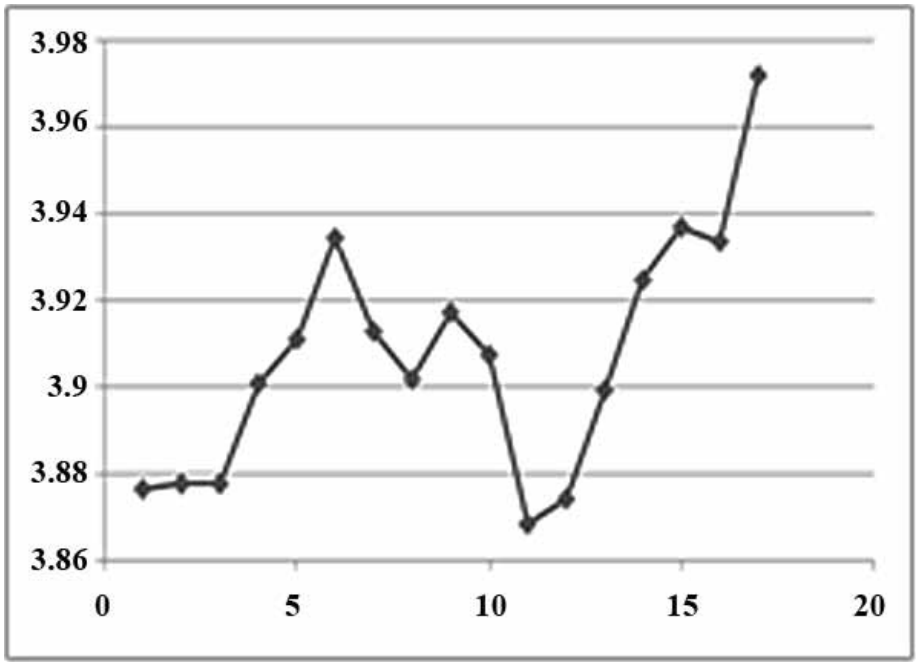
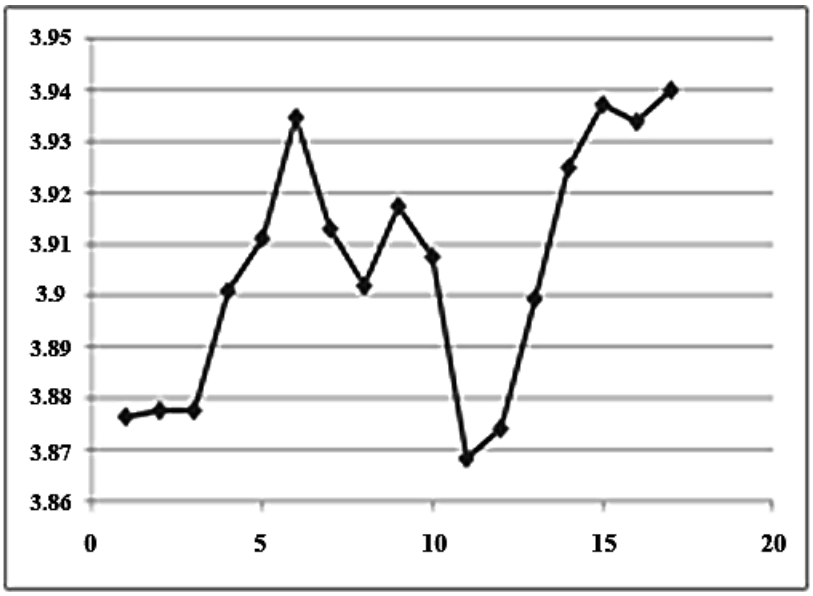


Figure 7. Extrapolated rate for day 17 (February 15th) using MHR method with 16 nodes and weighted mean value (30). Source: author, 2014



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. The true rate of euro for day 17 (February 15th). Source: author, 2014



In author's opinion, values extrapolated for next day 3.9694 (13) and 3.9721 (14) are good enough to be one of the factors for making a decision of buying or selling the currency.

Example 2

MHR calculations are done for true rates of US dollar at National Bank of Poland (NBP) from June 16th to July 8th, 2011 (Friday). If last four rates are considered: (1;2.7266), (2;2.7531), (3;2.7597) and (4;2.7505), MHR extrapolation with matrices of dimension $N = 2$ gives the result (5;2.7239):

$$M_2 \begin{bmatrix} 0.5503 & 9.13 \cdot 10^{-3} \\ 9.13 \cdot 10^{-3} & 0.5503 \end{bmatrix} \quad M_2 \cdot \begin{bmatrix} 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 1.6964 \\ 2.7239 \end{bmatrix}$$

So anticipated rate of US dollar on the day July 11th (Monday) is 2.7239.

If last eight rates are considered: (1;2.7877), (2;2.7517), (3;2.7273), (4;2.7156), (5;2.7266), (6; 2.7531), (7; 2.7597) and (8; 2.7505), MHR extrapolation with matrices of dimension $N = 4$ gives the result (9;2.8471):

$$M_2 = \begin{bmatrix} 0.3919 & -0.0032 & 0.0999 & 0.0527 \\ 0.0032 & 0.3919 & -0.0527 & 0.0999 \\ -0.0999 & 0.0527 & 0.3919 & 0.0032 \\ -0.0527 & -0.0999 & -0.0032 & 0.3919 \end{bmatrix} \quad M_2 \cdot \begin{bmatrix} 3 \\ 5 \\ 7 \\ 9 \end{bmatrix} = \begin{bmatrix} 2.3333 \\ 2.4999 \\ 2.7365 \\ 2.8471 \end{bmatrix}$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Anticipated rate of US dollar on the day July 11th is 2.8471. There are two extrapolated values for next day. Example 2 gives us two anticipated rates for tomorrow: 2.7239 and 2.8471. How these extrapolated values can be used in the process of decision making: to buy dollar or not, to sell dollar or not? The proposal final anticipated rate of US dollar (see Figure 9) on the day July 11th based on weighted mean value:

$$\frac{2 \cdot 2.7239 + 2.8471}{3} = 2.7650 \tag{31}$$

because the rate 2.7239 is calculated for $N = 2$, whereas 2.8471 is extrapolated for $N = 4$.

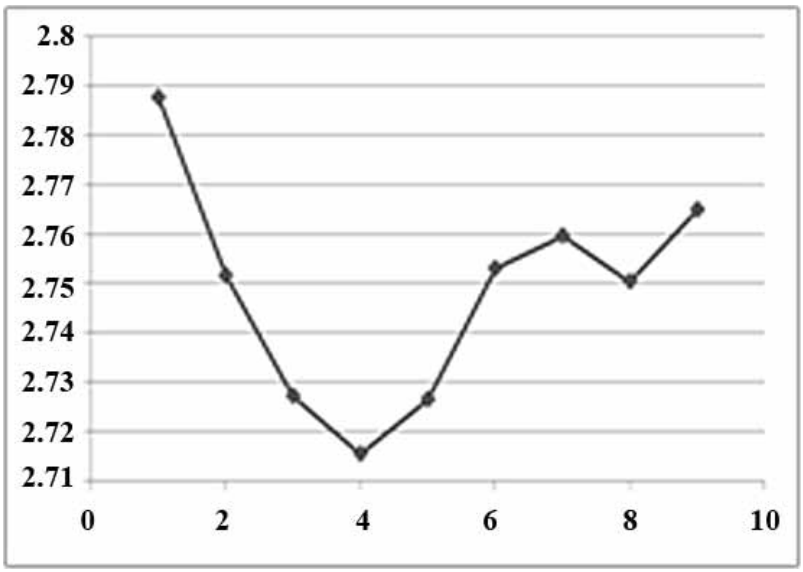
If last sixteen rates are considered, MHR extrapolation with matrices of dimension $N = 8$ has to be used. Here are the rates: (1;2.8069), (2;2.8077), (3;2.8058), (4;2.7776), (5;2.7661), (6;2.7914), (7;2.8201), (8;2.8055), (9;2.7877), (10;2.7517), (11;2.7273), (12;2.7156), (13;2.7266), (14;2.7531), (15;2.7597) and (16;2.7505). MHR extrapolation gives the result (17;2.7808). Anticipated rate of US dollar on the day July 11th is 2.7808.

MHR extrapolation has been done for three times ($N = 2$, 4 or 8) and anticipated values are 2.7239, 2.8471 and 2.7808 respectively. The proposal final anticipated rate of US dollar (see Figure 10) on the day July 11th based on weighted mean value:

$$\frac{4 \cdot 2.7239 + 2 \cdot 2.8471 + 2.7808}{7} = 2.7672 \tag{32}$$

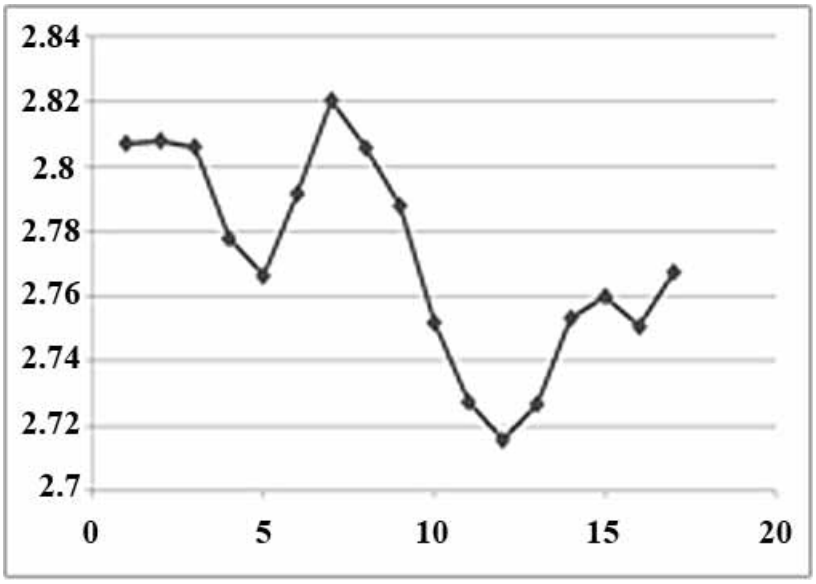
because the rate 2.7239 is calculated with last four data points, 2.8471 is extrapolated for last eight data points and 2.7808 is computed for last sixteen data points.

Figure 9. Extrapolated rate for day 9 (July 11th) using MHR method with 8 nodes and weighted mean value (31). Source: author, 2014



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

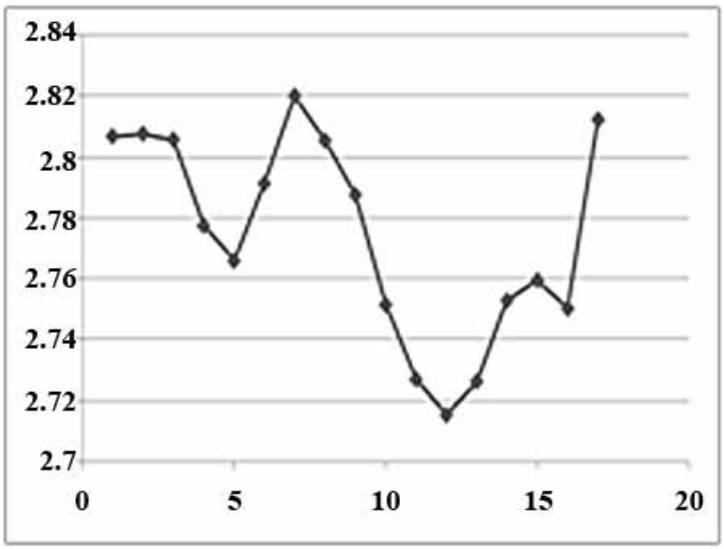
Figure 10. Extrapolated rate for day 17 (July 11th) using MHR method with 16 nodes and weighted mean value (32). Source: author, 2014



The true rate of US dollar on the day July 11th is 2.8123 (see Figure 11).

In author's opinion, extrapolated values 2.7650 (15) and 2.7672 (16) and anticipated rates in example 1 preserve the increasing trend and they are good enough to be one of the factors for making a decision of buying or selling the currency. Anticipated values, calculated by MHR method, are applied in the process of risk analysis and decision making: to follow the action or not, to do one thing or another. Extrapolated values can be used to make a decision in many branches of science and economics.

Figure 11. The true rate of US dollar for day 17 (July 11th). Source: author, 2014



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

DISCUSSION OF EXAMPLES

This chapter goes to first step of application for MHR method in point extrapolation and decision making. Figures and calculations show the trends of value: increasing or decreasing. This information could be used in analyzing risk and decision making, for example buying or selling when risk analysis is very difficult and very important.

The Length Estimation

Selection of the nodes is a key factor in the process of interpolation and extrapolation. The length of a curve is significant feature. Also the length estimation depends on the nodes. Having nodes (x_1,y_1), $(x_2,y_2),\ldots,(x_n,y_n)$ in MHR method (algorithm 1), it is possible to compute as many curve points as we want for any parameter $\alpha \in [0;1]$.

Assume that L is the number of reconstructed points plus n nodes. So curve consists of L points that could be indexed $(x_1',y_1'),(x_2',y_2'),\ldots,(x_L',y_L')$, where $(x_1',y_1') = (x_1,y_1)$ and $(x_L',y_L') = (x_n,y_n)$.

The length of curve, consists of L points, is estimated:

$$d(L) = \sum_{i=1}^{L-1} \sqrt{(x_{i+1}' - x_i')^2 + (y_{i+1}' - y_i')^2} \ . \tag{33}$$

For any accuracy of length estimation $\varepsilon > 0$, it is possible to use MHR method (algorithm 1) with suitable number and location of n nodes and reconstruct curve consists of L and L_1 points, where $|d(L) - d(L_1)| < \varepsilon$. So there is no need to compute

$$\int_{x_1}^{x_n} \sqrt{1+[f'(x)]^2}\,dx \tag{34}$$

as the length of curve f. Formula (33) has lower computational cost then (34).

SOLUTIONS AND RECOMMENDATIONS

The chapter deals with the problem of data interpolation and anticipation. None of the methods, that are used nowadays, applies the orthogonal basis for unknown value extrapolation and foreseeing, but proposed MHR method does as the solution. Orthogonality means a very important feature for stability in calculations. Also MHR point interpolation and extrapolation via simple computations with matrices is something new in the problem of anticipation, risk analysis and decision making. Extrapolated values show increasing or decreasing trend during the process of analyzing risk and decision making. MHR method as the solution is recommended in artificial intelligence and operational research, for example in scheduling, time-tabling or planning.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

FUTURE RESEARCH DIRECTIONS

Future works with MHR method are connected with the formula of extrapolated values. Also estimation of object area in the plane, using nodes of object contour, will be possible by MHR interpolation. Length of the curve and object area are significant features in many economic problems. Future works are dealing with smoothing the curve, parameterization of whole curve and possibility to apply MHR method to three-dimensional curves. Also case of equidistance nodes must be studied with all details. Another future research direction is to apply MHR method in artificial intelligence and operational research, for example scheduling, time-tabling, planning, identification of a specific person's face or fingerprint, character recognition (Lamdan, Schwartz & Wolfson, 1990) or image restoration (Pope & Lowe, 2004). There are several specialized tasks based on recognition to consider and it is important to use the shape of whole contour (Kriegman & Ponce, 1990) for identification and detection of persons, vehicles or other objects. Other applications of MHR method will be directed to computer graphics, modeling and image processing.

CONCLUSION

The method of Hurwitz-Radon Matrices leads to curve interpolation and value extrapolation depending on the number and location of data points. No characteristic features of curve are important in MHR method: failing to be differentiable at any point, the Runge's phenomenon or differences from the shape of polynomials. These features are very significant for classical polynomial interpolations and extrapolations. MHR method gives the possibility of reconstruction a curve and anticipation the data points. The only condition is to have a set of nodes according to assumptions in MHR method. Data representation and curve extrapolation by MHR method is connected with possibility of changing the nodes coordinates and reconstruction of new data or curve for new set of nodes. The same MHR interpolation and extrapolation is valid for discrete and continues data. Main features of MHR method are: accuracy of data reconstruction depending on number of nodes; interpolation or extrapolation of a curve consists of L points is connected with the computational cost of rank $O(L)$; MHR method is dealing with local operators: average OHR operators are built by successive 4, 8 or 16 data points, what is connected with smaller computational costs then using all nodes; MHR is not an affine interpolation.

REFERENCES

Brachman, R. J., & Levesque, H. J. (2004). *Knowledge representation and reasoning*. San Francisco, CA: Morgan Kaufman.

Cierniak, R. (2005). *Computed tomography*. Warsaw, Poland: Exit.

Citko, W., Jakóbczak, D., & Sieńko, W. (2005, September). *On Hurwitz - Radon matrices based signal processing*. Paper presented at the workshop Signal Processing at Poznan University of Technology, Poznań, Poland.

Dahlquist, G., & Bjoerck, A. (1974). *Numerical methods*. Englewood Cliffs, NJ: Prentice Hall.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Eckmann, B. (1999). Topology, algebra, analysis- relations and missing links. *Notices of the American Mathematical Society*, *5*(46), 520–527.

Fagin, R., Halpern, J. Y., Moses, Y., & Vardi, M. Y. (1995). *Reasoning about knowledge*. Cambridge, MA: MIT Press.

Jakóbczak, D. (2006). *Application of discrete, orthogonal operator of Hurwitz - Radon in compression and reconstruction of monochromatic images' contours.* (Unpublished doctoral dissertation). Polish - Japanese Institute of Information Technology, Warsaw, Poland.

Jakóbczak, D. (2007). 2D and 3D image modeling using Hurwitz - Radon matrices. *Polish Journal of Environmental Studies*, *4A*(16), 104–107.

Jakóbczak, D., & Kosiński, W. (2007). Hurwitz - Radon operator in monochromatic medical image reconstruction. *Journal of Medical Informatics & Technologies*, *11*, 69–78.

Jakóbczak, D., & Kosiński, W. (2007). Application of Hurwitz - Radon matrices in monochromatic medical images decompression. In Z. Kowalczuk & B. Wiszniewski (Eds.), *Intelligent data mining in diagnostic purposes: Automatics and informatics* (pp. 389–398). Gdansk, Poland: PWNT.

Kozera, R. (2004). *Curve modeling via interpolation based on multidimensional reduced data*. Gliwice, Poland: Silesian University of Technology Press.

Kriegman, D. J., & Ponce, J. (1990). On recognizing and positioning curved 3-D objects from image contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *12*(12), 1127–1137. doi:10.1109/34.62602

Lamdan, Y., Schwartz, J. T., & Wolfson, H. J. (1990). Affine invariant model-based object recognition. *IEEE Transactions on Robotics and Automation*, *5*(6), 578–589. doi:10.1109/70.62047

Latecki, L. J., & Lakaemper, R. (1999). Convexity rule for shape decomposition based on Discrete Contour Evolution. *Computer Vision and Image Understanding*, *3*(73), 441–454. doi:10.1006/cviu.1998.0738

Lowe, D. G. (1991). Fitting parameterized three-dimensional models to images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *5*(13), 441–450. doi:10.1109/34.134043

Lowe, D. G. (1999, September). *Object recognition from local scale-invariant features*. Paper presented at the International Conference on Computer Vision, Corfu, Greece.

Lowe, D. G. (2001). *Local feature view clustering for 3D object recognition*. Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition, Kauai, HI. 10.1109/CVPR.2001.990541

Markman, A. B. (1998). *Knowledge representation*. Mahwah, NJ: Lawrence Erlbaum Associates.

Pope, A. R., & Lowe, D. G. (2004). Probabilistic models of appearance for 3-D object recognition. *International Journal of Computer Vision*, *2*(40), 149–167.

Ralston, A. (1965). *A first course in numerical analysis*. New York: McGraw-Hill Book Company.

Sieńko, W., & Citko, W. (2002). *Hamiltonian Neural Net based signal processing.* Paper presented at the International Conference on Signal and Electronic System ICSES, Wrocław, Poland.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sieńko, W., Citko, W., & Jakóbczak, D. (2004). Learning and system modeling via Hamiltonian Neural Networks. In L. Rutkowski, J. Siekmann, R. Tadeusiewicz, & A. Zadeh (Eds.), *Lecture notes on artificial intelligence: Artificial intelligence and soft computing - ICAISC 2004* (pp. 266–271). Berlin: Springer - Verlag. doi:10.1007/978-3-540-24844-6_36

Sieńko, W., Citko, W., & Wilamowski, B. (2002). *Hamiltonian Neural Nets as a universal signal processor*. Paper presented at the 28th Annual Conference of the IEEE Industrial Electronics Society IECON, Sevilla, Spain. 10.1109/IECON.2002.1182910

Soussen, C., & Mohammad-Djafari, A. (2004). Polygonal and polyhedral contour reconstruction in computed tomography. *IEEE Transactions on Image Processing*, *11*(13), 1507–1523. doi:10.1109/TIP.2004.836159 PMID:15540458

Sowa, J. F. (2000). *Knowledge representation: logical, philosophical and computational foundations*. New York: Brooks/Cole.

Straffin, P. D. (1993). *Game theory and strategy*. Washington, DC: Mathematical Association of America.

Tang, K. (2005). Geometric optimization algorithms in manufacturing. *Computer-Aided Design and Applications*, *2*(6), 747–757. doi:10.1080/16864360.2005.10738338

Tarokh, V., Jafarkhani, H., & Calderbank, R. (1999). Space-Time Block Codes from orthogonal designs. *IEEE Transactions on Information Theory*, *5*(45), 1456–1467. doi:10.1109/18.771146

Ullman, S., & Basri, R. (1991). Recognition by linear combinations of models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *10*(13), 992–1006. doi:10.1109/34.99234

Watson, J. (2002). *Strategy – an introduction to game theory*. San Diego, CA: University of California Press.

ADDITIONAL READING

Basu, S., & Bresler, Y. (2000). $O(N^2 \log_2 N)$ filtered backprojection reconstruction algorithm for tomography. *IEEE Transactions on Image Processing*, *9*(10), 1760–1773. doi:10.1109/83.869187 PMID:18262914

Brankov, J. G., Yang, Y., & Wernick, M. N. (2004). Tomographic image reconstruction based on a Content – Adaptive Mesh Model. *IEEE Transactions on Medical Imaging*, *2*(23), 202–212. doi:10.1109/TMI.2003.822822 PMID:14964565

Brasse, D., & Defrise, M. (2004). Fast fully 3-D image reconstruction in PET using planograms. *IEEE Transactions on Medical Imaging*, *4*(23), 413–425. doi:10.1109/TMI.2004.824231 PMID:15084067

Cetin, M., Karl, W. C., & Willsky, A. S. (2002, September). *Edge – preserving image reconstruction for coherent imaging application.* Paper presented at the IEEE International Conference on Image Processsing, Rochester, NY, USA. 10.1109/ICIP.2002.1039992

Chlebus, E., & Cholewa, M. (1999). Rapid prototyping – rapid tooling. *CADCAM Forum, 11*, 23-28.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Cormen, T. H., Leiserson, C. E., & Rivest, R. L. (1996). *Introduction to algorithms*. Massachusetts, USA: the Massachusetts Institute of Technology Press and McGraw-Hill.

Defrise, M. (2001). A short reader's guide to 3D tomographic reconstruction. *Computerized Medical Imaging and Graphics*, *25*(2), 113–116. doi:10.1016/S0895-6111(00)00061-6 PMID:11137787

Dryja, M., Jankowska, J., & Jankowski, M. (1982). *Survey of numerical methods and algorithms. Part II*. Warsaw, Poland: WNT.

Eldar, Y. C. (2001). *Quantum Signal Processing.* (Unpublished doctoral dissertation). Massachusetts Institute of Technology, USA.

Eldar, Y. C., & Oppenheim, A. V. (2002). Quantum Signal Processing. *IEEE Signal Processing Magazine*, *6*(19), 12–32. doi:10.1109/MSP.2002.1043298

Fortuna, Z., Macukow, B., & Wąsowski, J. (1982). *Numerical methods*. Warsaw, Poland: WNT.

Jakóbczak, D. (2005). Hurwitz - Radon matrices and their children. *Computer Science*, *5*(8), 29–38.

Jankowska, J., & Jankowski, M. (1981). *Survey of numerical methods and algorithms. Part I*. Warsaw, Poland: WNT.

Kontaxakis, G., & Strauss, L. G. (1998). Maximum likelihood algorithms for image reconstruction in Positron Emission Tomography. *Radionuclides for Oncology – Current Status and Future Aspects, 1998,* 73-106.

Kowalczuk, Z., & Wiszniewski, B. (Eds.). (2007). *Intelligent data mining in diagnostic purposes: Automatics and informatics*. Gdansk, Poland: PWNT.

Kundur, D., & Hatzinakos, D. (1998). A novel blind deconvolution scheme for image restoration using recursive filtering. *IEEE Transactions on Signal Processing*, *2*(46), 375–390. doi:10.1109/78.655423

Laine, A., & Zong, X. (1996). *Border identification of echocardiograms via multiscale edge detection and shape modeling*. Paper presented at the IEEE International Conference on Image Processsing, Lausanne, Switzerland. 10.1109/ICIP.1996.560486

Lang, S. (1970). *Algebra*. Reading, MA: Addison-Wesley Publishing Company.

Le Buhan Jordan, C., Bossen, F., & Ebrahimi, T. (1997). *Scalable shape representation for content based visual data compression*. Paper presented at the International Conference on Image Processing, Santa Barbara, CA, USA. 10.1109/ICIP.1997.647962

Marker, J., Braude, I., Museth, K., & Breen, D. (2006). Contour-based surface reconstruction using implicit curve fitting, and distance field filtering and interpolation. *Volume Graphics*, *2006*, 1–9.

Meyer, Y. (1993). *Wavelets: algorithms & applications*. Philadelphia: Society for Industrial and Applied Mathematics.

Poggio, T., & Smale, S. (2003). The mathematics of learning: Dealing with data. *Notices of the American Mathematical Society*, *5*(50), 537–544.

Przelaskowski, A. (2005). *Data compression*. Warsaw, Poland: BTC.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Rutkowski, L., Siekmann, J., Tadeusiewicz, R., & Zadeh, A. (Eds.). (2004). *Lecture notes on artificial intelligence: Artificial intelligence and soft computing*. Berlin - Heidelberg, Germany: Springer - Verlag.

Vakhania, N. (1993). Orthogonal random vectors and the Hurwitz – Radon - Eckmann theorem. *Proc. of the Georgian Academy of Sciences - Mathematics, 1(1)*, 109-125.

Willis, M. (2000). *Algebraic reconstruction algorithms for remote sensing image enhancement*. Unpublished doctoral dissertation, Department of Electrical and Computer Engineering, Brigham Young University.

Xu, Fang, & Mueller, K. (2005). Accelerating popular tomographic reconstruction algorithms on commodity PC graphics hardware. *IEEE Transactions on Nuclear Science*, *3*(52), 654–661.

Zaletelj, J., & Tasic, J. F. (2003). *Optimization and tracking of polygon vertices for shape coding*. Berlin - Heidelberg, Germany: Springer - Verlag. doi:10.1007/978-3-540-45179-2_52

Zhang, J. K., Davidson, T., & Wong, K. M. (2004). Efficient design of orthonormal wavelet bases for signal representation. *IEEE Transactions on Signal Processing*, *7*(52), 1983–1996. doi:10.1109/TSP.2004.828923

KEY TERMS AND DEFINITIONS

Artificial Intelligence: Intelligence of machines and computers, as a connection of algorithms and hardware, which makes that a man – human being can be simulated by the machines in analyzing risk, decision making, reasoning, knowledge, planning, learning, communication, perception and the ability to move and manipulate objects.

Contour Modeling: Calculation of unknown points of the object contour having information about some points of the object contour.

Curve Interpolation: Computing new and unknown points of a curve and creating a graph of a curve using existing data points – interpolation nodes.

Data Extrapolation: Calculation of unknown values for the points situated outside the ranges of nodes.

Hurwitz – Radon Matrices: A family of skew – symmetric and orthogonal matrices with columns and rows that create, together with identical matrix, the base in vector spaces of dimensions $N = 2, 4$ or 8.

MHR Method: The method of curve interpolation and extrapolation using linear (convex) combinations of OHR operators.

OHR Operator: Matrix operator of Hurwitz – Radon built from coordinates of interpolation nodes.

Value Anticipation: Foreseeing next value when last value is known.

This research was previously published in Analyzing Risk through Probabilistic Modeling in Operations Research edited by Dariusz Jacek Jakóbczak; pages 1-31; copyright year 2016 by Business Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 68
Training Infrastructure to Participate in Real Life Institutions: Learning through Virtual Worlds

Pablo Almajano
University of Barcelona, Spain

Anna Puig
University of Barcelona, Spain

Maite Lopez-Sanchez
University of Barcelona, Spain

Maria Salamó Llorente
University of Barcelona, Spain

Inmaculada Rodriguez
University of Barcelona, Spain

Mireia Ribera
University of Barcelona, Spain

ABSTRACT

This chapter presents an example of a virtual training scenario that allows trainees to learn the rules of real human institutions before joining them, that is, before having to cope with the consequences of their binding actions in real (serious) settings. The development of the training scenario is based on a combination of Electronic Institutions technology with 3D Virtual Worlds, and is enhanced with tutoring agents that provide assistance services along the training process. The theoretical benefits that result from this approach are demonstrated through an authentic scenario that represents a market (i.e., an institution) devoted to trading water rights. This scenario has been tested with real users, and thus, its benefits have been empirically assessed.

INTRODUCTION

A real-life institution is a rather complex and structured environment where users participate with the aim of developing serious activities, such as the trading of goods. Therein, users interact by following well-defined communication protocols and playing different roles: staff roles are in charge of support-

DOI: 10.4018/978-1-7998-0951-7.ch068

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ing specific institutional tasks (i.e. register goods to trade or auction the registered goods), and external roles enter the institution to perform such tasks. Nevertheless, in these complex real-world scenarios, users may not know how to achieve their goals. Furthermore, users' actions usually have non-reversible and important consequences, such as, for example, the acquisition of legal obligations.

An approach to address users' participation in these intricate scenarios is to first immerse them in a "virtual" institution that reproduces the "real" one so that they can be trained to participate in the "real" one and to get familiar with its structure and functioning. Then, before their involvement in the "real" institution, users may learn by experiencing in a simulated environment, and this training would allow them to perform effectively and with confidence in the real situation.

In this context of adult learning, Kolb (Kolb, 1983, 2014) proposed a model that defines 4 experiential learning stages: concrete experience, reflection, abstract conceptualisation and active experimentation. That is, the learner lives a concrete experience, followed by a reflection on that experience. Then, the abstract conceptualisation allows either the application of known theories to the lived experience or the derivation of rules describing it. Finally, in the active experimentation the learner constructs a way of modifying the next occurrence of the experience, leading to a better performance.

These learning stages are usually leaded by the instructor and are experienced by the learner either inside or outside the class. On the one hand, typical in-person classes rely on the instructor to ensure the learner gathers and understands the needed information to reflect about, so that she is capable of creating an abstract conceptualization and to apply the gained understanding to the next experience. On the other hand, online (virtual) classes rely on a User Interface (UI) with the learner as the medium to go through the learning stages, with the timely support of teachers and workmates. This UI can be a web-based e-learning system such as moodle, or a 3D computer-generated simulation of an environment where the learner lives the "entire" learning experience. For example, a 3D Virtual World includes spaces for learning, tools to manipulate, virtual tutors to talk with, and other learners to interact with.

Some UIs consist in e-learning platforms, such as Moodle, where the last stage, active experimentation, occurs outside the e-learning space, that is in the learners' "real life"; whereas 3D Virtual Worlds (3D VWs) offer the possibility to do the active experimentation in the e-learning environment as well. Additionally to tutors and other learners, common to many e-learning spaces, 3D VWs offer spaces for learning, tools to manipulate, and virtual tutors. Moreover, virtual worlds grant the possibility to make an activity more than once at no cost and no risk, and they also result more attractive and ideal for social and collaborative activities.

Based on these benefits of 3D interfaces, this research sees 3D VWs as an adequate mean to instantiate this experiential learning cycle. That is, users can learn to participate in a real structured (complex) activity by acting themselves in a simulated scenario, experiencing the consequences of their actions, testing procedures and checking their understanding without harm (Hew & Cheung, 2010) nor consequences in the real world. As previous researches stated, virtual worlds could be much more appealing and engaging than traditional learning (Wrzesien & Alcañiz, 2010; Craven, 2015) . Hence, this chapter proposes to train real institutions' users by letting them experience their assigned duties within a 3D simulated space, fostering collaboration and engagement, and facilitating users' spatial situatedness and interactions.

Specifically, we present a Hybrid Structured 3D Virtual Environment where participants (both human users and software agents) develop, and are trained to develop, serious activities in rather complex institutions. To do so, we combine (i) an Organisation-Centred Multi-Agent System (OCMAS), to model the institution and regulate participants' interactions and (ii) a 3D training virtual environment, to facilitate learners interactions and engage them in the system.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Additionally, our training environment provides users with Personal Assistants devoted to help them during the learning process, and consequently, we named the entire proposal as Assisted Hybrid Structured 3D Virtual Environment. Personal Assistants are embodied tutoring agents who provide the users with different assistance services: information, justification, advice and estimation (Almajano, 2014; Almajano, López-Sánchez, Rodríguez & Trescak., 2014). They engage in conversation with learners to guide them in the consecution of their goals. These agents use an extension of AIML (Artificial Intelligence Modelling Language) for task-oriented conversations and let users rehearse in the system.

As a proof of concept, we present a training scenario of a virtual market for trading water rights: v-mWater. This scenario consists of several training activities reproducing those that users can perform in an actual water market (e.g., to register water rights or to auction them). During these activities, users interact with each other and staff agents, who are tutors in charge of assisting them.

The rest of this chapter is organized as follows. Firstly, a Background Section briefly presents the state of the art that contextualizes our work and highlights its potential benefits to training scenarios. Secondly, a System Model Overview Section describes the main components of our Assisted Hybrid Structured 3D Virtual Environment and how they relate to each other. Thirdly, the specific virtual water market scenario is introduced in the Training Scenario Section. Subsequent Technical Specification Section introduces the technical details related to the execution infrastructure and the assistance that tutors provide. Afterwards, Evaluation Section provides details on users testing. Next Section, entitled Discussion, is devoted to explain the limitations of the system and to present new research lines that can extend the system in several aspects, such as the incorporation of recommender systems, the inclusion of accessibility support, and the incorporation of gamification concepts to improve the learning experience. Finally, Conclusion Section presents some concluding remarks.

BACKGROUND

3D virtual worlds technologies were introduced by the seminal paper of Bartle(Bartle, 2003), and have been used in education since long ago. One of their strengths is their easiness of social interaction (San Chee, 2001), which can be done through chat, audio or even gestures. In particular audio and chat have proven to be particularly useful (Bouras & Tsiatsos, 2006). Nevertheless, single-user e-learning systems do not motivate users to explore thoroughly the possibilities of interaction (Monahan, McAdrle & Bertolotto, 2008)while teamwork help learners to reach higher achievements (Laister & Kober, 2002) as they feel as part of a supportive e-learning community and that fosters their motivation (McInnerney & Roberts, 2004). On the mentioned studies, on the overall, teachers and students find virtual environments intuitive and support the usefulness of 3D paradigm.

On the use of Virtual Worlds for training Dalgarno and Lee, in a seminal article on 2009 (2009) describe the learning affordances of 3D Virtual worlds as "enhanced spatial knowledge representation, greater opportunities for experiential learning, increased motivation/engagement, improved contextualisation of learning and richer/more effective collaborative learning" as compared to 2D counterparts.

Mostly, the literature agrees in the potentiality of 3D virtual worlds but it alerts about how to use them and the need to provide context and guidance to learners (Dickey, 2010)

Duncan, Miller and Jiang (2012) published the most general survey of the use of virtual worlds in education, and as a conclusion they recommend "the use of virtual support where real world experiential learning is difficult to achieve due to barriers of time, cost and place" although they identify some

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

disadvantages as the difficulty to monitor the educational process, the fact that students need more guidance, or their disadvantage in front of simulation in the real world and. A more recent literature survey, by Ghanbarzadeh et al. (2014), highlight the use of virtual worlds for both academic and professional training, and praise them as a natural evolution of Internet, distance learning and online gaming. The authors consider 3D virtual world "a demand of the new generation of students".

Recent articles show contradictory results. For example Hack research (2015) quantifies a decrease of failure in a distance learning module to health professionals from 11% to 5% thanks to the introduction of 3D virtual world training. On the contrary Richards and Taylor (2015), on an experiment with 129 biologists, found better learning gains in a 2D environment than in 3D virtual worlds.

It seems that 3D virtual worlds are justified in particularly costly and dangerous training scenarios, like police training (Moskaliuk, Bertram, Cress 2013), military, health or auctions as in the water-world case study.Also there is coincidence of a greater cognitive load in 3D scenarios. We think that the introduction of assistive agents, and natural language communication included in our proposal will be key for reducing both the learning curve and the learners' demands.A paradigmatic example of interactive e-learning is the city of Uruk, a simulation of the first city on earth, dating year 3000 BC (Bogdanovych, Rodríguez-Aguilar, Simoff, Cohen, 2010) (Bogdanovych, Kiran, Simoff, 2012) based on Second Life, where learners could interact with recreations of city buildings from Ziggurat or with the Uruk temple and also with their inhabitants. Posterior examples include medical education (Danforth, Procter, Chen, Johnson & Heller, 2009), or physics (Graesser, Chipman, Haynes & Olney, 2005) training where artificial dialogue agents were introduced.

Regarding personal assistants, Okamoto, Sycara and Scerri (2009) tested that agents' decision support to human users in hierarchical organizations improved success rate and efficiency. Many Multi-Agent System architectures have been proposed to provide assistance (Chalupsky et al., 2001; Yaich, Boissier, Picard & Jaillon, 2013; Dong, McCarthy, O'Mahony, Schaal & Smyth, 2012; Oh, Meneguzzi, Sycara & Norman, 2011), but they work in a restricted context where humans have limited and very specific tasks. In this chapter we offer a model for assistance in complex tasks related to Electronic Institutions[1], as defined by Esteva (Esteva, 2003).

Related to trainers-trainees interactions in virtual environments, natural language interactions constitute a suitable alternative to instructional/command based interactions. Embodied conversational agents, able to engage in a conversation with humans, are popular representatives of this type of interactions, which can be mostly found as virtual assistants in web environments. Specifically, Artificial Intelligence Mark-up Language (AIML) chatter bots ("Alicebot", n.d.) (Wallace, 2009) are well known reactive bots which follow basic dialogue structures defined in static files, but are able to store some general information (such as name or gender of the user) in a non-typed memory (i.e. a list of string values) and act accordingly. They have been conceived to give general information to users under request.

Finally, as for accessibility, i.e. the usability for people with disabilities, education presents a new challenge which is the enforced need to make any tool or material, virtual worlds included, valid for all profiles - blind, hard of hearing, with dyslexia, with motor impairments... - (Del Blanco, Marchiori, Torrente, Martínez-Ortiz & Fernández-Marijón, 2013; Torrente, Freire, Moreno-Ger & Fernández-Manjón, 2015). Unfortunately, although therapy benefits have been elicited about videogames (Vickers, Istance & Hyrskykari, 2013; Fikar, Schönauer & Kaufmann, 2013; "Transcript", n.d.), they are considered the less accessible multimedia platform. In fact, current development platforms do not facilitate the creation of universally designed videogames (Grammenos, Savidis & Stephanidis, 2007)and there still exists a wide gap in game accessibility between research and industry that has been repeatedly noted (Westin,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bierre, Gramenos & Hinn, 2011; Yuan, Folmer & Harris, 2010; Porter, 2014). In our model, the universal education challenge is not yet addressed, but the assistance layer is a promising mechanism to support customized solutions without the requirement of specific development platforms.

Combining and enriching the benefits of previous proposals such as sociability, entertainment, easy to use and intuitive, and also the support of intelligent agents who communicate through natural language, our proposal has an inmense potential as a common tool for training in education or professional environements. With virtual worlds, electronic institutions and intelligent agents you can model many real situations and give the learners the freedom to explore with full support.

SYSTEM MODEL OVERVIEW

As previously introduced, this research advocates for the use of an *Assisted Hybrid Structured 3D Virtual Environment* (VE) to support the training of real institutions' users. This section describes the main components of the system. An Assisted Hybrid Structured 3D VE is a persistent multi-user system where participants (both human users and software agents) develop and learn to conduct activities. In this system, the 3D interface graphically represents the system and facilitates human participation, and an Organisation Centred Multi-Agent System (OCMAS) structures participants' interactions. To do so, the OCMAS specifies the roles that participants can enact, the activities where complex tasks can be accomplished, and the communication protocols that enable the prosecution of such tasks.

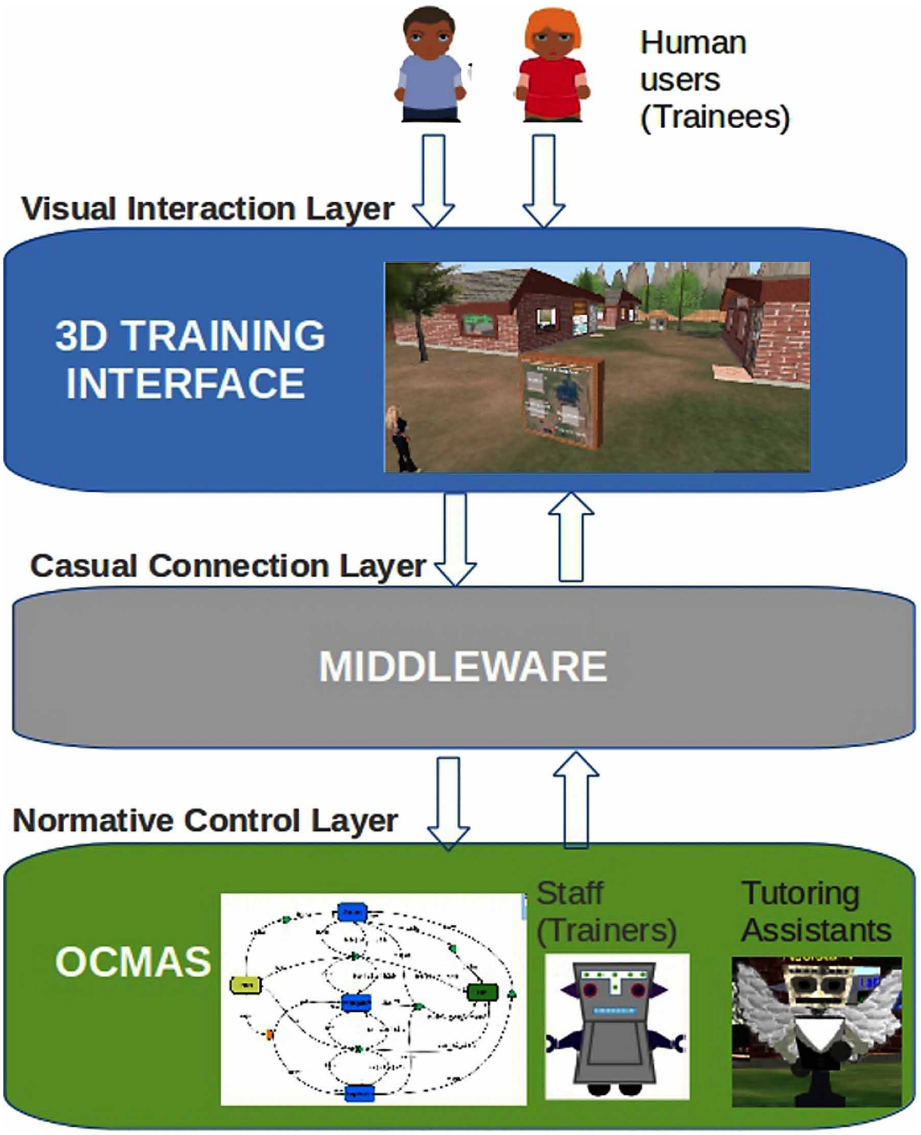
The system is, therefore, composed by three main components: the Normative Control Layer (OCMAS), the Visual Interaction Layer (3D interface), and a middleware to connect previous layers, which is the Causal Connection Layer (see Figure 1).

The division into a Normative Control Layer and a Visual Interaction Layer provides support to the conceptual separation between the normative control of interactions and the design of the virtual world, i.e., the design of the 3D interface. The Normative Control Layer is responsible for the institutional control of interactions among participants, while the Visual Interaction Layer focuses on the 3D representation of the institution. Both layers are causally connected. That means that, whenever one of them changes, the other one changes in order to maintain a consistent state (Maes & Nardi, 1988). Thus, the Causal Connection Layer keeps a consistent state between the model, represented by the Normative Control Layer, and its view, represented by the Visual Interaction Layer.

Regarding participants, humans (trainees) take part in the system by controlling an avatar on the Visual Interaction Layer. Software agents can be external, staff and assistant agents. External agents, likewise human participants, are connected to the system to perform activities on behalf of others. Staff agents are software agents devoted to support the institutional (system) activities. They act as trainers for human users. Tutoring Assistant agents are in charge of supporting human users in the consecution of their goals. These agents also act as trainers as they may guide users in their tasks. All software agents actually belong to the OCMAS within the Normative Control Layer; participate in the overall training system by exchanging messages; and are visualized as avatars in the Visual Interaction Layer in a way that can be distinctly identified. In this work, as depicted in the bottom part of Figure 1, we focus on those agents that support users during the training activities, that is, Staff and Tutoring Assistant agents.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. System components



TRAINING SCENARIO

This research presents as training scenario v-mWater, a virtual market based on trading water. This scenario can be used by governments with the objective of encouraging a more efficient use of surplus of water for irrigation, especially in those countries suffering from water scarcity problems. It is inspired by similar tools such as the one developed by Waterfind ("Waterfind...", n.d.).

In this market, participants negotiate water rights (Almajano, Trescak, Esteva, Rodríguez, López-Sánchez, 2012). In an agricultural context, a water right is associated to a farmland and refers to the right of an irrigator to use a given volume of water from a public water source (e.g., a river or a pond) to irrigate. An agreement is the result of a negotiation where a seller settles with a buyer to reallocate (part

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

of) the water from her or his water rights for a fixed period of time in exchange for a given amount of money. Specifically, sellers are required to first register their water rights for sale and buyers compete for acquiring water rights by participating in an auction. In the following section we introduce tutoring agents (Trainers) and training activities in v-mWater.

Tutoring Agents

As aforementioned, participants in our system can be both human users and software agents. Trainers are internal software agents devoted to train human users to participate in the real-life institution. Specifically, our system has available two types of trainers: i) staff agents and ii) tutoring assistants.

Staff agents are in charge of training users on how to participate in the system. To do so, they are programmed to perform automatic institutional tasks, following the communication protocols established in the system specification. Thus, human users learn how to behave in a real institution by interacting with staff agents in the simulated one. Concretely, v-mWater incorporates the market facilitator, who carries out water rights' registration and also acts as auctioneer; and the basin authority, in charge of officially announcing valid agreements (selling of water rights) in the market.

Tutoring personal assistants support trainees to understand the social model of the institution and to perform the complex tasks in a seamless way. One tutoring personal assistant provides direct and sole support to one trainee, who, can freely use the given support in its decision process. Since Personal Assistants are organisational agents, they are allowed to use organisational information in order to provide assistance.

Tutoring Personal assistants offer four types of services: information, justification, advice and estimation. First, information service provides trainees with processed information about the system specification and its execution states. Second, justification service, based on the system specification and its current state, provides the trainee with a justification related to the last action that she or he failed to execute. Third, estimation service, also based on the system specification and the current state, is able to estimate whether the action can be executed at current state or not and, in affirmative case, it also computes the next state after executing the given action. Fourth, advice service is the more sophisticated one, and provides trainees with a plan that leads to her or his goal.

Training Activities

There are three activities in the market that take place in three separate physical spaces accessible from the entrance (see Figure 2). First, Waiting and Information activity, depicted in Figure 3, represents a common area where irrigators (both sellers and buyers) can ask for information about the market progress to the market facilitator agent. Second, within the Registration activity detailed in Figure 4, the market facilitator staff agent is in charge of registering sellers' rights. In fact, the institution regulation establishes sellers to be the only external role that can access this activity, and thus doors in the virtual registration building do not open for buyers (i.e., they are not allowed to enter).

Finally, the last (but not least important) activity that is available to the users is the Auction, where the selling of water rights takes place (see Figure 5). Like in real settings, there are three roles involved in this activity: buyers, which correspond to the bidders; the market facilitator, which acts as the auctioneer; and the basin authority, in charge of officially announcing the valid agreements. There may be different Auction activities going on in the scenario, since the organisation creates one Auction scene

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. v-mWater activities



Figure 3. v-mWater Waiting & Information training room



Figure 4. v-mWater Registration training room



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. v-mWater Auction training room



for each area of the basin. Within each Auction activity, water rights are auctioned in consecutive rounds and only buyers with owned water rights belonging to the area of the auction are allowed to join it (i.e., the access is restricted for the rest of participants). Auctions follow a Japanese protocol, which is most suitable for trading water rights, but may require buyers to get familiar with it.

These three activities in the virtual world are populated and leaded by staff agents who support users' training. For example, an irrigator could connect to the training system to learn any of the activities that take place in the market. In what follows we describe the process by which an irrigator learns how to register water rights by experiencing (rehearsing) it. To do so, she first enters the 3D agricultural environment with a seller role and walks around the market area to be aware of both market activity locations and other participants connected to the system. Then, whenever she finds the Registration room, she goes into it and inspects the space, which is populated by a registration desk, some pictures of the basin, and a staff agent representing the registration manager. Then, the staff agent engages in a conversation with the human with the ultimate goal of performing a task. From now on we will refer to it as "task-oriented conversation".

The conversation proceeds as follows. The agent initiates the conversation by greeting the user, and then proceeds by asking her for all the data needed to register a water right. The user gives the conversational agent the required data. If the data is correct the conversation flows as defined in the OCMAS interaction protocol. If it is incorrect, the agent informs about the error, and waits for the new data. At the end of the process, the irrigator is informed about the result of the task, whether failed or succeed.

Furthermore, trainees can ask for help to their personal tutoring assistant agents. For example, let's imagine that Mary Smith arrives to the 3D virtual market and wants to register a water right 5 euros higher than the last transaction done in the market; nevertheless, she does not know where and how to be aware of the last transactions' price and how to register a water right; then, she asks the tutoring assistant agent for advice on how to do it and as a result, her personal assistant provides her with the following plan of actions to achieve her goal (Figure 6 depicts this advice in context):

Mary Smith, enter the Waiting And Information room.
Information about market transactions is publicly displayed at Information Panels. KEEP THE INFORMATION OBTAINED HERE (price) TO BE USED LATER.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. A tutoring assistant agent provides a user advice on how to achieve her goal



Mary Smith, exit the Waiting And Information room.
Mary Smith, enter the Registration room.
Mary Smith, ask for registering a water right to Registration Manager.

As we can see with this example Mary Smith will succeed in registering her water right at the desired price without previous knowledge of the rules and needed steps.

So far, this section has described a training scenario in the context of a water market. Nevertheless, it is worth noticing that, from the point of view of the user, this virtual scenario allows users to rehearse tasks, previous to the "real" tasks conducted in the real setting. Note also that, as previously mentioned, the OCMAS specification establishes who is allowed to enter to the different spaces, and, for example, an irrigator playing the role of buyer would not be allowed to enter the registration room, nor a seller would be allowed to enter an auction room.

The following section presents the technical details of the infrastructure that supports the execution of such a scenario.

TECHNICAL SPECIFICATION

This section describes the infrastructure that supports the execution of training sessions by means of the previously introduced Assisted Hybrid Structured 3D Virtual Environment. The core execution infrastructure component is VIXEE, which is first presented. Then, some technical details are provided about the software agents involved in the training process as well as about the mechanism deployed for task-oriented conversations.

Execution Infrastructure (VIXEE)

VIXEE, is a robust Virtual Institution eXEcution Environment that provides interesting features such as multi-verse communication and dynamic manipulation of the virtual world content (Trescak, Rodríguez, López-Sánchez&Almajano, 2013). VIXEE is a generic and domain-independent solution. As specified

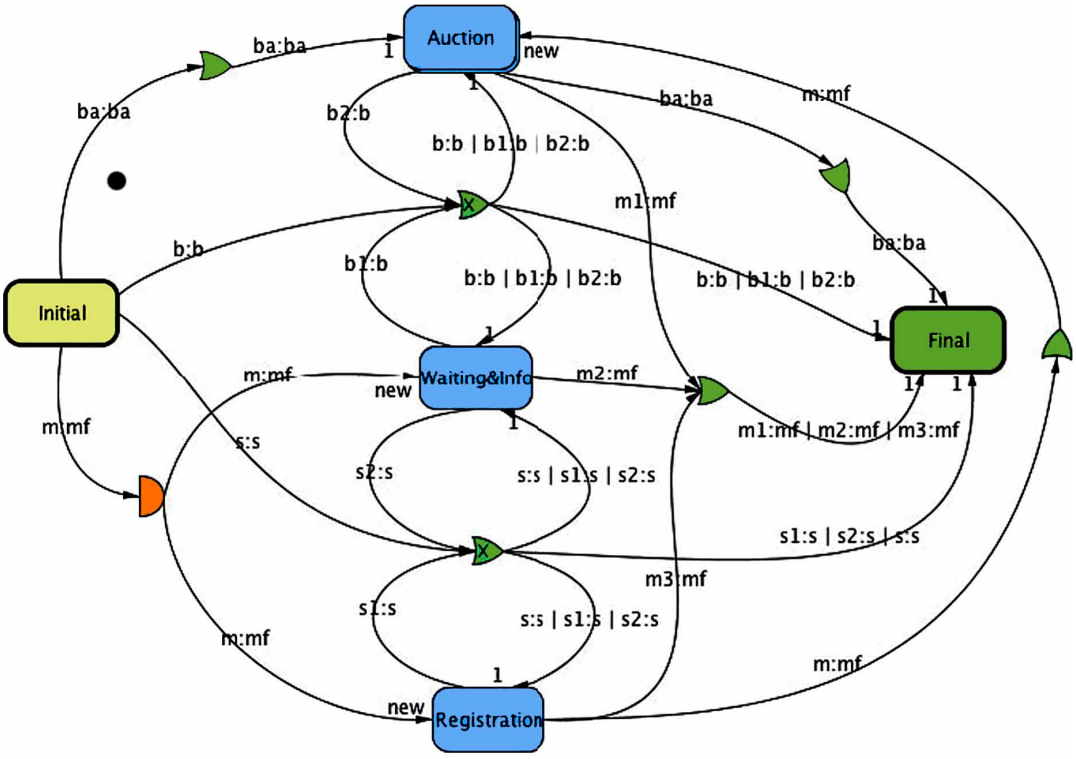
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

in the System Model Overview Section, and in particular, in Figure 1, its architecture is composed of three layers: i) Normative, ii) Visual Interaction and iii) Causal Connection.

The *Normative Layer* is composed by AMELI, the Electronic Institutions infrastructure that mediates agents' interactions while enforcing institutional rules (Esteva, de la Cruz & Sierra, 2002), the system specification and its running state. AMELI can be regarded as domain-independent because it can interpret any institution specification generated by ISLANDER tool (Esteva, de la Cruz & Sierra, 2002). It is implemented in JAVA and uses two TCP ports to communicate with the Causal Connection Layer.

Figure 7 depicts the specification of the Performative Structure of the v-mWater scenario. The so-called Performative Structure defines several dialogic activities (also referred as scenes) where agents participate enacting different roles. Additionally, it also specifies how agents can legally move among these activities depending on their roles. Specifically, a Performative Structure is defined as a graph where the nodes represent activities --and transitions, i.e. activity connectives used for synchronization-- linked by directed arcs, which represent the allowed agent movements between them. The example in Figure 7 depicts four different roles (mf: market facilitator; b: buyer; s: seller; and ba: basin authority) that can move between three scenario-dependant activities (Registration, Auction, and Waiting&Info). As depicted, the seller role is allowed to register water rights in the Registration activity and become informed about the market in the Waiting&Info activity. On the other hand, buyers can join the Auction activity once they have been informed about the registered water rights that will be auctioned. Finally, both the market facilitator and basin authority roles correspond to staff agents in charge of supporting these activities.

Figure 7. v-mWater Performative Structure



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

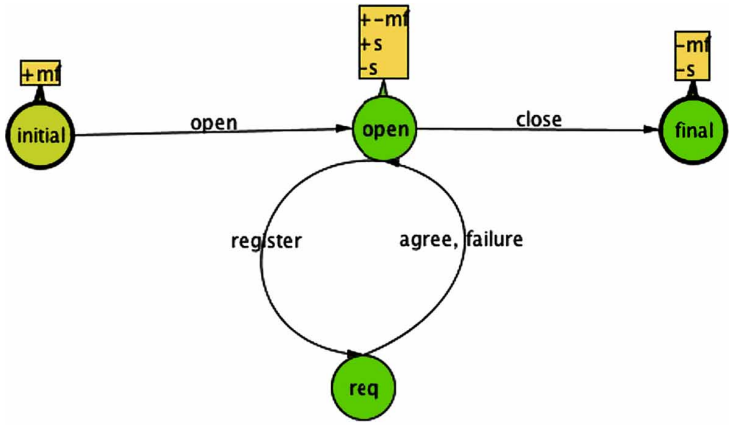
Virtual interactions within each activity are structured by means of well-defined protocols. These interaction protocols are specified in terms of Finite State Machines, whose states capture the possible activity states and whose transitions are followed depending on time-outs or whenever agents send the specific messages (i.e., illocution[2] schemes) specified in the edge labels. Since participants may enter or leave activities, protocol nodes also define the roles that are allowed to enter and exit at them. Figure 8 depicts the protocol followed within the Registration activity. Specifically, Figure 8 represents the four protocol's nodes (*initial*, *open*, *req*: registration requested, and *final*) and five illocutions (*open*, *close*, *register*, *agree* and *failure*). In particular, once a market facilitator staff agent (mf) has opened the activity and a seller has entered it, this seller asks for registering a water right by uttering the illocution *register*. This *register* illocution should specify: the identifier of the water right the participant wants to sell; the quantity of associated water; and a reference price for its negotiation in the Auction activity. Then, the protocol (activity) state changes from *open* to *req*. At state *req*, the market facilitator decides if the registration is valid by checking specific market regulations and sends back the result of the process to the seller: *agree* in case of success and *failure* otherwise. The seller can then leave the activity and the market facilitator agent will remain in the activity waiting for other sellers to register their water rights until he decides to close and leave the activity.

Note that the information specified in this protocol is used by tutoring assistant agents in the task-oriented conversation (see next section) in which they engage with human users, meanwhile the latters learn how to register a water right.

The *visual interaction layer* comprises several 3D virtual worlds. Each Virtual World can be implemented in a different programming language by using a different graphics technology. Common parts of a Virtual World are a Virtual World client and a Virtual World server. Such a server communicates with the causal connection layer by using a standard protocol (e.g. UDP, TCP or HTTP). In particular, our example scénario has been built with Open Simulator, an open source multi-platform, multi-user 3D Virtual World server ("OpenSimulator", n.d.).

Finally, the *causal connection layer* (Trescak, Rodríguez, López-Sánchez&Almajano, 2013) causally connects the visual interaction and the normative layers, that is, whenever one of them changes, the other one changes in order to maintain a consistent state (Maes & Nardi, 1988). As Figure 9 shows, this layer implements a Virtual Worlds Manager which is a multi-verse communication mechanism that allows

Figure 8. Registration protocol



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

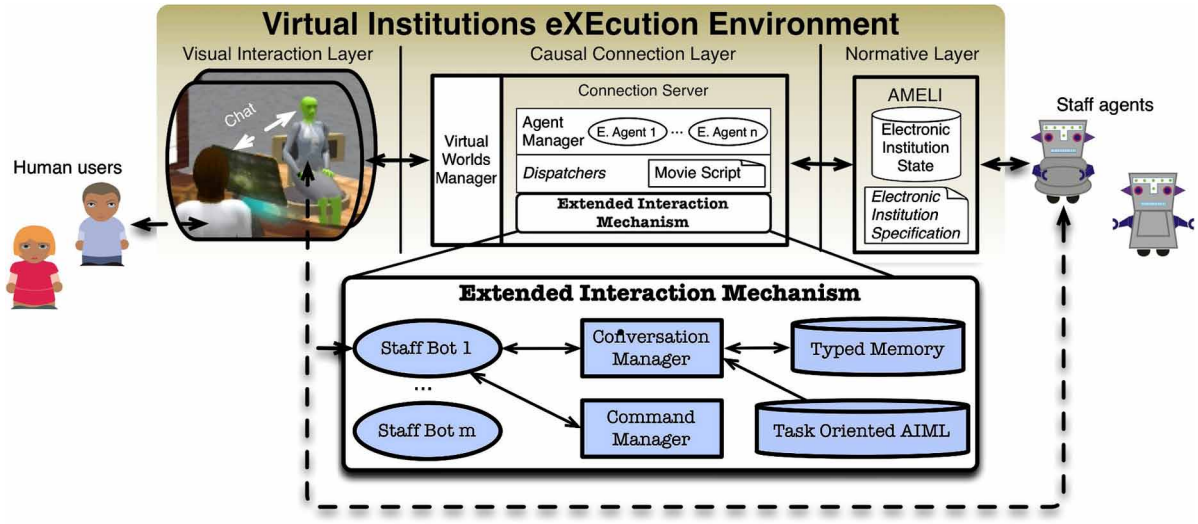
users from different virtual worlds to participate in the same Virtual Institution. Thus, the Agent Manager creates an External Agent for each connected user. Furthermore, dispatchers are in charge of the mapping between Virtual World actions and AMELI events -and vice versa- defined in the so-called movie script. On one hand, a message generated from AMELI provokes a Virtual World (VW) action so that the visualisation in all connected VWs is updated. On the other hand, for each institutional (regulated) action performed by a human avatar in the VW, a dispatcher sends the corresponding message to AMELI.

Human-agent interactions are ultimately supported by means of an the Interaction Mechanism at the bottom of Figure 9. Staff bots (connected to both staff agents and VW characters) are endowed with a textual communication system, which can be based on commands or/and on natural language dialogs. In the command based system, a Command Manager is in charge of mapping human commands to/from AMELIE messages (illocutions). As for the natural language dialogues, a Conversation Manager (CM) uses an extension of Artificial Intelligence Markup Language (AIML) to manage task-oriented conversations (Almajano, Mayas, Rodríguez & López-Sánchez, 2014). In particular, the Conversation Manager takes care of: i) identifying the task at hand (e.g., water right registration); ii) asking the human the required (if any) information (e.g., water right identifier, quantity or price of the associated water); iii) validating and storing provided inputs (if any) by considering both protocol's specification and the current state of the institution; and, finally, iv) informing about the result, which may imply composing an illocution from the gathered information or translating an illocution into natural language.

Task-Oriented Conversational Interaction

In real life institutions, users follow standardized communication protocols by means of natural language conversations in order to complete complex tasks. With the aim of providing a training scenario where the real institution is faithfully represented, our system has incorporated an interaction mechanism that allows tutoring agents (both staff and personal assistant agents), from now on referred as trainers, to converse in natural language with trainees (i.e. human users).

Figure 9. Overview of VIXEE Architecture

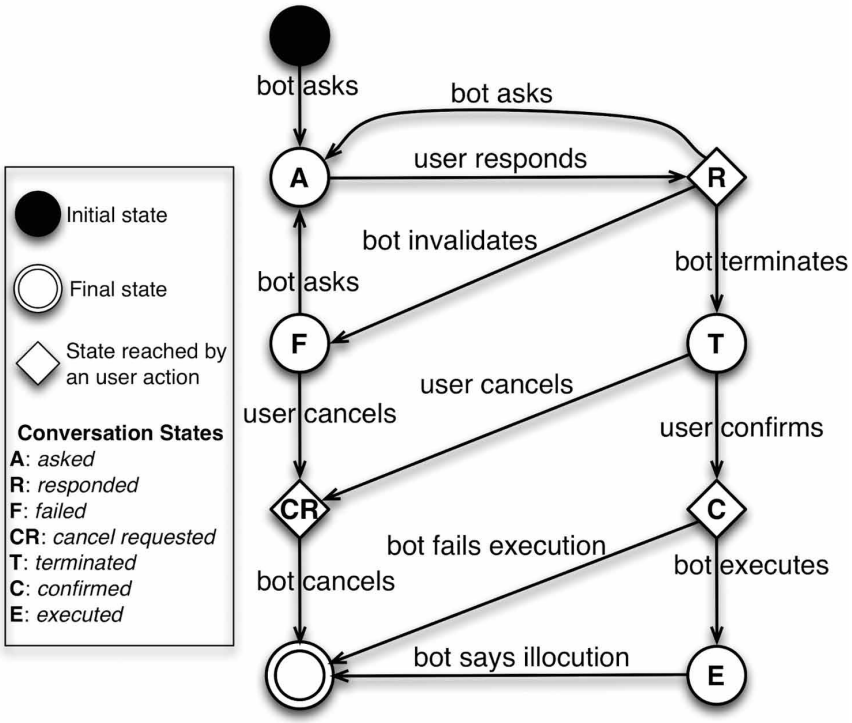


EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Using the conversational system, trainees can express their intentions using natural language dialogues and trainers have to understand trainees' entries and support them in their task completion. Trainer-trainee interaction in Assisted Hybrid Structured Virtual Environments have a complex structure, and thus, conversations must follow the communication protocols in the system specifications. As previously introduced, if a trainee initiates the interaction with a trainer, this requires the corresponding trainee to: First, identify trainee's task; Second, request required trainee entries and validate (and store) them based on both the system specification and current state; and Third, to redirect the compiled institutional message to the trainer so that the trainee's task can be completed. On the other direction, a trainer could also initiate an interaction with a trainee, if needed, by sending her or him a chat message in the simulated 3D environment.

This bidirectional process requires a trainer to dynamically control the flow and state of a task-oriented conversation with a trainee based on the trainee's entries, the system specification and its current state. The trainer delegates in a Conversation Manager the updating of the conversation's states, the storage of trainee's responses into a Typed Memory (with the types defined in the system specification) and the interpretation of the conversational knowledge. This conversational knowledge is based on Artificial Intelligence Mark-up Language (AIML), but we extended it in order to support structured/regulated conversations with the inclusion of some task-oriented special tags. This is so because basic AIML has been conceived to program reactive chatter bots that follow basic conversations with users, but it falls short for such complex interactions.

Figure 10. Finite State Machine depicting conversation's states during task stage



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Conversation Structure

In Assisted Hybrid Structured 3D Virtual Environments, a task-oriented conversation consists of three main interaction stages: welcoming, task, and farewell.

On the one hand, welcoming and farewell correspond to initial and final stages of the conversation. At these stages, the trainer is reactive and interacts with the trainee, basically, to greet and inform about the activity (welcoming stage), and say goodbye (farewell stage). On the other hand, as Finite State Machine in Figure 10 depicts, trainer-trainee interaction during the task stage is more complex. The change to this task stage is triggered when, during the welcoming stage, the trainee interacts with the trainer by sending a message, and the trainer recognizes this message content as a request to perform a task. During the task stage the human will send of messages to the corresponding trainer agent, and the trainer is an automata which, for a particular conversation state, selects the next action to perform based on the trainees' entries, messages defined in the system specification, and the allowed values of message contents at current state of the institution.

Once the trainer has identified the task (i.e. register), the task stage begins and the bot takes the initiative by asking to the trainee the data needed to complete the task. This interaction transits from the initial state to the asked state (A) in the Finite State Machine in Figure 10. In this state, the trainer waits for the trainee response, which will be processed in the responded (R) state.

Notice that both the format and value of trainee's entries (e.g., registration details) have to be checked against the system specification and current system state respectively. If these entries have incorrect format along a number of times, the conversation state becomes failed (F) and the trainer informs the trainee, asking for a confirmation to continue with the task. If the trainee confirms that she or he wants to continue with the task, the trainer repeats the last question and the conversation reaches again the asked state. Nevertheless, if she or he decides not to continue (cancel requested: CR state), it provokes the trainer to finish the task, reaching then the final state of task stage.

Alternatively, if all information is successfully gathered, the trainer terminates the questions (terminated state T) by presenting to the trainee the obtained information and requesting her or his confirmation before executing the actual message in the system. Then, the trainee can either request the cancellation with the same consequences as for previous failed state (F) or confirm that task's data is correct (confirmed state C). If the trainee user confirms, the trainer agent facilitates the execution of the corresponding message (i.e. register illocution) within the system. Upon successful message execution, the executed (E) state is reached, and the trainer sends a natural language message (response) back to the trainee, finishes the conversation, and changes to farewell stage.

EVALUATION

In order to check our scénario with real users we did a twofold evaluation, one about the virtual environment as a whole, where human-agent (i.e., trainee-trainer) interaction was based on commands; and one specific for the conversational mechanism, where command-based interaction was compared against the natural language dialogues. A total number of 37 participants were recruited to perform three separate tests that focus on studying specific aspects of our approach. The ages of the users were chosen within the range of working ages in many countries, as we intend the application to be useful both in education institutions and professional settings.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Virtual Environment Evaluation

Initially, the virtual environment was tested with an exploratory usability test (Rubin & Chisnell, 2008), mainly to get qualitative data on how the 3D virtual institution helped on the learning of the real institution. In the test, trainees had to accomplish a unique task: "Act as a seller, and register a water right for a price which is 5 euros higher than the price of the last transaction done", but performing it meant to plan her navigation through the 3D virtual scenario (the trainee had to visit two activity rooms), to get information in the Waiting&Info room, to work out an adequate registration price, and to achieve registration successfully.

The test aimed to know the effectiveness of the tool to reach trainees' goals and their willingness to use it, but also to observe human users interacting with the application, observe the mechanisms they used to obtain information, and value the easiness of getting that information.

The test was delivered to 13 users, ages between 18 to 54, 7 of them females and 6 males. As for experience with computers: 2 of the users had basic computer experience; 5 medium experience; and 6 advanced experience; whereas 1 had high experience with virtual environments, 3 some experience with them and 9 none. In order to facilitate interaction, users did a basic training session on functions and operation of the virtual environment before the test. During the test every interaction was recorded and analysed to observe the navigation paths of users and the obstacles they found to get theirs goals. Finally, a post-test questionnaire was administered to participants in order to know their opinion and satisfaction on the experience.

Despite the complexity of the task at hand, all users were able to complete the task, 10 of them with a correct price. Participants did not find any obstacle that prevented them completing the task, although 5 of them did not figure out they had to check prices before registering their water right.

The post test survey included 13 questions about:

Q1: Situatedness and movement in 3D
Q2: Virtual Environment walking (2.1) and teleport (2.2) comfortability
Q3: Info gathering (panel/bot)
Q4: Human-bot interaction
Q5: Bot visual distinction
Q6: Chat-based bot communication
Q7: Task easiness
Q8: Immersiveness in 3D
Q9: Improved opinion of 3D VWs
Q10: Likeliness of future usage
Q11: 3D interface usefulness
Q12: Overall system opinion
Q13: User's comments

Figure 11 shows the average answers to the post-test questionnaire. As a summary, trainees evaluated very positively the facilities to gather information, both information panels and information bots, they had a good overall impression about interaction with bots but they did not like command-based dialogues. Observing the records of the tasks it was noted that older users found the environment difficult

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

but also more rewarding than younger players. Users without experience found it harder to play and, on the overall, computer experience was a direct indicator of the attained level of immersion. Users gave an average value of 3.5 over 5 on their intention to use a similar system in the future.

After the initial test, a second one was conducted where the assistance service provided by software trainers was included, and thus, its evaluation focuses on this service. The task was the same as in the initial test, but the tutoring assistants provided plan advices such as the one depicted in previous Figure 6. The second test script also consisted in a training session, the test and an ad-hoc post-test questionnaire. In this case there were 14 users (different from the ones in previous test), ages between 23 to 66, 4 of them females and 10 males. As of experience with computers, 4 of the users had basic computer experience and 10 advanced experience, while 8 had high experience with virtual environments, and 6 none.

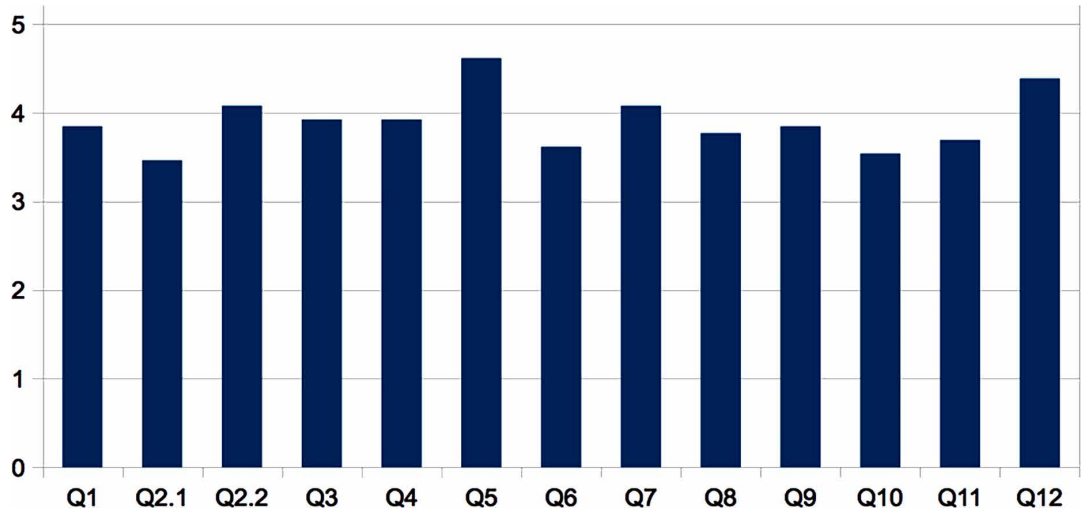
The post test survey included 8 questions about:

Q1: Info gathering (panel/bot)
Q2: Human-bot interaction
Q3: Bot visual distinction
Q4: Dialogue-based bot communication
Q5: Overall system opinion
Q6: Assistance usefulness
Q7: Tutoring assistant interaction
Q8: Advice Understanding

And an open question to gather user's comments

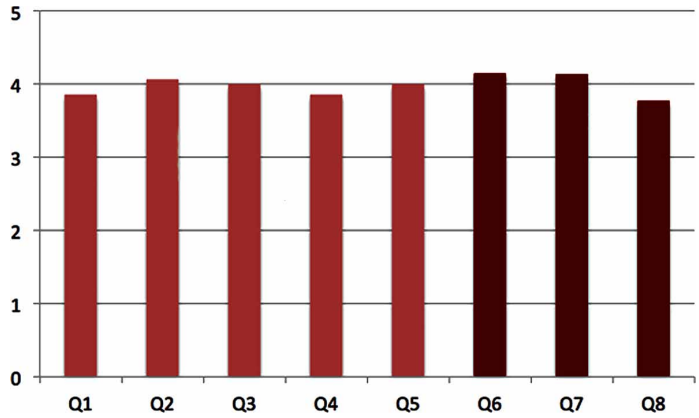
Questions in the post-test questionnaire were focused on assistance. Figure 12 depicts its average results. Overall results were similar to the previous ones, but it is worth noticing the questions relating to the assistance service got very high scores.

Figure 11. First post-test questionnaire results
X Axis: Questions From Q1 to Q12. Y axis: average values



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 12. Second post-test questionnaire results.
X axis: questions from Q1 to Q8. Y axis: average values



Along the tests, 13 out of 14 participants requested assistance voluntarily in different moments of the task, the interaction with bots was very natural and users could ask and understand the given information. In particular, plans were followed as expected. With the assistance, users reduced significantly the needed steps to reach their goal (from 10.8 in test 1 in average to 7 in test 2).

Conversational Interaction Evaluation

On a final test, the conversational mechanism was tested against the command-based interaction mechanism with a summative and comparative evaluation method (Bowman, Gabbard & Hix, 2002). 10 additional participants were recruited for this third test, ages between 21 to 45, 3 of them females and 7 males. Moveover, half of them (5) had previous experience with command-based systems, and the other half (5) were novices.

This third test followed a within-subject design and each subject did the same two tasks. Task 1 was defined as "Ask the Information Manager about the last 2 transactions in the market"; and Task 2 was defined as "ask the Registration Manager to register a water right, identified as wr1, for a price of 25€". Task order was alternated to avoid learning and fatigue effects, and thus half the participants started by using the command-based interaction style, whereas the other half started by using the conversational style.

Before the test a short training session was delivered to teach how to interact with objects and bots and how to move in 3D environments, and as in the other two evaluations, there was also an ad hoc post-test questionnaire. This post-test consisted in the following 7 questions:

Q1: I did not feel I needed help while talking to the bot
Q2: I did not feel frustrated while talking to the bot
Q3: What the bot said to me made sense
Q4: I did know what to answer to the bot
Q5: How comfortable was the communication
Q6: I felt that the bot understood me
Q7: The bot had answers I expected

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Possible answers to these questions were: 1 never / very uncomfortable; 2 sometimes/uncomfortable; 3 regularly/normal; 4 often/comfortable; 5 always/very comfortable.

Figure 13 depicts average values for the answers to these 7 different questions. As we can observe in the Figure, overall satisfaction was good. In particular all dialog related questions had an average value above 4. Taking a closer look to the figures, we observed that novices valued most the conversational mechanisms while experts rated almost identically both mechanisms. The test was also useful to detect that bots should greet users on its own initiative as well as to be able to extract and process multiple data from a single conversational message.

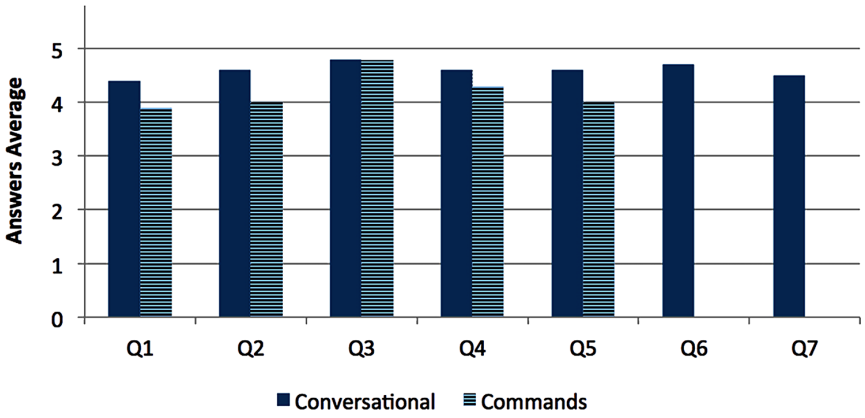
Regarding effectiveness, the conversational interaction style obtained a task completion rate of 100%, while the command-based approach only a 70% rate. Additionally, errors were much lower in the conversational system (0.3 errors versus 2.2 errors in average in command-based system). Relating to efficiency, the number of messages in the conversational system was superior (that was forced by the design of the conversation interaction), but the messages were more focused and driven by the goal.

More tests and a quantitative approach would have yield more precise results, and maybe a segmentation of users based on age or experience could orient new paths for improvement. Nevertheless these evaluations gave us enough data to orient our future work.

DISCUSSION AND FUTURE WORK

We can see our training infrastructure as an authentic learning approach applied to learn specific skills and knowledge to perform tasks in a job (Donovan et al., 1998) as it provides a concise and immediate learning process. The term *authentic learning* refers to an "*educational and instructional techniques focused on connecting what students are taught in school to real-world issues, problems, and applications*" ("Authentic...", n.d.). Authentic learning methods include techniques that allow trainers to explore, discuss and build concepts and relationships in real-worlds situations.

Figure 13. Third post-test questionnaire results
X axis: questions from Q1 to Q7
Y axis: average values for both conversational and commands interfaces (where applicable)



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Herrington, Reeves and Oliver (2014) developed the authentic technology-based learning environment model that defines nine practical characteristics derived from the cognitive and comprehensive learning experience. Next, we analyse these characteristics in our training scenario, showing its own advantages and limitations:

- **An Authentic Context Should be Used:** The scenario should reflect the real environment of the training, not a simplified version of the reality. Our training scenario defines a real complex and sustained trading situation with norms and roles that replicate the complexity of the irrigation market. The use of a virtual world gives the immersion and the strong feeling of "being there" while also allowing trainees to safety learn in a virtual environment.
- **Authentic Tasks Based on Ill-Defined Problems[3]:** This characteristic requires that trainers define tasks and sub-tasks that require to solve challenges faced by trainees. Also, tasks should allow competing solutions and a diversity of outcomes instead of one single correct answer. In our case, tasks can be solved by performing different plans, and thus, their solutions are not fixed. For instance, when a participant starts the water market session, she can freely navigate through the environment and access to either one of the three activities involved in the water rights' negotiation: Waiting and Information, Registration or Auction. Moreover, trainees can think different strategies to solve some tasks. Some participants enter the Waiting and Information Room to get the actual price of a water right, whereas others only watch the panel information from the window in order to get the desired information faster.
- **Allow to Access to Expert Performances and Modeling Processes:** the environment replicates the forms of support in the "real" context, as well as it is able to support different levels of expertise. In a 3D virtual world, expertise and performance can be captured and analyzed easily to reinforce some skills and improve their performance. In our case, tutoring assistants help trainees on-demand using a conversational system where trainees can express their needs using natural language dialogues and trainers understand trainees' entries and support them in their task completion, using information of the system specification and their current state. Such tutoring assistant agents use an extension of AIML to support task-oriented conversations. Furthermore, we may use adaptive agents to react to different expertise levels or to different behaviour. Moreover, the selection of the adequate tutoring of different levels could be based on a Recommender System. Specifically, Recommender Systems (RSs) represent one approach to develop more personalised information systems, which provide users with suggestions taking into account user requirements and community preferences (Konstan & Riedl, 2012). We consider that RSs can also be a significant part of e-learning systems for personalization and recommendation of appropriate materials to the learner considering its level of expertise(Parveen, Rubina & Anant, 2012). Although the use of RSs has been extensively focused on web-based scenarios, more recently, there is an interest in the study of their integration within 3D virtual environments, particularly in an e-commerce context (Arroyo, Serradilla & Calvo, 2010; Xu & Yu, 2010; Contreras, Salamó, Rodríguez & Puig, 2014; Contreras, Salamó, Rodríguez & Puig, 2013).
- **Multiple Roles and Perspectives Should be Provided in Order to Provide Trainees with the Opportunity to Examine the Problem from Different Perspectives Using a Variety of Resources:** In our case, participants can enact different roles in the virtual world so to learn different expected behaviours and perspectives of the simulated market. On the other hand, it is worth mentioning the potentiality of the assistance layer to offer message personalization for dif-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ferent profiles of users, in particular for users with disabilities. The personalization could consist of more verbose textual explanations for blind users, specific plans designed for users with motor impairments with less interactions, plain language instructions for users with cognitive or learning disabilities, etc. and thus, each user will solve the problem in a different way, taking profit of his strengths and minimizing his shortcomings.

- **Collaborative Construction of Knowledge:** Market institutions are generally characterised by a rather competitive coordination schema, and thus, its regulations and protocols scarcely support collaboration. Nevertheless, the overall learning approach we propose is scenario independent, and thus, activities could be applied to other domains and so specified to support collaborative knowledge creation and knowledge sharing.
- **Reflection:** this property provides trainees with the opportunity to examine their performances so to self-evaluate and repeat some parts of the training whenever required. We can use gamification strategies based on badges, scores and feedback of the realized actions to provide trainees with an immediate response from the system, as suggested in (Teräs, Teräs & Reiners, 2014) . Additionally, our virtual world could include game-based feedback(Deterding, Dixon, Khaled & Nacke, 2011). which stores points and achievements accumulated over time in a profile that provides a trainee with a potentially powerful representation of their avatar identity.
- **Articulation:** it provides mechanisms to communicate ideas between pairs and to reason some concepts. Currently, our 3D virtual environments are endowed with a chat-based communication mechanism that allows participants to exchange textual messages, and these chats can support both sharing ideas between participants as well as reasoning about specific concepts that should be learned. In our environment, though, these chat-based messages are allowed but they are not regulated by the virtual institution norms, and, in particular, they are not explicitly forbidden by any norm or protocol. Chat-based communication can be viewed as a complementary way to share knowledge and experiences among participants.
- **Scaffolding and Coaching:** it allows to integrate some type of assessment or assistance into the activities rather than employing external tests. In this respect, we could consider tutoring assistant agents as a first step towards this direction, since they monitor trainees' status and provide meaningful assistance.
- **Authentic Assessment Allows Trainees to Demonstrate Their Acquired Skills and Their Progression in the World:** This last learning characteristic can be naturally accomplished in our "real" immersive scenario.

This authenticity, and the technical feasibility of the model, as demonstrated by the text, makes this option a real alternative to live training for professional training both in educational and professional settings. The cost, flexibility and capacity of the System to be enriched will make it very attractive for lifelong learning and job orientation in a ever-changing world with increasing demands of professional skills.

As previously discussed, our future plans include to extend our current system to support a full authentic learning experience, such as using recommender systems, gamification strategies and adaptive learning systems to support a full authentic learning experience.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

In real life institutions, people gather in physical spaces to develop complex tasks, such as trading or paperwork handling. Therein, institutional members, i.e. staff, are in charge of enabling the complex institutional activities while fulfilling the norms that regulate each particular institution. External users enter the institution and achieve their goals by interacting with both other external users and staff as well as with the environment. These interactions follow intricate communication protocols, and may have irreversible consequences, such as legal commitments.

This chapter has presented a training system that teaches human users to participate in such a complex scenarios. In particular, we propose to simulate a real institution and enable assistance mechanisms by developing it as an *Assisted Hybrid Structured 3D Virtual Environment*. The institutional physical spaces are represented in a 3D virtual environment, and the interactions between participants are regulated by means of an Organisation Centered Multi-Agent System (OCMAS).

The training system is populated by tutoring agents that are programmed to train human users (i.e. trainees) not only on how to participate within the real institution --by playing staff roles-- but also to understand the social model underlying it --by providing personal assistance. A personal assistant directly and solely supports a trainee, providing her or him with services such as information about the institution or advice on how to proceed along it. Additionally, tutoring agents are able to communicate with trainees in natural language conversations by means of our proposed Task-Oriented AIML. As a result, human users are trained to perform complex real-life tasks within an immersive scenario, in a seamless way and avoiding the fear to fail.

We have empirically tested our proposal by simulating a water market, where sellers enter to register water rights, and buyers negotiate them within an auction. Tutoring agents play staff roles market facilitator and basin authority, and tutoring personal assistants provide advices on how to perform a registration task. The performed experiments show that when the training system incorporates tutoring agents and conversational command styles its usability, efficiency, and efficacy increases.

To sum up, the training system presented in this chapter is an authentic learning approach for real institutions with a number of promising research lines still open. For example, techniques such as recommender systems, gamification strategies and adaptive learning systems taking into account accessibility issues, will be used to support full authenticity and education requirements. In its corrent State of development it is already a real alternative to live training, and it could be very attractive for higher education institutions of big companies looking for longlife learning. The immersive and supportive approach presented here can have a big impact on how training is modelized from now on.

REFERENCES

Alicebot. (n.d.). Retrieved May 28, 2015, from http://www.alicebot.org/

Almajano, P. (2014). *Assisted Hybrid Structured 3D Virtual Environments*. Universitat de Barcelona.

Almajano, P., Lopez-Sanchez, M., Rodriguez, I., & Trescak, T. (2014). Assistant agents to advice users in hybrid structured 3D virtual environments. *Computer Animation and Virtual Worlds, 25*(3-4), 495–504. Retrieved from http://doi.wiley.com/10.1002/cav.1598

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Almajano, P., Mayas, E., Rodriguez, I., & Lopez-Sanchez, M. (2014b). Conversational Structured Hybrid 3D Virtual Environments. In *Proceedings of the XV International Conference on Human Computer Interaction - Interacción '14* (pp. 1–8). New York: ACM Press. Retrieved from http://dl.acm.org/citation.cfm?id=2662253.2662316

Almajano, P., Trescak, T., Esteva, M., Rodríguez, I., & López-Sánchez, M. (2012). v-mWater: a 3D virtual market for water rights (demonstration). In Procedings of AAMAS (pp. 1483-1484).

Arroyo, A., Serradilla, F., & Calvo, Ó. (2010). Modeling Spatial-Temporal Context Information in Virtual Worlds. In N. García-Pedrajas, F. Herrera, C. Fyfe, J. M. Benítez, & M. Ali (Eds.), Trends in Applied Intelligent Systems (Vol. 6096, pp. 437–447). Berlin: Springer Berlin Heidelberg. doi:10.1007/978-3-642-13022-9_44

Authentic Learning Definition - The Glossary of Education Reform. (2014). Retrieved May 28, 2015, from http://edglossary.org/authentic-learning/

Bartle, R. A. (2003). *Designing Virtual Worlds*. New Riders.

Bogdanovych, A., Kiran, I., & Simoff, S. (2012). The City of Uruk: Teaching Ancient History in a Virtual World. In *Intelligent Virtual Agents: 12th International Conference, IVA 2012, Santa Cruz, CA, USA, September, 12-14, 2012. Proceedings* (Vol. 7502, pp. 28–35). Springer. doi:10.1007/978-3-642-33197-8_3

Bogdanovych, A., Rodriguez-Aguilar, J. A., Simoff, S., & Cohen, A. (2010). Authentic interactive re-enactment of cultural heritage with 3d virtual worlds and artifi cial intelligence. *Applied Artificial Intelligence*, *24*(6), 617–647. doi:10.1080/08839514.2010.492172

Bouras, C., & Tsiatsos, T. (2006). Educational virtual environments: Design rationale and architecture. *Multimedia Tools and Applications*, *29*(2), 153–173. doi:10.100711042-006-0005-7

Bowman, D. A., Gabbard, J. L., & Hix, D. (2002). A Survey of Usability Evaluation in Virtual Environments: Classification and Comparison of Methods. *Presence (Cambridge, Mass.)*, *11*(4), 404–424. doi:10.1162/105474602760204309

Chalupsky, H., Gil, Y., Knoblock, C. A., Lerman, K., Oh, J., Pynadath, D. V., … Tambe, M. (2001). Electric Elves: Applying Agent Technology to Support Human Organizations. In *Proceedings of the Thirteenth Conference on Innovative Applications of Artificial Intelligence Conference* (pp. 51–58). AAAI Press.

Contreras, D., Salamó, M., Rodríguez, I., & Puig, A. (2013). Integrating Conversational Recommender System within a 3D Collaborative Environment. In. *XV Conf. of Spanish Association for Artificial Intelligence*, (pp. 89–98).

Contreras, D., Salamó, M., Rodríguez, I., & Puig, A. (2014). An approach to improve user experience with conversational recommenders through a 3D virtual environment. In *Proceedings of the XV International Conference on Human Computer Interaction - Interacción '14* (pp. 1–4). New York: ACM Press. 10.1145/2662253.2662319

Craven, D. (2015). Gamification in Virtual Worlds for Learning: A Case Study of PIERSiM for Business Education. In L. C. W. Torsten Reiners (Ed.), *Gamification in Education and Business* (pp. 385–401). Springer; doi:10.1007/978-3-319-10208-5_19

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Dalgarno, B., & Lee, M. J. W. (2010). What are the learning affordances of 3-D virtual environments? *British Journal of Educational Technology*, *41*(1), 10–32. doi:10.1111/j.1467-8535.2009.01038.x

Danforth, D. R., Procter, M., Chen, R., Johnson, M., & Heller, R. (2009, July 19). Development of Virtual Patient Simulations for Medical Education. *Journal of Virtual Worlds Research*. doi:10.4101/jvwr.v2i2.707

Del Blanco, Á., Marchiori, E. J., Torrente, J., Martínez-Ortiz, I., & Fernández-Manjón, B. (2013). Using e-learning standards in educational video games. *Computer Standards & Interfaces*, *36*(1), 178–187. doi:10.1016/j.csi.2013.06.002

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness. In *Proceedings of the 15th International Academic MindTrek Conference on Envisioning Future Media Environments - MindTrek '11* (p. 9). New York: ACM Press. 10.1145/2181037.2181040

Dickey, M. D. (2010). Teaching in 3D: Pedagogical Affordances and Constraints of 3D Virtual Worlds for Synchronous Distance Learning. *Distance Education*, *24*(1), 105–121. doi:10.1080/01587910303047

Dong, R., McCarthy, K., O'Mahony, M., Schaal, M., & Smyth, B. (2012). Towards an intelligent reviewer's assistant. In *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces - IUI '12* (p. 159). New York: ACM Press. 10.1145/2166966.2166995

Donovan, M., Suzanne, E., Bransford, J. D. E., & Pellegrino, J. W., E. (1998). How People Learn: Bridging Research and Practice. Washington, DC: National Academy Press.

Duncan, I., Miller, A., & Jiang, S. (2012). A taxonomy of virtual worlds usage in education. *British Journal of Educational Technology*, *43*(6), 949–964. doi:10.1111/j.1467-8535.2011.01263.x

Esteva, M. (2003). *Electronic institutions. from specification to development*. Universitat Politècnica de Catalunya. Retrieved from http://www.iiia.csic.es/~marc/phd.pdf.zip

Esteva, M., de la Cruz, D., & Sierra, C. (2002). ISLANDER. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems part 3 - AAMAS '02* (p. 1045). New York: ACM Press. doi:10.1145/545056.545069

Fikar, P., Schönauer, C., & Kaufmann, H. (2013). The Sorcerer's Apprentice: A serious game aiding rehabilitation in the context of Subacromial Impingement Syndrome. In *Proceedings of the ICTs for improving Patients Rehabilitation Research Techniques: PervasiveHealth '13* (pp. 327–330). Brussels, Belgium: IEEE. doi:10.4108/icst.pervasivehealth.2013.252224

Ghanbarzadeh, R., Ghapanchi, A. H., Blumenstein, M., & Talaei-Khoei, A. (2014). A decade of research on the use of three-dimensional virtual worlds in health care: A systematic literature review. *Journal of Medical Internet Research*, *16*(2), e47. doi:10.2196/jmir.3097 PMID:24550130

Graesser, A. C., Chipman, P., Haynes, B. C., & Olney, A. (2005). AutoTutor: An Intelligent Tutoring System With Mixed-Initiative Dialogue. *IEEE Transactions on Education*, *48*(4), 612–618. doi:10.1109/TE.2005.856149

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Grammenos, D., Savidis, A., & Stephanidis, C. (2007). Unified design of universally accessible games. In C. Stephanidis (Ed.), *Universal Access in Human-Computer Interaction. Applications and Services: 4th International Conference on Universal Access in Human-Computer Interaction, UAHCI 2007 Held as Part of HCI International 2007* (*Vol. 4556*, pp. 607–616). Berlin: Springer-Verlag. doi:10.1007/978-3-540-73283-9_67

Hack, C. J. (2015). The benefits and barriers of using virtual worlds to engage healthcare professionals on distance learning programmes. *Interactive Learning Environments*, 1–14. doi:10.1080/10494820.2015.1057743

Herrington, J., Reeves, T. C., & Oliver, R. (2014). Authentic Learning Environments. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), Handbook of Research on Educational Communications and Technology (pp. 401–412). New York, NY: Springer New York. doi:10.1007/978-1-4614-3185-5_32

Hew, K. F., & Cheung, W. S. (2010). Use of three-dimensional (3-D) immersive virtual worlds in K-12 and higher education settings: A review of the research. *British Journal of Educational Technology*, *41*(1), 33–55. http://doi.wiley.com/10.1111/j.1467-8535.2008.00900.x doi:10.1111/j.1467-8535.2008.00900.x

Kolb, D. A. (1983). *Experiential Leraning: Experience as the source of learning*. Prentice Hall.

Kolb, D. A. (2014). *Experiential Learning: Experience as the Source of Learning and Development* (2nd ed.). FT Press.

Konstan, J. A., & Riedl, J. (2012). Recommender systems: From algorithms to user experience. *User Modeling and User-Adapted Interaction*, *22*(1-2), 101–123. doi:10.100711257-011-9112-x

Laister, J., & Kober, S. (2002). Social Aspects of Collaborative Learning in Virtual Learning Environments. In *Networked learning*. Sheffield, UK: Lancaster University. Retrieved from http://www.networkedlearningconference.org.uk/past/nlc2002/proceedings/papers/19.htm

Maes, P., & Nardi, D. (1988). *Meta-level architectures and reflection*. New York: Elsevier Science Pub. Co. Inc.

McInnerney, J. M., & Roberts, T. S. (2004). Online Learning: Social Interaction and the Creation of a Sense of Community. *Journal of Educational Technology & Society*, *7*(3), 73–81.

Monahan, T., & McAdrle, G. (2008). Virtual reality for collaborative e-learning. *Computers & Education*, *50*(4), 1339–1353. doi:10.1016/j.compedu.2006.12.008

Moskaliuk, J., Bertram, J., & Cress, U. (2013). Impact of virtual training environments on the acquisition and transfer of knowledge. *Cyberpsychology, Behavior, and Social Networking*, *16*(3), 210–214. doi:10.1089/cyber.2012.0416 PMID:23363227

Oh, J., Meneguzzi, F., Sycara, K., & Norman, T. J. (2011). Prognostic normative reasoning in coalition planning. In *The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 3: AAMAS '11* (pp. 1233–1234). International Foundation for Autonomous Agents and Multiagent Systems.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Okamoto, S., Sycara, K., & Scerri, P. (2009). Personal assistants for human organizations. In V. Dignum (Ed.), *Handbook of Research on Multi-Agent Systems: Semantics and dynamics of organizational models* (pp. 514–540). IGI Global; doi:10.4018/978-1-60566-256-5.ch021

OpenSimulator. (n.d.). Retrieved May 5, 2015, from http://opensimulator.org/wiki/Main_Page

Parveen, R., Jaiswal, A. K., & Kant, V. (2012). E-Learning Recommendation Systems – A Survey. *International Journal of Engineering Research and Development, 4*(12), 10–12.

Porter, J. R. (2014). Understanding and addressing real-world accessibility issues in mainstream video games. *ACM SIGACCESS Accessibility and Computing*, (108), 42–45. doi:10.1145/2591357.2591364

Richards, D., & Taylor, M. (2015). A Comparison of learning gains when using a 2D simulation tool versus a 3D virtual world: An experiment to find the right representation involving the Marginal Value Theorem. *Computers & Education*, *86*, 157–171. doi:10.1016/j.compedu.2015.03.009

Rubin, J., & Chisnell, D. (2008). Handbook of Usability Testing: Howto Plan, Design, and Conduct Effective Tests (2nd ed.). Wiley Publishing.

San Chee, Y. (2001). Networked virtual environments for collaborative learning. In *Proceedings of ICCE/ SchoolNet 2001Ninth International Conference on Computers in Education* (Vol. 311).

Teräs, H., Teräs, M., & Reiners, T. (2014). Giving virtual world learning a pedagogical boost with a gamified authentic learning design. In *Proceedings of the 26th EdMedia: World Conference on Educational Media & Technology* (pp. 1688–1693). Tampere, Finland: Association for the Advancement of Computing in Education (AACE).

Torrente, J., Freire, M., Moreno-Ger, P., & Fernández-Manjón, B. (2015). Evaluation of semi-automatically generated accessible interfaces for educational games. *Computers & Education*, *83*, 103–117. doi:10.1016/j.compedu.2015.01.002

Transcript: wheeling in second life. (n.d.). Retrieved April 13, 2015, from http://uiaccess.com/transcripts/ wheeling_in_second_life.html

Trescak, T., Rodriguez, I., Lopez Sanchez, M., & Almajano, P. (2013). Execution infrastructure for normative virtual environments. *Engineering Applications of Artificial Intelligence*, *26*(1), 51–62. doi:10.1016/j.engappai.2012.09.016

Vickers, S., Istance, H., & Hyrskykari, A. (2013). Performing Locomotion Tasks in Immersive Computer Games with an Adapted Eye-Tracking Interface. *ACM Transactions on Accessible Computing*, *5*(1), 2. doi:10.1145/2531922.2514856

Wallace, R. S. (2009). The Anatomy of A.L.I.C.E. In R. Epstein, G. Roberts, & G. Beber (Eds.), *Parsing the Turing Test* (pp. 181–210). Dordrecht: Springer Netherlands; doi:10.1007/978-1-4020-6710-5_13

Waterfind Australia | Water Trading Made Easy. (n.d.). Retrieved May 5, 2015, from http://www.waterfind.com.au/

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Westin, T., Bierre, K., Gramenos, D., & Hinn, M. (2011). Advances in game accessibility from 2005 to 2010. In C. Stephanidis (Ed.), *Universal Access in Human-Computer Interaction. Users Diversity: 6th International Conference, UAHCI 2011, Held as Part of HCI International 2011* (pp. 400–409). Springer-Verlag. 10.1007/978-3-642-21663-3_43

Wrzesien, M., & Alcañiz Raya, M. (2010). Learning in serious virtual worlds: Evaluation of learning effectiveness and appeal to students in the E-Junior project. *Computers & Education*, *55*(1), 178–187. doi:10.1016/j.compedu.2010.01.003

Xu, B., & Yu, Y. (2010). A Personalized Assistant in 3D Virtual Shopping Environment. In *2010 Second International Conference on Intelligent Human-Machine Systems and Cybernetics* (Vol. 2, pp. 266–269). IEEE. 10.1109/IHMSC.2010.167

Yaich, R., Boissier, O., Picard, G., & Jaillon, P. (2013). Adaptiveness and social-compliance in trust management within virtual communities. *Web Intelligence and Agent Systems*, *11*(4), 315–338.

Yuan, B., Folmer, E., & Harris, F. C. (2010). Game accessibility: A survey. *Universal Access in the Information Society*, *10*(1), 81–100. doi:10.100710209-010-0189-5

KEY TERMS AND DEFINITIONS

3D Virtual World / Environment (3DVW): Persistent virtual locations where participants, represented as animated figures in a three dimensional visual environment, may interact with other participants or objects in the environment in real time.

Accessibility: The usability of an application for users with disabilities who often use assistive tools and different modalities of interaction (only text, no audio, keyboard instead of mouse…)

Agent: (Intelligent) autonomous entity (piece of software) that interacts within an environment.

Artificial Intelligence Markup Language: (AIML): An XML dialect for creating natural language software agents.

Assisted Hybrid Structured 3D Virtual Environment (3DVW): Where participants, which can be human and software agents, interact following the rules of an organization (an OCMAS) and are assisted accordingly.

Conversational Interaction: Interaction conducted in a dialogical way, by exchanging natural language messages.

Hybrid Structured 3D Virtual Environment (3DVW): Where participants, which can be human and software agents, interact following the rules of an organization (an OCMAS)

Multi-Agent System (MAS): Computerized system composed of multiple interacting (intelligent) agents within an environment.

Organization Centered Multi-Agent System (OCMAS): MAS endowed with an organization or institution, where agents interact in a structured way which is defined by the organization.

Recommender Systems: Systems that provide users with suggestions of items, taking into account user's requirements or preferences.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ENDNOTES

[1] An open Organisation-Centred Multi-Agent System, where agents may not belong to the institution and interact dialogically, by sending illocutions.

[2] An illocution is any structured message defined for an institution that participants utter when interacting within it.

[3] By ill-defined problem we mean problems that can be solved by different solution paths.

This research was previously published in the Handbook of Research on 3-D Virtual Environments and Hypermedia for Ubiquitous Learning edited by Francisco Milton Mendes Neto, Rafael de Souza, and Alex Sandro Gomes; pages 192-219; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 69
Semantic Search Exploiting Formal Concept Analysis, Rough Sets, and Wikipedia

Yuncheng Jiang
South China Normal University, China

Mingxuan Yang
South China Normal University, China

ABSTRACT

This article describes how the traditional web search is essentially based on a combination of textual keyword searches with an importance ranking of the documents depending on the link structure of the web. However, one of the dimensions that has not been captured to its full extent is that of semantics. Currently, combining search and semantics gives birth to the idea of the semantic search. The purpose of this article is to present some new methods to semantic search to solve some shortcomings of existing approaches. Concretely, the authors propose two novel methods to semantic search by combining formal concept analysis, rough set theory, and similarity reasoning. In particular, the authors use Wikipedia to compute the similarity of concepts (i.e., keywords). The experimental results show that the authors' proposals perform better than some of the most representative similarity search methods and sustain the intuitions with respect to human judgements.

1. INTRODUCTION

The World Wide Web is the world's most valuable information resource and has become the world's largest database with search being the main tool that enables organisations and individuals to exploit its huge amounts of information that is freely offering (Virgilio et al., 2012). Thus, having a successful mechanism for finding and retrieving the most relevant information to a task at hand is of major importance. Traditionally, Web search is essentially based on a combination of textual keyword search with an importance ranking of the documents depending on the link structure of the Web (Fazzinga and

DOI: 10.4018/978-1-7998-0951-7.ch069

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Lukasiewicz 2010). However, one of the dimensions that has not been captured to its full extent is that of semantics. Combining search and semantics gives birth to the idea of the semantic search. Semantic search can be described in a sentence as the effort of improving the accuracy of the search process by understanding the context and limiting the ambiguity (Melo et al., 2016; Virgilio et al., 2012; Vocht et al., 2017). In fact, the issue of semantics has become the grand challenge for the next-generation World Wide Web (Jindal et al., 2014; Storey et al., 2008).

The development of semantic search is currently an extremely hot topic (Binding et al. 2015; Fazzinga and Lukasiewicz, 2010; Formica, 2012; Formica et al., 2013; Likavec et al., 2015; Melo et al., 2016; Storey et al., 2008; Virgilio et al., 2012; Vocht et al., 2017). For example, Formica (2012) showed how rough set theory (Pawlak 1991) could be employed in combination with fuzzy formal concept analysis to perform Semantic Web search and discovery of information in the web. Formica et al. (2013) pointed out that there were many proposals on semantic search and retrieval in the literature, but there was still no specific solution that clearly emerged. For this reason, they presented a proposal based on a notion that is currently gaining momentum in the field: the Information Content (IC) approach (Resnik, 1995; Jiang et al., 2017).

However, there are still some limitations in the above approaches. For example, similarity reasoning in (Formica et al., 2013) was based on ontology based IC computation (Sanchez et al., 2011). The fact that ontology based IC computation only relies on ontological knowledge is a drawback because it completely depends on the degree of coverage and detail of the unique input ontology (Sanchez and Batet, 2013). Especially, with the emergence of social networks (Martinez-Gil and Aldana-Montes 2013), a lot of (sets of) concepts or terms are not included in domain ontologies (Kumar et al., 2017), therefore, IC computation that is based on these kinds of domain ontologies cannot be used in these tasks. On the other hand, the prerequisite of ontology based IC computation is the existence of several predefined domain ontologies. Clearly, the construction of domain ontologies is time-consuming and error-prone and maintaining these ontologies also requires a lot of effort from experts. Thus, the methods of ontology based IC computation are also limited in scope and scalability (Jiang et al., 2015). These limitations of similarity reasoning are the motivation behind the new techniques presented in this paper which implement semantic search by exploiting Wikipedia (Hovy et al., 2013; Medelyan et al., 2009). It should be noted that the authors' approaches are only suitable for less specific domains because Wikipedia does not guarantee the coverage for specific domains. In the case of a very specific domain, a domain specific reference resource such as ontologies is needed. On the other hand, in the case of less specific domains, semantic search relies on some knowledge sources such as WordNet (Fellbaum, 1998), DBpedia (Lehmann et al., 2009; Li et al., 2017b), Wikipedia (Hovy et al., 2013; Medelyan et al., 2009), and YAGO (Hoffart et al., 2013; Suchanek et al., 2008). In this paper, the authors will exploit Wikipedia to implement semantic search. In fact, the authors' approaches can also be based on other knowledge sources such as DBpedia or YAGO by simple adjustment.

The purpose of this paper is to present some new methods to semantic search to solve the shortcomings of existing approaches. Concretely, along the line of the research of (Formica 2012; Formica et al. 2013), the authors will propose some novel methods to semantic search by combining formal concept analysis (Ganter and Wille, 1999), rough set theory (Pawlak, 1991), and Wikipedia based similarity reasoning (Jiang et al., 2017; Li et al., 2017a; Ponzetto and Strube, 2007; Taieb et al., 2013; Yazdani and Popescu-Belis, 2013). On the other hand, semantic search usually adopts user-friendly unstructured (Rocha et al., 2004) or semi-structured (Anyanwu et al., 2005) query strategies. However, these reported approaches are still much less attractive than keyword search, which is the most effective and successful paradigm

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

for modern information retrieval (Ning et al., 2009). Research on semantic search based on keywords has been very active and many results have been achieved (Ning et al., 2009; Zenz et al., 2009). Thus, this paper focuses on the problem of supporting effective keyword search for semantic search.

2. PRELIMINARIES

In this section the authors briefly recall some basic notions of formal concept analysis and rough sets.

2.1. Formal Concept Analysis

Formal Concept Analysis (FCA) is a theoretical framework which supplies a basis for conceptual data analysis and knowledge processing (Ganter and Wille, 1999). In this subsection the authors give the basic of FCA (Ganter and Wille, 1999; Valverde-Albacete et al., 2016) as far as they are needed for this paper.

A formal context $K = (G, M, I)$ consists of two sets G and M and a binary relation I between G and M. The elements of G are called the objects and the elements of M are called the attributes of the context. gIm or $(g, m) \in I$ can be used to express that an object g is in a relation I with an attribute m.

A small context can be easily represented by a cross table, i.e., by a rectangular table the rows of which are headed by the object names and the columns headed by the attribute names. A cross in row g and column m means that the object g has the attribute m.

For a set $A \subseteq G$ of objects we define $A' = \{m \in M | \ gIm \text{ for all } g \in A\}$. Correspondingly, for a set B of attributes we define $B' = \{g \in G | \ gIm \text{ for all } m \in B\}$.

A formal concept of the formal context $K = (G, M, I)$ is a pair (A, B) with $A \subseteq G$, $B \subseteq M$, $A' = B$ and $B' = A$. The researchers call A the extent and B the intent of the formal concept (A, B). $\mathrm{B}(G, M, I)$ (or $\mathrm{B}(K)$) denotes the set of all formal concepts of the formal context $K = (G, M, I)$, $\mathrm{Int}(K)$ and $\mathrm{Ext}(K)$ denote the sets of intents and extents of all formal concepts of the formal context $K = (G, M, I)$, respectively.

If (A_1, B_1) and (A_2, B_2) are formal concepts of the formal context $K = (G, M, I)$, (A_1, B_1) is called a subconcept of (A_2, B_2) (or (A_2, B_2) is a superconcept of (A_1, B_1)), provided that $A_1 \subseteq A_2$ (or $B_2 \subseteq B_1$) and is denoted by $(A_1, B_1) \leq (A_2, B_2)$. The relation $\leq$ is called the hierarchical order (or simply order) of the formal concepts. The set of all formal concepts of (G, M, I) ordered in this way is denoted by $\underline{\mathrm{B}}(G, M, I)$ and is called the concept lattice of the formal context (G, M, I).

2.2. Rough Sets

Rough set theory is an extension of classical set theory with additional operators, namely approximation operators. Among the various formulations that can be found in the literature, below the one given in (Jiang et al., 2009; Yao, 1996; Zhang et al., 2015) is briefly recalled.

Let U denote a finite and non-empty set called the universe, and let $\mathrm{R} \subseteq U \times U$ denote an equivalence relation on U. The pair $apr = (U, \mathrm{R})$ is called an approximation space. The equivalence relation R partitions the set U into disjoint subsets. Such a partition of the universe is denoted by U/R. If two elements x, y in U belong to the same equivalence class, the researchers say that x and y are indistinguishable. The equivalence class of R and the empty set ϕ are called the elementary sets in $apr = (U, \mathrm{R})$. Any finite union of elementary sets is called a definable set.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Given an arbitrary set $A \subseteq U$, it may not correspond to a definable set because A may include and exclude objects that belong to different definable sets. In this case, one may characterize A by a pair of lower and upper approximations:

$$\underline{apr}(A) = \bigcup_{[x]_R A} [x]_R \text{ and } \overline{apr}(A) = \bigcup_{[x]_R \cap A \neq \varnothing} [x]_R$$

where $[x]_R = \{y | \, x R y\}$, is the equivalence class containing x. The pair $\left(\underline{apr}(A), \overline{apr}(A)\right)$ is called the rough set w.r.t. A.

The notion of approximation can be naturally introduced into FCA (Formica 2012). In line with (Formica 2012; Yao and Chen 2006), given a concept lattice L, let Int(L) be the family of all the intents of L. Given a set of attributes B that is not the intent of any concept in L, intuitively the upper approximation $\overline{apr}(B)$ is the smallest set in Int(L) that contains B, whereas the lower approximation $\underline{apr}(B)$ is the largest set in Int(L) that is contained in B. Formally,

$$\overline{apr}(B) = \cap \{X | \, X \in \text{Int}(L), B \subseteq X\},$$

$$\underline{apr}(B) = \{X | \, X \in \text{Int}(L), X \subseteq B, \forall Y \in \text{Int}(L) \, (X \subseteq Y \Rightarrow Y \not\subseteq B)\}.$$

3. SEMANTIC SEARCH FOR QUERY KEYWORDS

In this section the authors will propose two novel approaches to semantic search for query keywords.

3.1. Semantic Search Using FCA and Similarity Reasoning

To implement semantic search by exploiting FCA, the authors assume that users have a formal context and that Web queries are expressed by sets of query keywords. In fact, it is similar to the assumption of (Formica 2012). From a practical point of view this is a limitation users can live with, especially taking into account that the formal contexts are used to achieve some specific domain applications, and thus, users certainly can construct a corresponding formal context by transforming the original information.

In the following the authors assume that users have a formal context $K = (G, M, I)$, where $G = \{g_1, g_2, \ldots, g_m\}$, $M = \{a_1, a_2, \ldots, a_n\}$, and I is a relation between G and M (i.e., $I \subseteq G \times M$). According to the algorithms of concept lattice construction (Ganter and Wille 1999; Godin et al. 1995; Nourine and Raynaud 2002), users can obtain the corresponding concept lattice $L = \underline{\mathrm{B}}(K)$ (or $\underline{\mathrm{B}}(G, M, I)$) w.r.t K. On the other hand, Web queries are expressed by sets of keywords, formally, a query $Q = \{q_1, q_2, \ldots, q_s\}$, where q_i $(1 \leq i \leq s)$ is a keyword. Theoretically, in an open environment such as Internet a query keyword may be any string. It is the same as the query keyword in Search Engines such as Google. For example, $Q =$ {*Semantic search*, *Ontology*, *Rough sets*} is a set of query keywords. Clearly, Q contains two multiword expressions *Semantic search* and *Rough sets*. Users can use different search interfaces such as "*Semantic search*"+"*Ontology*"+"*Rough sets*" to extract the keywords correctly. Of course, for any query string

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

such as *Semantic search Ontology Rough sets*, users need to discover the keywords correctly by means of some techniques such as automatic term extraction and association measures.

Obviously, for any query $Q = \{q_1, q_2, \ldots, q_s\}$, if there is a formal concept $(A, B) \in L$ such that $Q = B$, then users know that A is the result for Q. However, if for any formal concept $(A, B) \in L$, users have $Q \neq B$. Then, the goal is to find the formal concepts of L whose intents "better approximate" the set of attributes specified by the query and, therefore, whose extents are closer to the expected answer. In the following the authors will propose two approaches to semantic search by "better approximating" the query keywords. The first method is provided by using FCA and Wikipedia based similarity reasoning. The authors call this method the Semantic Search for query Keywords by making use of FCA and Wikipedia based similarity reasoning (SSKFW). Essentially, SSKFW is implemented by evaluating the similarity between intents of formal concepts in L and the set Q of query keywords.

Definition 1: Let $K = (G, M, I)$ be a formal context, where $M = \{a_1, a_2, \ldots, a_n\}$ and $Q = \{q_1, q_2, \ldots, q_s\}$ be a query. The semantic similarity $sim_{att}(a_i, q_j)$ between a_i and q_j $(1 \leq i \leq n, 1 \leq j \leq s)$ is defined as follows:

(1) if $a_i, q_j \in WikiCon$, $sim_{att}(a_i, q_j) = \begin{cases} 1, & if\ Synonyms\left(a_i\right) \cap Synonyms\left(q_j\right) \neq \varnothing \\ sim_{fea}\left(feature\left(a_i\right), feature\left(q_j\right)\right), & otherwise \end{cases}$ and

(2) if $a_i \notin WikiCon$ or $q_j \notin WikiCon$, $sim_{att}(a_i, q_j) = \begin{cases} 1, & if\ a_i = q_j \\ 0, & otherwise \end{cases}$,

where: (1) *WikiCon* is the set of all Wikipedia concepts.

(3) *Synonyms*(X) is the set of synonyms of X. Users can get the value of *Synonyms*(X) by using the redirect page of Wikipedia. Technically, users also may exploit the property *wikiPageRedirects* of DBpedia (Lehmann et al. 2009) to obtain *Synonyms*(X). On the other hand, users can use Wikipedia anchor texts to add the elements of *Synonyms*(X). If DBpedia is considered as the knowledge source, users can also add the elements of *Synonyms*(X) by exploiting the entity linking system of DBpedia.

(4) *feature*(X) is the set of features of X. Technically, *feature*(X) = *words*(Y), where Y is the first paragraph of text of X in Wikipedia page. Users may obtain *words*(Y) by transforming Y into a set of words.

(5) $sim_{fea}(Set_1, Set_2) = \dfrac{\left|Set_1 \cap Set_2\right|}{\left|Set_1 \cup Set_2\right|}$.

Remark 1: According to the theoretical principles and the way in which knowledge are analyzed to estimate similarity, different methods can be identified (Martinez-Gil 2014; Petrakis et al. 2006). Currently, there are lots of approaches to Wikipedia based similarity computation (Jiang et al. 2017; Li et al. 2017a; Ponzetto and Strube 2007; Taieb et al. 2013; Yazdani and Popescu-Belis 2013). In Definition 1 users can use these existing methods to compute the similarity $sim_{att}(a_i, q_j)$ between a_i and q_j. However, the computation of these methods needs more time. For example, several approaches to Wikipedia based semantic similarity computation in (Ponzetto and Strube 2007) are based on the whole Wikipedia category structure. Thus, in Definition 1 the authors present a novel approach to similarity computation. The definition of sim_{fea} is based on *X-similarity* approach (Petrakis et al. 2006). Of course, users also can use other approaches to define sim_{fea}. For example, according to R&E approach (Rodriguez and Egenhofer 2003), sim_{fea} may be defined as follows:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

$$sim_{fea}(Set_1, Set_2) = \frac{\left|Set_1 \cap Set_2\right|}{\left|Set_1 \cap Set_2\right| + \alpha\left|Set_1 \setminus Set_2\right| + (1-\alpha)\left|Set_2 \setminus Set_1\right|},$$

where $0 \leq \alpha \leq 1$ and the value of α is specified by users or experts ($\alpha = 0.5$ in default).

Since the intent $B = \{a_1, a_2, \ldots, a_t\}$ of formal concept (A, B) and the query $Q = \{q_1, q_2, \ldots, q_s\}$ are two sets of keywords, so the authors can give the notion of the similarity between B and Q by using sim_{att}. With regard to the computation of similarity for sets of keywords, the approach has been inspired by the method for evaluating concept similarity defined within the SymOntos (Formica and Missikoff 2002) and FCA (Formica 2006). It is based on the maximum weighted matching problem in bipartite graphs, that can be solved in polynomial time (Galil 1986).

Consider the intent $B = \{a_1, a_2, \ldots, a_t\}$ of some formal concept (A, B) and the query $Q = \{q_1, q_2, \ldots, q_s\}$. Let a candidate set of pairs be a subset of $B \times Q$ such that there are no two pairs in the set sharing an element. Then consider the set of all candidate sets of pairs, that is indicated as $CS(B, Q)$. For each candidate set of $CS(B, Q)$, consider the sum of the sim_{att} of the pairs of keywords. One of the candidate sets having the maximum among all the computed sums is chosen. Formally,

Definition 2: Let $B = \{a_1, a_2, \ldots, a_t\}$ be the intent of some formal concept (A, B) and $Q = \{q_1, q_2, \ldots, q_s\}$ be a query.

If $t \leq s$, the set $CS(B, Q)$ of the candidate sets of pairs is defined as follows:

$CS(B, Q) = \{\{\langle a_1, q_m\rangle, \ldots, \langle a_t, q_n\rangle\} \mid a_h \in B, q_l \in Q, \forall h = 1, \ldots, t, \forall l = m, \ldots, n, \{m, \ldots, n\} \subseteq \{1, \ldots, s\}$, and $a_i \neq a_j, q_i \neq q_j, \forall i \neq j\}$.

If $s \leq t$, the set $CS(B, Q)$ of the candidate sets of pairs is defined as follows:

$CS(B, Q) = \{\{\langle a_m, q_1\rangle, \ldots, \langle a_n, q_s\rangle\} \mid a_h \in B, q_l \in Q, \forall h = m, \ldots, n, \{m, \ldots, n\} \subseteq \{1, \ldots, t\}, \forall l = 1, \ldots, s$, and $a_i \neq a_j, q_i \neq q_j, \forall i \neq j\}$.

Definition 3: Let $K = (G, M, I)$ be a formal context and (A, B) be a formal concept of K, where $B = \{a_1, a_2, \ldots, a_t\} \subseteq M$. Given any query $Q = \{q_1, q_2, \ldots, q_s\}$, the similarity of B and Q, denoted as $sim_{FW}(B, Q)$, is defined as follows:

$$sim_{FW}(B, Q) = \frac{1}{\max(t, s)} \max_{cs \in CS(B,Q)} \left(\sum_{a_i, q_j \in cs} sim_{att}(a_i, q_j) \right).$$

Now the authors can provide the following SSKFW method in Algorithm 1.

Algorithm 1. *SSKFW method for semantic search*

Input: a formal context $K = (G, M, I)$, where $G = \{g_1, g_2, \ldots, g_m\}$, $M = \{a_1, a_2, \ldots, a_n\}$, and $I \subseteq G \times M$; a query $Q = \{q_1, q_2, \ldots, q_s\}$

Output: a result set $S \subseteq G$ for Q

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

(1) generate the concept lattice $L = \underline{B}(K)$ (or $\underline{B}(G, M, I)$) of K;
(2) if there exists a formal concept $(A, B) \in L$, where $A = \{g_1, g_2, \ldots, g_w\}$, $B = \{a_1, a_2, \ldots, a_t\}$, such that $B = Q$, then assign A to S (i.e., $S = A$) and Goto (5); elsewise (i.e., for each formal concept $(A, B) \in L$, $B \neq Q$);
(3) for each formal concept $(A, B) \in L$, compute the similarity $sim_{FW}(B, Q)$;
(4) let $S = \cup\{A \mid (A, B) \in L \wedge sim_{FW}(B, Q) = \max\{sim_{FW}(D, Q) \text{ for all } (C, D) \in L\}\}$;
(5) return S.

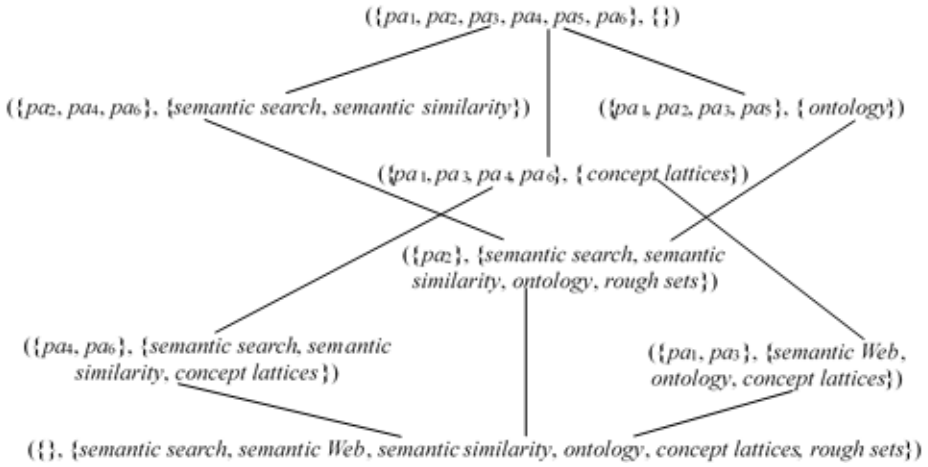
Remark 2: In Algorithm 1, the step (4) is to find the results whose corresponding intents "better approximate" the set of keywords specified by the query and, therefore, whose extents are closer to the expected answer. In fact, users can relax the condition of "better approximation". Concretely, users may specify a threshold value α ($0 \leq \alpha \leq 1$) for "better approximation". Formally, the step (4) in Algorithm 1 can be modified as follows:
(4′) let $S = \cup\{A \mid (A, B) \in L \wedge sim_{FW}(B, Q) \geq \alpha\}$.

3.2. Semantic Search Using FCA, Rough Sets and Similarity Reasoning

From the step (3) in Algorithm 1 the researchers know that users have to compute the similarity between intents of all formal concepts in L and query. In particular, when some complex approaches to Wikipedia based similarity computation such as the methods presented in (Taieb et al. 2013; Yazdani and Popescu-Belis 2013) are used in SSKFW, it will waste a lot of time for computation of $sim_{FW}(B, Q)$ for each formal concept $(A, B) \in L$. The authors will give another method by a small amount of computation. The authors call this method Semantic Search for query Keywords using FCA, Rough sets, and Wikipedia based similarity reasoning (SSKFRW).

Example 1: Suppose that users have the concept lattice L related to research papers of Figure 1 and suppose the user is looking for a paper on semantic search based on similarity reasoning. This query can be represented by $Q = \{semantic\ search, ontology, semantic\ similarity\}$.

Figure 1. Concept lattice L of research papers



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Clearly, in the concept lattice L there are no concepts defined by Q. That is, there are no answers in L for the query Q. Similarly, to the method presented in (Formica 2012), users can look for the formal concepts whose intents better approximate it. According to the definitions given in Section 2.2, the smallest intent in L containing Q is {*semantic search*, *semantic similarity*, *ontology*, *rough sets*}, i.e., the concept whose intent is an upper approximation of Q is:

($\{pa_2\}$, {*semantic search, semantic similarity, ontology, rough sets*}).

Analogously to the upper approximation of Q, lower approximations of Q can also be obtained. Two lower approximations are identified in L. They correspond to the following concepts whose intents are the largest sets contained in Q:

($\{pa_2, pa_4, pa_6\}$, {*semantic search, semantic similarity*}) and ($\{pa_1, pa_2, pa_3, pa_5\}$, {*ontology*}).

Here, users can select as answer, on the basis of their needs, one of the concept extents associated with the above upper and lower approximations. That is, users can choose $\{pa_2\}$, $\{pa_2, pa_4, pa_6\}$, or $\{pa_1, pa_2, pa_3, pa_5\}$ as the answer. Obviously, there is a problem here. Since lower approximations may not be unique, so it is still difficult to choose the answer. Naturally, users should be able to implement generation of answers from the upper and lower approximations automatically using Wikipedia based similarity computation. Intuitively, users first obtain the upper and lower approximations of the query by using rough set theory. Then users get the results of query by computing the similarity between the intents of upper and lower approximations and the query. However, there is still one problem.

Assume another query is represented by $Q' = \{$*semantic searching*, *ontologies*, *semantic relatedness*$\}$, users know that the set of formal concepts whose intents are upper or lower approximations of Q' is an empty set. Formally,

$$\{(A, B) | (A, B) \in L \wedge (\overline{apr}(Q') = B \vee \underline{apr}(Q') = B)\} = \phi.$$

Obviously, this result is not correct from a semantics point of view since there exist some semantic relationships between the intents of formal concepts in L and the query Q'. Concretely, according to Wikipedia or DBpedia users have the following:

semantic searching ≡ *semantic search, ontologies* ≡ *ontology,* and *semantic relatedness* ≡ *semantic similarity.*

Therefore, Q and Q' are equivalent from a semantics point of view. That is, the upper and lower approximations of Q' should be the same as those of Q. In SSKFRW, the authors will consider these cases.

Definition 4: A Semantically equivalent Set of Keywords based on Wikipedia (SSKW) is an unordered collection of strings $SSKW = \{key_1, key_2, \ldots, key_u\}$ such that

$\forall key_i, key_j \in SSKW, i \neq j \rightarrow key_i \not\equiv key_j$ (i.e., Synonyms(key_i)∩Synonyms(key_j) = ϕ). That is, there do not exist $key_i, key_j \in SSKW$ such that $i \neq j \wedge key_i \equiv key_j$ (i.e., Synonyms(key_i)∩Synonyms(key_j)≠ϕ).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In what follows the authors only consider semantically equivalent sets of keywords.

Definition 5: Given two semantically equivalent sets of keywords (SSKWs) $SSKW_1$ and $SSKW_2$, $SSKW_1$ is said to be a semantic subset of $SSKW_2$ (denoted by $SSKW_1 \subseteq_S SSKW_2$) if and only if the following condition holds:

$\forall key_1 \in SSKW_1$, there exists $key_2 \in SSKW_2$ satisfies $key_1 \equiv key_2$.

$SSKW_1$ and $SSKW_2$ is said to semantically equivalent (denoted by $SSKW_1 \equiv_S SSKW_2$) if and only if $SSKW_1 \subseteq_S SSKW_2$ and $SSKW_2 \subseteq_S SSKW_1$.

Given a keyword *key* and an SSKW *SSKW*, if there exists a keyword $key' \in SSKW$ such that $key \equiv key'$, *key* is called a semantic element of *SSKW*, and *SSKW* is said to semantically contain *key*.

The authors write $key \in_S SSKW$ to denote that *key* is a semantic element of *SSKW*. The notation $key \notin_S SSKW$ denotes that *key* is not a semantic element of *SSKW*.

Similarly, to $key \in_S SSKW$ (resp., $key \notin_S SSKW$), the authors can define $skey \in_S SSSKW$ (resp., $skey \notin_S SSSKW$), where *skey* is a set of keywords and *SSSKW* is a set of *SSKW*.

Definition 6: Given a concept lattice L and a set of attributes (i.e., keywords) B, $B \notin_S \text{Int}(L)$, the semantic upper and lower approximations of B, $\overline{apr}(B)$ and $\underline{apr}(B)$, are defined as follows, respectively:

$$\overline{apr}(B) = \cap\{X \mid X \in \text{Int}(L), B \subseteq_S X\},$$

$$\underline{apr}(B) = \{X \mid X \in \text{Int}(L), X \subseteq_S B, \forall Y \in \text{Int}(L)\ (X \subseteq Y \Rightarrow Y \not\subseteq_S B)\}.$$

Now the authors provide the SSKFRW method.

Algorithm 2. *SSKFRW method for semantic search*

Input: a formal context $K = (G, M, I)$, where $G = \{g_1, g_2, \ldots, g_m\}$, $M = \{a_1, a_2, \ldots, a_n\}$, and $I \subseteq G \times M$; a query $Q = \{q_1, q_2, \ldots, q_s\}$

Output: a result set $S \subseteq G$ for Q

(1) construct the concept lattice $L = \underline{\mathrm{B}}(K)$ (or $\underline{\mathrm{B}}(G, M, I)$) of K;
(2) if there exists a formal concept $(A, B) \in L$, where $A = \{g_1, g_2, \ldots, g_w\}$, $B = \{a_1, a_2, \ldots, a_t\}$, such that $B \equiv_S Q$, then assign A to S (i.e., $S = A$) and Goto (9); elsewise (i.e., for each formal concept $(A, B) \in L$, $B \not\equiv_S Q$);
(3) compute the semantic upper approximation *UA* and the set of semantic lower approximations *LA* of Q;
(4) assign $\{UA\} \cup LA$ to *CS*, i.e., $CS = \{UA\} \cup LA$;
(5) for each intent $int \in CS$, obtain the corresponding formal concept (ext, int) in L, i.e., $(ext, int) \in L$;
(6) let $FC = \{(ext, int) \mid (ext, int) \in L \wedge int \in CS\}$;
(7) for each formal concept $(A, B) \in FC$, compute the similarity $sim_{FW}(B, Q)$;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

(8) let $S = \cup\{A | (A, B) \in FC \wedge sim_{FW}(B, Q) = \max\{sim_{FW}(D, Q) \text{ for all } (C, D) \in FC\}\}$;
(9) return S.

Remark 3: Similarly, to Remark 2, users can specify a threshold value α ($0 \leq \alpha \leq 1$) in step (8) of Algorithm 2. Formally, the step (8) in Algorithm 2 can be modified as follows:
(8′) let $S = \cup\{A | (A, B) \in FC \wedge sim_{FW}(B, Q) \geq \alpha\}$.

Compared to Algorithm 1, although users need to spend time to calculate the semantic upper and lower approximations in Algorithm 2, however, users only need to compute the similarity between intents of part of formal concepts in L (i.e., the upper and lower approximations of Q) and query, respectively in step (7) of Algorithm 2 (Algorithm 1 needs to compute the similarity between the query and the intents for each formal concept), thus, SSKFRW is an improvement of SSKFW in computation time.

4. EVALUATION

To validate the effectiveness of SSKFW and SSKFRW, in this section the authors will use real-world data in their experiments. It is well known that evaluating semantic search is a problem that has been studied in the literature (Elbedweihy et al. 2015; Fazzinga and Lukasiewicz 2010; Formica 2012; Formica et al. 2013; Storey et al. 2008; Virgilio et al. 2012). However, these semantic search methods have inherently different underlying assumptions, so the authors cannot perform benchmarks and comparative results for their approaches and the existing proposals. In fact, Formica (Formica 2012) analyzed various representative papers on the basis of the different topics addressed in semantic (Web) search and pointed out that an evaluation of these works is difficult to be performed. Similarly to evaluation method in (Formica et al. 2013), the authors evaluate their approaches by comparing SSKFW, SSKFRW, against the Dice, Jaccard, Cosine, and the Weighted Sum, which are some of the most representative proposals defined in the literature and are defined as follows (for any two sets $X = \{X_1, \ldots, X_m\}$ and $Y = \{Y_1, \ldots, Y_n\}$ of keywords):

$$Dice(X, Y) = \frac{2\left|X \cap Y\right|}{\left|X\right| + \left|Y\right|},$$

$$Jaccard(X, Y) = \frac{\left|X \cap Y\right|}{\left|X \cup Y\right|},$$

$$Cosine(X, Y) = \frac{\left|X \cap Y\right|}{\left|X\right| \times \left|Y\right|}, \text{ and}$$

$$WeightedSum(X, Y) = \frac{2\sum Aff\left(X_i, Y_j\right)}{\left|X\right| + \left|Y\right|},$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

where the affinity of X_i and Y_j is $Aff(X_i, Y_j)$ is:

$Aff(X_i, Y_j) = 1$ if $X_i = Y_j$; $= 0.5$

if X_i is broader or narrower concept of Y_j; $= 0$ otherwise.

The authors also compare SSKFW and SSKFRW against human judgment (HJ). Similarly to HJ evaluation in (Formica et al. 2013), in order to evaluate HJ of the authors' experiments, they ask to a group of 10 people, randomly selected from their Institute staff, to indicate the similarity among each query Q_i ($1 \leq i \leq 4$, see Table 1) and the real-world research papers pa_j ($1 \leq j \leq 30$, see Table 2), defined in their keywords. The semantic similarity between each paper and each query is assessed by 10 people in a scale between 0 (semantically unrelated) and 1 (highly synonymous), respectively. Thus, four final sets of paper and query pairs are rated with the average of the similarity values provided by their Institute staff and are shown in Tables 3−6.

More concretely, the authors consider four query, namely Q_i ($1 \leq i \leq 4$), representing the query request of four research papers, shown in Table 1. The authors also consider 30 research papers, namely pa_j ($1 \leq j \leq 30$), representing the digital resources by using different keywords, shown in Table 2. The HJ values, and the results obtained according to the above-mentioned similarity methods (Dice, Jaccard, and Cosine) and their approaches (SSKFW and SSKFRW), are shown in Tables 3−6, corresponding to the experimental results about Q_i ($1 \leq i \leq 4$), respectively.

Now the authors discuss the precision and recall of the approaches SSKFW, SSKFRW, the Dice, Jaccard, Cosine, and the Weighted Sum. In Tables 7−10, the digital resources are listed according to the values assigned by HJ, Weighted Sum, and Dice from the highest up to the threshold (the authors fix the threshold to 0.50), and the results returned by SSKFW and SSKFRW. Note that Jaccard and Cosine have not been addressed because their values are all less than the given threshold. The authors apply the typical evaluation measures, namely precision and recall, to the returned digital resources (the definitions of recall and precision in Weighted Sum and Dice are the same as traditional information retrieval). Assume that the ranked papers of HJ are $\{pa_1, pa_2, \ldots, pa_n\}$ and the relation of HJ values of the ranked papers is $HJ(pa_1) \geq HJ(pa_2) \geq \ldots \geq HJ(pa_n)$, the returned papers of SSKFW (or SSKFRW) are $\{pa'_1, pa'_2, \ldots, pa'_m\}$. If users use max function in Algorithm 1 (or Algorithm 2), recall and precision are defined as follows:

$$recall = \begin{cases} \dfrac{\left|\left\{pa_1, pa_2, \ldots, pa_m\right\} \cap \left\{pa'_1, pa'_2, \ldots, pa'_m\right\}\right|}{m}, & if\ n \geq m \\ \dfrac{\left|\left\{pa_1, pa_2, \ldots, pa_n\right\} \cap \left\{pa'_1, pa'_2, \ldots, pa'_m\right\}\right|}{n}, & if\ n < m \end{cases} \quad \text{and}$$

$$precision = \begin{cases} \dfrac{\left|\left\{pa_1, pa_2, \ldots, pa_m\right\} \cap \left\{pa'_1, pa'_2, \ldots, pa'_m\right\}\right|}{m}, & if\ n \geq m \\ \dfrac{\left|\left\{pa_1, pa_2, \ldots, pa_n\right\} \cap \left\{pa'_1, pa'_2, \ldots, pa'_m\right\}\right|}{m}, & if\ n < m \end{cases}.$$

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Query keywords

q_1: {*Ontologies*, *Concept lattice*, *Rough set*}
q_2: {*Ontologies*, *Semantic Web*, *Semantic relatedness*}
q_3: {*Description logics*, *Ontologies*}
q_4: {*Information content*, *Semantic similarity*}

Table 2. Annotation of research papers

pa_1: {*Information content*, *Semantic similarity*, *Ontologies*, *MeSH*, *SNOMED CT*}
pa_2: {*Semantic similarity*, *Semantic relatedness*, *Ontologies*, *Feature-based similarity*, *WordNet*}
pa_3: {*Formal Concept Analysis*, *Semantic Web*, *Domain ontologies*, *Similarity reasoning*}
pa_4: {*Semantic Web*, *Formal concept analysis*, *Fuzzy information*, *Rough set theory*, *Formal concept*}
pa_5: {*Similarity reasoning*, *Weighted Reference Ontology*, *Information content*, *Digital resources*}
pa_6: {*Web search*, *Semantic search*, *RDF*, *Semantic Web*, *Linked data*}
pa_7: {*Semantic Web search*, *Ontologies*, *Description logics*, *Conjunctive queries*, *Semantic Web*, *Web search*}
pa_8: {*Natural language queries*, *Semantic relatedness*, *Vocabulary independent queries*, *RDF*, *Linked data*}
pa_9: {*Web of data*, *Linked data*, *Knowledge extraction*, *Wikipedia*, *RDF*}
pa_{10}: {*Provenance*, *Linked data*, *DBpedia*, *Wikipedia*}
pa_{11}: {*Formal concept analysis*, *Semantic Web*, *Information content*, *Similarity reasoning*}
pa_{12}: {*Semantic search*, *Ranking*, *Semantic relationship*, *Ontologies*, *Semantic Web*}
pa_{13}: {*Semantic similarity*, *Multiple ontologies*, *MeSH*, *WordNet*}
pa_{14}: {*Semantic similarity*, *Evolutionary computation*, *Semantic Web*, *Synonym recognition*, *Differential evolution*}
pa_{15}: {*Knowledge representation*, *Description logics*, *Structured objects*, *Ontologies*}
pa_{16}: {*Description logics*, *Rough description logics*, *Rough set theory*, *Concepts*, *Approximate concepts*}
pa_{17}: {*Semantic search*, *Query logs*, *Information retrieval*, *Relevance feedback*, *Evaluation*, *Linked data*}
pa_{18}: {*eScience*, *Ontologies*, *Linked data*, *Grid*, *Big data*}
pa_{19}: {*Granular computing*, *Granular information*, *Ontologies*, *Description logic*, *Rough set theory*, *Ontologies editor*}
pa_{20}: {*Semantic relatedness*, *Wikipedia*, *Wikipedia category graph*, *Word relatedness*, *Semantic analysis*}
pa_{21}: {*Wikipedia*, *Toolkit*, *Ontology extraction*, *Semantic relatedness*, *Disambiguation*, *Annotation*}
pa_{22}: {*Rough set theory*, *Formal concept analysis*, *Data mining*, *Choosing behaviours*, *Personal investment portfolios*}
pa_{23}: {*Granular computing*, *Formal concept analysis*, *Rough set*, *Diachronism*, *Knowledge system*}
pa_{24}: {*Rough set*, *Concept lattice*, *Reduction*, *Consistent set*, *Formal concept analysis*}
pa_{25}:{*Semantic Web*, *Ontologies*, *Description logics*, *Matchmaking*}
pa_{26}: {*Information content*, *Semantic similarity*, *Ontologies*, *Taxonomic knowledge*, *WordNet*}
pa_{27}: {*Semantic similarity*, *Ontologies*, *Product information retrieval*, *Formal concept analysis*}
pa_{28}: {*Ontology learning*, *Semantic role labeling*, *Information extraction*, *Semantic Web*, *Ontologies*}
pa_{29}: {*Semantic Web*, *Keyword-based*, *Query*, *Semantic indexing*, *Ontologies*}
pa_{30}: {*Personalized information retrieval*, *Semantic Web*, *Ontologies*, *Reasoning*, *SPARQL*}

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Experimental results about q_1

Papers	HJ	SSKFW		SSKFRW		Weighted Sum	Dice	Jaccard	Cosine
		max	$\alpha = 0.5$	max	$\alpha = 0.5$				
pa_1	0.22					0.38	0.25	0.14	0.07
pa_2	0.25					0.38	0.25	0.14	0.07
pa_3	0.10					0.29	0.00	0.00	0.00
pa_4	0.83	√	√	√	√	0.38	0.00	0.00	0.00
pa_5	0.10					0.13	0.00	0.00	0.00
pa_6	0.10					0.14	0.00	0.00	0.00
pa_7	0.30					0.33	0.25	0.14	0.07
pa_8	0.10					0.00	0.00	0.00	0.00
pa_9	0.10					0.00	0.00	0.00	0.00
pa_{10}	0.10					0.00	0.00	0.00	0.00
pa_{11}	0.36					0.29	0.00	0.00	0.00
pa_{12}	0.30					0.38	0.25	0.14	0.07
pa_{13}	0.15					0.29	0.00	0.00	0.00
pa_{14}	0.16					0.25	0.00	0.00	0.00
pa_{15}	0.22					0.29	0.29	0.17	0.08
pa_{16}	0.40					0.25	0.00	0.00	0.00
pa_{17}	0.16					0.00	0.00	0.00	0.00
pa_{18}	0.20					0.25	0.29	0.17	0.08
pa_{19}	0.60					0.33	0.22	0.13	0.06
pa_{20}	0.10					0.00	0.00	0.00	0.00
pa_{21}	0.10					0.11	0.00	0.00	0.00
pa_{22}	058	√	√	√	√	0.25	0.00	0.00	0.00
pa_{23}	0.62	√	√	√	√	0.50	0.25	0.14	0.07
pa_{24}	0.78	√	√	√	√	0.63	0.50	0.33	0.13
pa_{25}	0.38					0.57	0.29	0.17	0.08
pa_{26}	0.20					0.38	0.25	0.14	0.07
pa_{27}	0.75		√			0.43	0.29	0.17	0.08
pa_{28}	0.23					0.38	0.25	0.14	0.07
pa_{29}	0.20					0.38	0.29	0.17	0.08
pa_{30}	0.25					0.38	0.25	0.14	0.07

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Experimental results about q_2

Papers	HJ	SSKFW		SSKFRW		Weighted Sum	Dice	Jaccard	Cosine
		max	$\alpha = 0.5$	max	$\alpha = 0.5$				
pa_1	0.73	√	√	√	√	0.38	0.25	0.14	0.07
pa_2	0.82	√	√	√	√	0.63	0.50	0.33	0.13
pa_3	0.38					0.43	0.00	0.00	0.00
pa_4	0.40					0.25	0.25	0.14	0.07
pa_5	0.15					0.14	0.00	0.00	0.00
pa_6	0.25					0.29	0.29	0.17	0.08
pa_7	0.55		√	√	√	0.44	0.44	0.29	0.11
pa_8	0.20					0.38	0.25	0.14	0.07
pa_9	0.10					0.13	0.00	0.00	0.00
pa_{10}	0.10					0.14	0.00	0.00	0.00
pa_{11}	0.22					0.43	0.29	0.17	0.08
pa_{12}	0.65		√	√	√	0.63	0.50	0.33	0.13
pa_{13}	0.33					0.14	0.00	0.00	0.00
pa_{14}	0.51					0.38	0.25	0.14	0.07
pa_{15}	0.20					0.29	0.29	0.17	0.08
pa_{16}	0.10					0.13	0.00	0.00	0.00
pa_{17}	0.11					0.11	0.00	0.00	0.00
pa_{18}	0.23					0.38	0.25	0.14	0.07
pa_{19}	0.25					0.33	0.22	0.13	0.06
pa_{20}	0.20					0.25	0.25	0.14	0.07
pa_{21}	0.30					0.22	0.25	0.14	0.07
pa_{22}	0.10					0.13	0.00	0.00	0.00
pa_{23}	0.10					0.13	0.00	0.00	0.00
pa_{24}	0.15					0.13	0.00	0.00	0.00
pa_{25}	0.61		√	√	√	0.57	0.57	0.40	0.17
pa_{26}	0.71	√	√	√	√	0.38	0.25	0.14	0.07
pa_{27}	0.69		√	√	√	0.43	0.29	0.17	0.08
pa_{28}	0.53		√	√	√	0.50	0.50	0.33	0.13
pa_{29}	0.55		√	√	√	0.50	0.50	0.33	0.13
pa_{30}	0.55		√	√	√	0.50	0.50	0.33	0.13

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 5. Experimental results about q_3

Papers	HJ	SSKFW		SSKFRW		Weighted Sum	Dice	Jaccard	Cosine
		max	$\alpha = 0.5$	max	$\alpha = 0.5$				
pa_1	0.52		√			0.29	0.29	0.17	0.10
pa_2	0.52		√			0.29	0.29	0.17	0.10
pa_3	0.28					0.17	0.00	0.00	0.00
pa_4	0.20					0.14	0.00	0.00	0.00
pa_5	0.18					0.17	0.00	0.00	0.00
pa_6	0.10					0.17	0.00	0.00	0.00
pa_7	0.83	√	√	√	√	0.50	0.50	0.33	0.17
pa_8	0.10					0.00	0.00	0.00	0.00
pa_9	0.10					0.00	0.00	0.00	0.00
pa_{10}	0.10					0.00	0.00	0.00	0.00
pa_{11}	0.15					0.17	0.00	0.00	0.00
pa_{12}	0.50		√			0.29	0.29	0.17	0.10
pa_{13}	0.32					0.17	0.00	0.00	0.00
pa_{14}	0.20					0.14	0.00	0.00	0.00
pa_{15}	0.95	√	√	√	√	0.67	0.67	0.50	0.25
pa_{16}	0.62		√			0.43	0.29	0.17	0.10
pa_{17}	0.13					0.00	0.00	0.00	0.00
pa_{18}	0.65		√			0.29	0.29	0.17	0.10
pa_{19}	0.80		√			0.38	0.25	0.17	0.08
pa_{20}	0.10					0.00	0.00	0.00	0.00
pa_{21}	0.17					0.13	0.00	0.00	0.00
pa_{22}	0.15					0.14	0.00	0.00	0.00
pa_{23}	0.15					0.14	0.00	0.00	0.00
pa_{24}	0.15					0.14	0.00	0.00	0.00
pa_{25}	0.90	√	√	√	√	0.67	0.67	0.17	0.25
pa_{26}	0.52		√			0.29	0.29	0.17	0.10
pa_{27}	0.55		√			0.33	0.33	0.20	0.13
pa_{28}	0.44		√			0.29	0.29	0.17	0.10
pa_{29}	0.52		√			0.29	0.29	0.17	0.10
pa_{30}	0.48		√			0.29	0.29	0.17	0.10

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 6. Experimental results about q_4

Papers	HJ	SSKFW		SSKFRW		Weighted Sum	Dice	Jaccard	Cosine
		max	$\alpha = 0.5$	max	$\alpha = 0.5$				
pa_1	0.95	√	√	√	√	0.57	0.57	0.40	0.20
pa_2	0.50		√		√	0.43	0.29	0.17	0.10
pa_3	0.33					0.33	0.00	0.00	0.00
pa_4	0.10					0.00	0.00	0.00	0.00
pa_5	0.70		√		√	0.50	0.33	0.20	0.13
pa_6	0.22					0.17	0.00	0.00	0.00
pa_7	0.16					0.13	0.00	0.00	0.00
pa_8	0.61		√		√	0.14	0.00	0.00	0.00
pa_9	0.10					0.00	0.00	0.00	0.00
pa_{10}	0.10					0.00	0.00	0.00	0.00
pa_{11}	0.65		√		√	0.50	0.33	0.20	0.13
pa_{12}	0.42					0.29	0.00	0.00	0.00
pa_{13}	0.62		√		√	0.33	0.33	0.20	0.13
pa_{14}	0.53		√		√	0.29	0.29	0.17	0.10
pa_{15}	0.11					0.00	0.00	0.00	0.00
pa_{16}	0.10					0.00	0.00	0.00	0.00
pa_{17}	0.13					0.00	0.00	0.00	0.00
pa_{18}	0.10					0.00	0.00	0.00	0.00
pa_{19}	0.10					0.00	0.00	0.00	0.00
pa_{20}	0.60		√		√	0.14	0.00	0.00	0.00
pa_{21}	0.58		√		√	0.13	0.00	0.00	0.00
pa_{22}	0.10					0.00	0.00	0.00	0.00
pa_{23}	0.12					0.00	0.00	0.00	0.00
pa_{24}	0.15					0.00	0.00	0.00	0.00
pa_{25}	0.25					0.17	0.00	0.00	0.00
pa_{26}	0.95	√	√	√	√	0.57	0.57	0.40	0.20
pa_{27}	0.51		√		√	0.33	0.33	0.20	0.13
pa_{28}	0.47					0.14	0.00	0.00	0.00
pa_{29}	0.39					0.14	0.00	0.00	0.00
pa_{30}	0.38					0.14	0.00	0.00	0.00

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

If users use threshold in Algorithm 1 (or Algorithm 2), recall is defined as follows:

$$recall = \frac{\left|\left\{pa_1, pa_2, \ldots, pa_n\right\} \cap \left\{pa_1^{'}, pa_2^{'}, \ldots, pa_m^{'}\right\}\right|}{n} \text{ and}$$

precision is defined as follows:

$$precision = \frac{\left|\left\{pa_1, pa_2, \ldots, pa_n\right\} \cap \left\{pa_1^{'}, pa_2^{'}, \ldots, pa_m^{'}\right\}\right|}{m}.$$

From Tables 3−10 users know that the precision of SSKFW and SSKFRW approaches in most cases is equal to 1 except for three cases (0.75, SSKFW for q_1 using max; 0.75, SSKFRW for q_1 using max; and 0.86, SSKFW for q_3 using threshold). Similarly, the precision of Weighted Sum and Dice in most cases is also equal to 1 except for one case (0.67, Weighted Sum for q_1). This shows the fact that all methods (SSKFW, SSKFRW, Weighted Sum, and Dice) always can retrieve the relevant resources correctly.

Concerning the recall, the recall of SSKFW and SSKFRW in most cases is greater than or equal to 0.67 except for one case (0.25, SSKFRW for q_3 using threshold). Especially, for SSKFW and SSKFRW with max, the recall of 6 cases (a total of 8 cases) is equal to 1; for SSKFW and SSKFRW with threshold, the recall of 6 cases (a total of 8 cases) is greater than 0.80. However, the best recall of Weighted Sum and Dice is 0.55 (Weighted Sum and Dice for q_2). In particular, for Weighted Sum and Dice, the recall of 6 cases (a total of 8 cases) is less than 0.40, where the worst recall is 0.17 (0.18) in the case of Dice for q_1 (q_4). Clearly, the recall of SSKFW and SSKFRW approaches is always greater than the recall of Weighted Sum and Dice.

There are two points that the authors have to point out here. Firstly, regarding the computation time and space complexity, SSKFW and SSKFRW must be greater than Dice, Jaccard, Cosine, and Weighted Sum, because SSKFW and SSKFRW implement search on the semantic level, thus, SSKFW and SSKFRW need more computing than Dice, Jaccard, Cosine, and Weighted Sum. In fact, all approaches for

Table 7. Precision and recall for q_1

HJ		SSKFW		SSKFRW		Weighted Sum		Dice	
Ranked papers	Values	Returned papers		Returned papers		Ranked papers	Values	Ranked papers	Values
		max	$\alpha = 0.5$	max	$\alpha = 0.5$				
pa_4	0.83	pa_4	pa_4	pa_4	pa_4	pa_{23}	0.50	pa_{24}	0.50
pa_{24}	0.78	pa_{22}	pa_{22}	pa_{22}	pa_{22}	pa_{24}	0.63		
pa_{27}	0.75	pa_{23}	pa_{23}	pa_{23}	pa_{23}	pa_{25}	0.57		
pa_{23}	0.62	pa_{24}	pa_{24}	pa_{24}	pa_{24}				
pa_{19}	0.60		pa_{27}						
pa_{22}	0.58								
		Precision = 0.75	Precision = 1.00	Precision = 0.75	Precision = 1.00	Precision = 0.67		Precision = 1.00	
		Recall = 0.75	Recall = 0.83	Recall = 0.75	Recall = 0.67	Recall = 0.33		Recall = 0.17	

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 8. Precision and recall for q_2

HJ		SSKFW		SSKFRW		Weighted Sum		Dice	
Ranked papers	Values	Returned papers		Returned papers		Ranked papers	Values	Ranked papers	Values
		max	$\alpha = 0.5$	max	$\alpha = 0.5$				
pa_2	0.82	pa_1	pa_1	pa_1	pa_1	pa_2	0.63	pa_2	0.50
pa_1	0.73	pa_2	pa_2	pa_2	pa_2	pa_{12}	0.63	pa_{12}	0.50
pa_{26}	0.71	pa_{26}	pa_7	pa_7	pa_7	pa_{25}	0.57	pa_{25}	0.57
pa_{27}	0.69		pa_{12}	pa_{12}	pa_{12}	pa_{28}	0.50	pa_{28}	0.50
pa_{12}	0.65		pa_{25}	pa_{25}	pa_{25}	pa_{29}	0.50	pa_{29}	0.50
pa_{25}	0.61		pa_{26}	pa_{26}	pa_{26}	pa_{30}	0.50	pa_{30}	0.50
pa_7	0.55		pa_{27}	pa_{27}	pa_{27}				
pa_{29}	0.55		pa_{28}	pa_{28}	pa_{28}				
pa_{30}	0.55		pa_{29}	pa_{29}	pa_{29}				
pa_{28}	0.53		pa_{30}	pa_{30}	pa_{30}				
pa_{14}	0.51								
		Precision = 1.00	Precision = 1.00	Precision = 1.00	Precision = 1.00	Precision = 1.00		Precision = 1.00	
		Recall = 1.00	Recall = 0.91	Recall = 1.00	Recall = 0.91	Recall = 0.55		Recall = 0.55	

Table 9. Precision and recall for q_3

HJ		SSKFW		SSKFRW		Weighted Sum		Dice	
Ranked papers	Values	Returned papers		Returned papers		Ranked papers	Values	Ranked papers	Values
		max	$\alpha = 0.5$	max	$\alpha = 0.5$				
pa_{15}	0.95	pa_7	pa_1	pa_7	pa_7	pa_7	0.50	pa_7	0.50
pa_{25}	0.90	pa_{15}	pa_2	pa_{15}	pa_{15}	pa_{15}	0.67	pa_{15}	0.67
pa_7	0.83	pa_{25}	pa_7	pa_{25}	pa_{25}	pa_{25}	0.67	pa_{25}	0.67
pa_{19}	0.80		pa_{12}						
pa_{18}	0.65		pa_{15}						
pa_{16}	0.62		pa_{16}						
pa_1	0.52		pa_{18}						
pa_2	0.52		pa_{19}						
pa_{26}	0.52		pa_{25}						
pa_{27}	0.55		pa_{26}						
pa_{29}	0.52		pa_{27}						
pa_{12}	0.50		pa_{28}						
			pa_{29}						
			pa_{30}						
		Precision = 1.00	Precision = 0.86	Precision = 1.00	Precision = 1.00	Precision = 1.00		Precision = 1.00	
		Recall = 1.00	Recall = 1.00	Recall = 1.00	Recall = 0.25	Recall = 0.25		Recall = 0.25	

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 10. Precision and recall for q_4

HJ		SSKFW		SSKFRW		Weighted Sum		Dice	
Ranked papers	Values	Returned papers		Returned papers		Ranked papers	Values	Ranked papers	Values
		max	$\alpha = 0.5$	max	$\alpha = 0.5$				
pa_1	0.95	pa_1	pa_1	pa_1	pa_1	pa_1	0.57	pa_1	0.57
pa_{26}	0.95	pa_{26}	pa_2	pa_{26}	pa_2	pa_5	0.50	pa_{26}	0.57
pa_5	0.70		pa_5		pa_5	pa_{11}	0.50		
pa_{11}	0.65		pa_8		pa_8	pa_{26}	0.57		
pa_{13}	0.62		pa_{11}		pa_{11}				
pa_8	0.61		pa_{13}		pa_{13}				
pa_{20}	0.60		pa_{14}		pa_{14}				
pa_{21}	0.58		pa_{20}		pa_{20}				
pa_{14}	0.53		pa_{21}		pa_{21}				
pa_{27}	0.51		pa_{26}		pa_{26}				
pa_2	0.50		pa_{27}		pa_{27}				
		Precision = 1.00	Precision = 1.00	Precision = 1.00	Precision = 1.00	Precision = 1.00		Precision = 1.00	
		Recall = 1.00	Recall = 1.00	Recall = 1.00	Recall = 1.00	Recall = 0.36		Recall = 0.18	

semantic search are faced with this problem. From Algorithms 1 and 2 users know that SSKFW and SSKFRW are feasible in practical applications because concept lattice can be generated off line and users may select different approaches to similarity computation for different applications. Secondly, SSKFW and SSKFRW need to use Wikipedia to implement semantic search. However, the content of Wikipedia is constantly changing. Generally speaking, the content is getting better and better, thus, with the development of Wikipedia content, SSKFW and SSKFRW can get better results.

5. CONCLUSION

In this paper the authors propose two novel methods to semantic search by combining formal concept analysis, rough set theory, and Wikipedia based similarity reasoning. The experimental results show that the authors' proposals perform better than some of the most representative similarity search methods and sustain the intuitions with respect to human judgements. As future work, the authors are planning to investigate the approaches to unambiguous Wikipedia concepts selection by automatic disambiguation methods and to query keywords discovery by means of automatic term extraction or association measures. In addition, the authors plan to validate the SSKFW and SSKFRW approaches in some application domains such as social networks.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ACKNOWLEDGMENT

The authors would like to thank the anonymous referees and Editor in Chief (Professor Miltiadis D. Lytras) for their valuable comments and helpful suggestions which greatly improved the exposition of the paper. The works described in this paper are supported by The National Natural Science Foundation of China under Grant Nos. 61772210 and 61272066; The key project in universities in Guangdong Province of China under Grant No. 2016KZDXM024; The Innovation project of postgraduate education in Guangdong Province of China under Grant No. 2016SFKC_13; The Natural Science Foundation of Guangdong Province of China under Grant No. S2012030006242.

REFERENCES

Anyanwu, K., Maduko, A., & Sheth, A. (2005). SemRank: ranking complex relationship search results on the semantic web. In *Proceedings of the 14th international conference on World Wide Web (WWW 2005)* (pp. 117-127). 10.1145/1060745.1060766

Binding, C., Charno, M., Jeffrey, S., May, K., & Tudhope, D. (2015). Template Based Semantic Integration: From Legacy Archaeological Datasets to Linked Data. *International Journal on Semantic Web and Information Systems*, *11*(1), 1–29. doi:10.4018/IJSWIS.2015010101

Elbedweihy, K. M., Wrigley, S. N., Clough, P., & Ciravegna, F. (2015). An overview of semantic search evaluation initiatives. *Journal of Web Semantics*, *30*, 82–105. doi:10.1016/j.websem.2014.10.001

Fazzinga, B., & Lukasiewicz, T. (2010). Semantic Search on the Web. *Semantic Web – Interoperability, Usability*. *Applicability*, *1*.

Fellbaum, C. (1998). *WordNet: An Electronic Lexical Database*. Cambridge, MA: Academic Press.

Formica, A. (2006). Ontology-based concept similarity in Formal Concept Analysis. *Information Sciences*, *176*(18), 2624–2641. doi:10.1016/j.ins.2005.11.014

Formica, A. (2012). Semantic Web search based on rough sets and Fuzzy Formal Concept Analysis. *Knowledge-Based Systems*, *26*, 40–47. doi:10.1016/j.knosys.2011.06.018

Formica, A., & Missikoff, M. (2002). Concept similarity in SymOntos: An enterprise ontology management tool. *The Computer Journal*, *45*(6), 583–594. doi:10.1093/comjnl/45.6.583

Formica, A., Missikoff, M., Pourabbas, E., & Taglino, F. (2013). Semantic search for matching user requests with profiled enterprises. *Computers in Industry*, *64*(3), 191–202. doi:10.1016/j.compind.2012.09.007

Galil, Z. (1986). Efficient algorithms for finding maximum matching in graphs. *ACM Computing Surveys*, *18*(1), 23–38. doi:10.1145/6462.6502

Ganter, B., & Wille, R. (1999). *Formal Concept Analysis: Mathematical Foundations*. Berlin, Germany: Springer-Verlag. doi:10.1007/978-3-642-59830-2

Godin, R., Missaoui, R., & Alaoui, H. (1995). Incremental concept formation algorithms based on Galois (concept) lattices. *Computational Intelligence*, *11*(2), 246–267. doi:10.1111/j.1467-8640.1995.tb00031.x

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hoffart, J., Suchanek, F. M., Berberich, K., & Weikum, G. (2013). YAGO2: A spatially and temporally enhanced knowledge base from Wikipedia. *Artificial Intelligence*, *194*, 28–61. doi:10.1016/j.artint.2012.06.001

Hovy, E., Navigli, R., & Ponzetto, S. P. (2013). Collaboratively built semi-structured content and Artificial Intelligence: The story so far. *Artificial Intelligence*, *194*, 2–27. doi:10.1016/j.artint.2012.10.002

Jiang, Y., Bai, W., Zhang, X., & Hu, J. (2017). Wikipedia-based information content and semantic similarity computation. *Information Processing & Management*, *53*(1), 248–265. doi:10.1016/j.ipm.2016.09.001

Jiang, Y., Wang, J., Tang, S., & Xiao, B. (2009). Reasoning with rough description logics: An approximate concepts approach. *Information Sciences*, *179*(5), 600–612. doi:10.1016/j.ins.2008.10.021

Jiang, Y., Zhang, X., Tang, Y., & Nie, R. (2015). Feature-based approaches to semantic similarity assessment of concepts using Wikipedia. *Information Processing & Management*, *51*(3), 215–234. doi:10.1016/j.ipm.2015.01.001

Jindal, V., Bawa, S., & Batra, S. (2014). A review of ranking approaches for semantic search on Web. *Information Processing & Management*, *50*(2), 416–425. doi:10.1016/j.ipm.2013.10.004

Kumar, S., Baliyan, N., & Sukalikar, S. (2017). Ontology cohesion and coupling metrics. *International Journal on Semantic Web and Information Systems*, *13*(4), 1–26. doi:10.4018/IJSWIS.2017100101

Lehmann, J., Bizer, C., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R., & Hellmann, S. (2009). DBpedia — a crystallization point for the web of data. *Journal of Web Semantics*, *7*(3), 154–165. doi:10.1016/j.websem.2009.07.002

Li, P., Jiang, Y., Wang, J., & Yin, Z. (2017a). Semantic extension of query for the Linked Data. *International Journal on Semantic Web and Information Systems*, *13*(4), 109–133. doi:10.4018/IJSWIS.2017100106

Li, P., Xiao, B., Ma, W., Jiang, Y., & Zhang, Z. (2017b). A graph-based semantic relatedness assessment method combining Wikipedia features. *Engineering Applications of Artificial Intelligence*, *65*, 268–281. doi:10.1016/j.engappai.2017.07.027

Likavec, S., Osborne, F., & Cena, F. (2015). Property-based Semantic Similarity and Relatedness for Improving Recommendation Accuracy and Diversity. *International Journal on Semantic Web and Information Systems*, *11*(4), 1–40. doi:10.4018/IJSWIS.2015100101

Martinez-Gil, J. (2014). An overview of textual semantic similarity measures based on web intelligence. *Artificial Intelligence Review*, *42*(4), 935–943. doi:10.100710462-012-9349-8

Martinez-Gil, J., & Aldana-Montes, J. F. (2013). Semantic similarity measurement using historical google search patterns. *Information Systems Frontiers*, *15*(3), 399–410. doi:10.100710796-012-9404-7

Medelyan, O., Milne, D., Legg, C., & Witten, I. H. (2009). Mining meaning from Wikipedia. *International Journal of Human-Computer Studies*, *67*(9), 716–754. doi:10.1016/j.ijhcs.2009.05.004

Melo, D., Rodrigues, I. P., & Nogueira, V. B. (2016). Using a Dialogue Manager to Improve Semantic Web Search. *International Journal on Semantic Web and Information Systems*, *12*(1), 62–78. doi:10.4018/IJSWIS.2016010104

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ning, X., Jin, H., Jia, W., & Yuan, P. (2009). Practical and effective IR-style keyword search over semantic web. *Information Processing & Management*, *45*(2), 263–271. doi:10.1016/j.ipm.2008.12.005

Nourine, L., & Raynaud, O. (2002). A fast incremental algorithm for building lattices. *Journal of Experimental & Theoretical Artificial Intelligence*, *14*(2-3), 217–227. doi:10.1080/09528130210164152

Pawlak, Z. (1991). *Rough Sets: Theoretical Aspects of Reasoning about Data*. Dordrecht: Kluwer Academic Publishers. doi:10.1007/978-94-011-3534-4

Petrakis, E. G. M., Varelas, G., Hliaoutakis, A., & Raftopoulou, P. (2006). X-Similarity: Computing semantic similarity between concepts from different ontologies. *Journal of Digital Information Management*, *4*(4), 233–237.

Ponzetto, S. P., & Strube, M. (2007). Knowledge derived from Wikipedia for computing semantic relatedness. *Journal of Artificial Intelligence Research*, *30*, 181–212.

Resnik, P. (1995). Using information content to evaluate semantic similarity in a taxonomy. In *Proceedings of International Joint Conference for Artificial Intelligence (IJCAI 1995)* (pp. 448-453).

Rocha, C., Schwabe, D., & Aragao, M. P. (2004). A hybrid approach for searching in the semantic web. In *Proceedings of the 13th international conference on World Wide Web (WWW 2004)* (pp. 374-383). 10.1145/988672.988723

Rodriguez, M. A., & Egenhofer, M. J. (2003). Determining semantic similarity among entity classes from different ontologies. *IEEE Transactions on Knowledge and Data Engineering*, *15*(2), 442–456. doi:10.1109/TKDE.2003.1185844

Sanchez, D., & Batet, M. (2013). A semantic similarity method based on information content exploiting multiple ontologies. *Expert Systems with Applications*, *40*(4), 1393–1399. doi:10.1016/j.eswa.2012.08.049

Sanchez, D., Batet, M., & Isern, D. (2011). Ontology-based information content computation. *Knowledge-Based Systems*, *24*(2), 297–303. doi:10.1016/j.knosys.2010.10.001

Storey, V. C., Burton-Jones, A., Sugumaran, V., & Purao, S. (2008). CONQUER: A methodology for context-aware query processing on the World Wide Web. *Information Systems Research*, *19*(1), 3–25. doi:10.1287/isre.1070.0140

Suchanek, F. M., Kasneci, G., & Weikum, G. (2008). YAGO: A large ontology from Wikipedia and WordNet. *Journal of Web Semantics*, *6*(3), 203–217. doi:10.1016/j.websem.2008.06.001

Taieb, M. A. H., Aouicha, M. B., & Hamadou, A. B. (2013). Computing semantic relatedness using Wikipedia features. *Knowledge-Based Systems*, *50*, 260–278. doi:10.1016/j.knosys.2013.06.015

Valverde-Albacete, F. J., Gonzalez-Calabozo, J. M., Penas, A., & Pelaez-Moreno, C. (2016). Supporting scientific knowledge discovery with extended, generalized Formal Concept Analysis. *Expert Systems with Applications*, *44*, 198–216. doi:10.1016/j.eswa.2015.09.022

Virgilio, R. D., Guerra, F., & Velegrakis, Y. (2012). *Semantic Search over the Web*. Berlin, Germany: Springer. doi:10.1007/978-3-642-25008-8

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Vocht, L. D., Softic, S., Verborgh, R., Mannens, E., & Ebner, M. (2017). Social semantic search: A case study on Web 2.0 for science. *International Journal on Semantic Web and Information Systems*, *13*(4), 155–180. doi:10.4018/IJSWIS.2017100108

Yao, Y. Y. (1996). Two views of the theory of rough sets in finite universes. *International Journal of Approximate Reasoning*, *15*(4), 291–317. doi:10.1016/S0888-613X(96)00071-0

Yao, Y. Y., & Chen, Y. (2006). Rough set approximations in formal concept analysis. *Transactions on Rough Sets V. Lecture Notes in Computer Science*, *4100*, 285–305. doi:10.1007/11847465_14

Yazdani, M., & Popescu-Belis, A. (2013). Computing text semantic relatedness using the contents and links of a hypertext encyclopedia. *Artificial Intelligence*, *194*, 176–202. doi:10.1016/j.artint.2012.06.004

Zenz, G., Zhou, X., Minack, E., Siberski, W., & Nejdl, W. (2009). From keywords to semantic queries — Incremental query construction on the semantic web. *Journal of Web Semantics*, *7*(3), 166–176. doi:10.1016/j.websem.2009.07.005

Zhang, X., Dai, J., & Yu, Y. (2015). On the union and intersection operations of rough sets based on various approximation spaces. *Information Sciences*, *292*, 214–229. doi:10.1016/j.ins.2014.09.007

This research was previously published in the International Journal on Semantic Web and Information Systems (IJSWIS), 14(3); edited by Miltiadis D. Lytras; pages 99-119; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 70
The Rise of the Virtual Human

Wendy A. Powell
University of Portsmouth, UK

Natalie Corbett
University of Portsmouth, UK

Vaughan Powell
University of Portsmouth, UK

ABSTRACT

Virtual Humans are here to stay. From the voice in your satNav to Apple's "Siri", we are accustomed to engaging in some level of conversation with our technology, and it is rapidly becoming apparent that natural language interfaces have potential in a wide range of applications. Whilst audio-only communication has its place, most natural conversations take place face to face, and believable embodiment of virtual humans is the necessary next step for them to be fully integrated into our lives. Much progress has been made in the creation of relatable characters for film, but real-time facial animation presents a unique set of design challenges. This chapter examines the role of the virtual human, its history, and approaches to design and creation. It looks at ways in which they can be brought to life, interacting, assisting and learning. It concludes with a view into popular culture and perceptions of the future, where fact and fiction meet.

INTRODUCTION

With the rapid advances in technology in the past 40 years, the vision of the virtual human is coming closer to reality, with improvements in visual appearance, speech and emotion, and in the application of this into real-world settings. As can already be observed in pre-rendered animations, near-photorealism is now possible, with some avatars almost indistinguishable (Alexander, Rogers, Lambeth, Chiang & Debevec, 2009; Borshukov, Piponi, Larsen, Lewis & Tempelaar-Lietz, 2005). Although total realism not yet possible, many studies are working toward this goal, with even real-time avatars slowly becoming more and more realistic. This chapter introduces a range of virtual humans currently in development or

DOI: 10.4018/978-1-7998-0951-7.ch070

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

being used in real-world interaction. The history of animated humans and the recent rapid advances are discussed. The chapter will then examine some of the technical challenges involved in creating virtual humans, particularly with regards to facial expression and speech. Finally, the role of the virtual human in popular culture is explored, looking at fears and hopes for the future, and how perceptions are changing over time.

BACKGROUND: THE ROLE OF THE VIRTUAL HUMAN

This section will take a look at avatars already being used in real-life settings, interacting independently with users, learning and remembering information, and even assisting in the treatment of patients.

Avatars Who Interact

Since the early days of computing there has been a fascination with the concept of communication between computers and human beings. As early as 1950 Alan Turing raised the question of whether a computer could appear to think like a human (Turing, 1950), and even today the 'Turing Test' is considered to be a benchmark by which we measure the success of virtual humans. Whilst early work focused on text-based communication due to the poor quality of speech generation and animation, improvements in technology have raised our expectations to the point where we anticipate that a humanised avatar could pass the Turing test within the next generation.

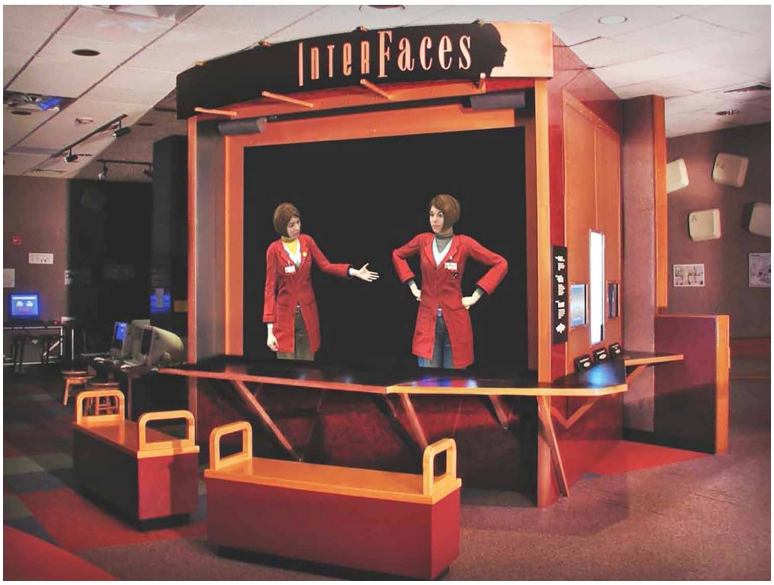
Early research on avatar development and behaviour focused mainly on speech, with the eyes and the motion of the head receiving little attention. However, in a 2004 study, a group developed an avatar to interact with video feeds rather than directly with humans, watching and following movement in a variety of situations. The team used a neurobiological model of visual attention in order to create realistic avatar eye and head animation (Itti, Dhavale & Pighin, 2003). They postulated that this type of non-verbal communication was an essential part of human interaction, and aimed to recreate this digitally in the most realistic way possible, using a procedural, science based approach. A reflexive approach was used to guide the avatars attention toward visual targets, in the form of either video recordings or video game output. Automatically generated eye blinks were added to aid realism, as well as accurate deformation of the facial tissue during movement, and a photorealistic facial model. Although the avatar was not always human-like in its point of focus, results of this early work showed potential, in using attention-directed behaviour to enhance the realism of a virtual human.

More recent work saw the development of virtual twins Ada and Grace (Figure 1), who were deployed in a museum to interact with each other and with visitors using both verbal and non-verbal communication (Swartout et al., 2010).

To communicate with the characters, the operator could push a button then speak into a microphone. The audio was then sent to an automatic speech recogniser (ASR), which translated the speech into text, sending this on to a language understanding module. An algorithm used the text to select an appropriate response for the characters to use. The characters have a set of about 400 responses to draw from, with dialogue management software ensuring that responses are not repeated. Animation sequences were also created and stored, with these being called upon at the same time as the responses, with the relevant animations being called upon dependent on the response. Viewing video footage of Ada and Grace being posed questions by museum visitors (USC ICT, 2010), it can be seen that there is rapid retrieval of

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Virtual twins Ada and Grace act as museum guides
© USC Institute for Creative Technologies. Used with permission.



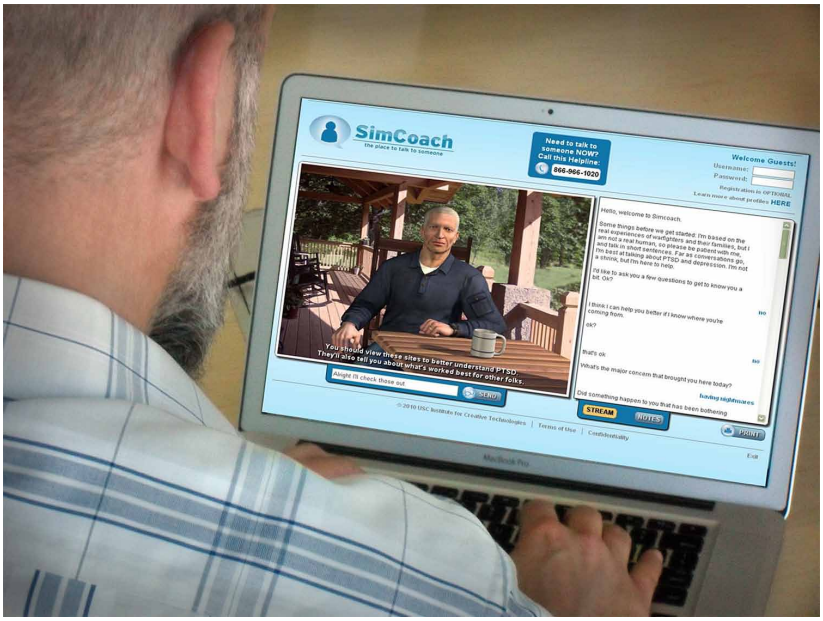
responses with appropriate answers being given to visitors. However, the animation is clearly constrained by the demands of real-time performance, with repetition of a small gesture set and relatively low fidelity facial performance and, partciularly in the synchronisation of mouth movements to speech elements. The lack of movement of the facial musculature when speaking, results in a somewhat robotic effect. As can be seen within the Digital Emily project (Alexander, et.al., 2009), the USC ICT is more than capable of creating incredibly lifelike animations, with their light stage also used to great effect in many films (Debevec, n.d.). However, there does seem to be a trade-off between the achievement of high graphical fidelity, and the ability to work in real-time, with a distinct drop in graphical fidelity when increasing the demand of real-time processes. However, as technology improves and graphical rendering capabilities increase then we are likely to see rapid improvements in this visual fidelity.

Also emerging from USC ICT is the virtual support agent or 'SimCoach' (Figure 2), providing information and support for military personnel and their families (Rizzo et al., 2011). The user will be able to select from a gallery the character that they feel most comfortable talking with. The characters will use speech, gesture and emotion to introduce the system to the user, as well as asking questions about the user's history and medical concerns in order to provide the right kind of support. A prototype SimCoach has been deployed on a veterans website (Brave Heart Veterans, n,d.). At the time of writing, interaction was limited to text-based questioning, and responses are limited and somewhat lacking in visual fidelity. Nevertheless, it is already demonstrating the potential to provide support and advice via a humanised interface.

Researchers in the UK are developing the RITA avatar (Figure 3), which is designed to interact with older adults. Although only in early stages of development at the time of writing, RITA demonstrates a significantly higher level of visual fidelity and realism than many earlier real-time avatars. The prototype uses a library of pre-recorded speech segments and blend-shape animation to communicate with realistic emotional expression. Vision-based facial and emotion recognition software identifies known

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. SimCoach provides information and support for military personnel.
© USC Institute for Creative Technologies. Used with permission.



users and interprets their current emotional state, triggering responses which vary according to both the mood and identity of the user. Technology still needs to advance considerably in order to support the full planned functionality of RITA (RITA, n.d.)., but nevertheless, it gives us a glimpse of the future potential of nearly-real virtual humans.

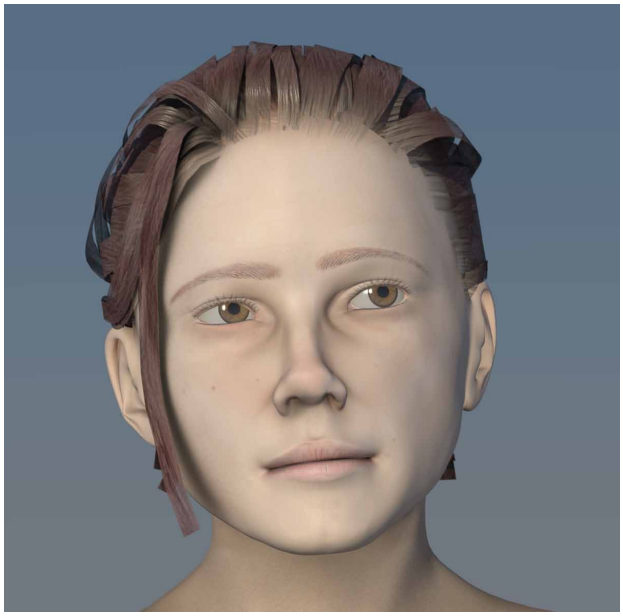
Whilst the ability of a virtual human to hold a conversation is an impressive achievement, it is still of relatively limited application without some further purpose. Conversational agents are deployed in many areas, from marketing websites through to exhibition guides, but development has not stopped here. More recent development is bringing us intelligent agents who teach, learn, and even care for people.

Avatars Who Help

An example of a successful avatar created for use within the field of education, although not a true example of a virtual human, comes in the form of 'Steve' (Soar Training Expert for Virtual Environments), an avatar used for training Navy personnel to operate a high-pressure air compressor (HPAC) on board a ship (Johnson & Rickel, 1997). Steve consists of a virtual hand, inhabiting a virtual representation of a ship, complete with the HPAC that students must learn to operate. The students are allowed a 3D view of this world through use of a head mounted camera, interacting with either a 3D mouse or data gloves. Steve was designed to help students learn to perform tasks such as operating the HPAC, and repairing equipment, but also to enable them to apply these procedures to a variety of situations, by understanding the rationale behind each step. Steve was given the ability to manipulate and point at objects to help students, as well as constantly monitoring the environment. For demonstrating or monitoring tasks, Steve is equipped with a pre-programmed knowledge of tasks in the form of hierarchical plans, learning the processes required to complete each task. Steve can be switched between monitoring and demonstra-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. The RITA avatar, designed to support older adults
© 2014, University of Portsmouth. Used with permission.



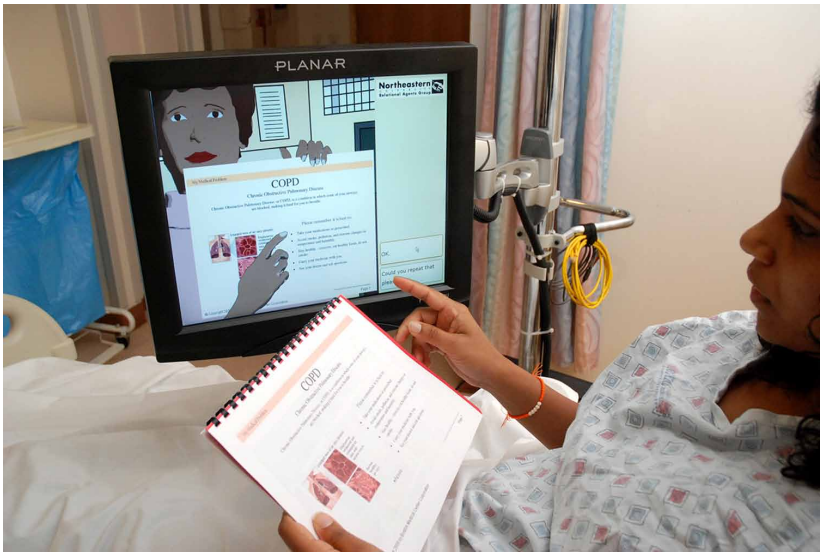
tion modes, allowing for a smooth and flexible student and tutor relationship in which learning can take place at the pace of each individual. Although not a full virtual human, the learning algorithms used in Steve's design have the potential to be used in other avatar applications, particularly as the sophistication of the technology increases.

More recently, a group researchers has been developing virtual nurses to aid in the communication of discharge information to patients with low health literacy skills (Bickmore, Pfeifer & Jack, 2009). Health literacy refers to the ability to understand healthcare information, and to apply this understanding to health situations. The lack of understanding of discharge information can lead to patients being readmitted to hospital after not taking medication correctly, or not understanding side effects. The greatest aid to those with low health literacy seems to be in face-to-face communication, used in conjunction with written instructions, allowing the patient to ask questions, and for the nurse to elaborate on information. The use of this delivery in the form of a virtual nurse is beneficial, as it allows the patient to take as much time as they need, without feeling that they are wasting anyone's time, something which seems to be a recurrent feature in the use of avatars within the healthcare sector (Rizzo et al., 2011). Two virtual nurses were created, with the head and torso displayed for use in the system. These nurses were very stylised, and spoke using a synthetic voice, with scripted dialogue. Animation was also implemented for lip sync, and to allow the hands of the nurse to point at relevant information. Interaction for the patient was in the form of the selection of utterance options, displayed on a touch screen. A mobile kiosk containing the patient information and the virtual nurse was wheeled into the patient's room, with the patient then left alone to conduct their conversation (Figure 4).

The nurse could be programmed with information about the patient, calling them by name, and firstly asking a few questions relating to their interests, in order for the patient to feel more at ease. The virtual nurse would then move on to the discharge information, explaining each point whilst pointing at relevant

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Virtual nurse assisting patient to understand information leaflet
© 2008, Northeastern University, Boston. Used with permission.



parts of the leaflet, and encouraging the patient to follow in their paper copy. The results of the study were largely positive, with the virtual nurse tested on both hospitalised and non-hospitalised participants. 92% of the participants claimed that they felt comfortable with the nurse, with 70% stating that they either preferred speaking with the virtual nurse, or were indifferent. When dealing with doctors and nurses, it may often be the case that patients feel that they are wasting time, particularly in the elderly, who may put off going to their GP or to the hospital because of this. The results of this study support the notion that people like, and often prefer, talking to a virtual avatar, particularly within the healthcare sector.

Avatars within the healthcare industry have been implemented for use in a variety of settings, ranging from communication training for children with autism (Chops, 2014; Kandalaft, Didehbani, Krawczyk, Allen & Chapman, 2013), to treatment of speaking (Slater, Pertaub & Steed, 1999; Kang, Brinkman, van Riemsdijk & Neerincx, 2011), and the treatment of post-traumatic stress disorder (Rizzo et al., 2011).

Autistic children often have difficulty in social communication, and may find it easier to communicate with a virtual human. CHOPS (2014) have created animated characters to aid in treating autistic children. The children can either interact with the animated character, with a speech pathologist used to reply to the children in real-time, through the animated character, or learn to control the animations themselves. The character can be used by therapists in a number of ways, for example, to tell stories, teach lessons, and demonstrate facial expressions or body language. As the avatars used are far less threatening for a child with autism than a real person, the student can remain engaged and will be much more open to learning new concepts. The novelty of the animated characters can also aid in keeping their attention for a longer period of time.

Another example of the use of avatars to aid in the treatment of Autism is through the use of the online virtual-reality world, 'Second-life.' Researchers at the University of Texas Centre for Brain Health, paired clinicians with Autism patients within the Second-life world to deliver therapy (Kandalaft, Didehbani, Krawczyk, Allen & Chapman, 2013). The study aimed to enhance the social skills, social cognition and social functioning among eight young adults diagnosed with high functioning autism. Virtual-reality

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

therapy for Autism patients can be beneficial, as social interactions can be simulated and easily repeated, without fear of mistakes or rejections and with lower levels of stress. Avatars were created to resemble each of the participants within the study as well as the coach involved. After logging into the system for each session, participants would each be instructed to go to a specific location to meet a specific person with whom they would be interacting. This person would be in the form of another clinician, playing the part of a person, such as a potential employer delivering an interview, or a confrontational roommate. The participant would then take part in the selected interaction, with the coach remotely observing, taking notes on the social objectives of the participants performance. After the interaction had taken place, the participant would then meet with the virtual coach to discuss performance, receiving feedback from the coach, which they would then attempt to implement in a second attempt at the same interaction. The results of this study found largely positive feedback among participants, with all stating that they found the sessions enjoyable, and reporting social benefits, such as participating more within social situations and finding it easier to see things from other people's perspectives. However, the very small sample size, as well as the lack of a control group, limits the generalisation of the study. The use of only participants from the high functioning end of the autistic spectrum also means that this study may not be applicable to a very wide range of people within other parts of the autistic spectrum. However, the findings are of interest, and do present the idea that virtual-reality, and virtual avatars can be of benefit within the general healthcare system when used alongside real life medical practitioners, and should not be overlooked within this sector.

Virtual-reality has also proved useful within in the treatment of phobias, providing a safe, non-threatening environment, in which clinicians can have complete control over the stimulus and situation, allowing patients to take therapy at their own pace, and feel more secure than in a real-life phobia inducing situation. One such study aimed to aid in the treatment of public speaking anxiety (Slater et al, 1999). This study aimed to utilise a virtual environment occupied with virtual avatars to explore their effectiveness in psychotherapy for social phobias. To create the environment in which the experiment would take place, researchers constructed a virtual seminar room, matching the one in which participants met with the researchers and completed questionnaires. The virtual seminar room was populated with an audience of eight avatars, placed in a semicircle to face the speaker. The avatars were programmed to display six primary facial expressions, the same as those outlined within Ekman's 'Unmasking the Face' (Ekman & Friesen, 1975), as well as yawns, sleepy faces, and random autonomous actions such as twitches, blinks and nods. Avatars also were programmed with the ability to stand up, clap, and walk out of the seminar room. To increase the realism of the avatars, an unseen operator would trigger the audience responses at key moments during the experiment. Participants were asked to give a prepared talk three times, receiving one positive reaction from the avatars, one negative response, and one mixed. Interestingly, the study found that the level of interest of the audience was important, and this influenced the impact of the positive or negative response, even though the audience consisted of relatively low fidelity avatars. Although this was a fairly early study in the field, it does indicate that virtual humans may be of use in treating social phobia and presentation anxiety.

A more recent study into the therapy of anxiety disorders is the 2011 study by Kang et al. (2011). This study, although not yet completed, aims to use an internet-delivered system to treat multiple patients through the use of exposure therapy, a type of therapy in which the patient is exposed to the cause of their anxiety (c.f. Parsons & Rizzo, 2008). The project is a four year research plan, which once complete, will provide an internet-delivered therapy system, with a virtual therapist implemented who will monitor the patient, making adjustments to the virtual-reality scenario, changing the level of expo-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

sure relative to the patients levels of anxiety. The project in its current stage will only focus on social anxiety, within a public speaking scenario, much like the study of Slater et al (1999). The avatars used within this project, as with the aforementioned study, will take the form of members of an audience, programmed to react to the participant speaking. They will also be designed to portray four different attitudes; interested, neutral, bored, and critical. Each avatar will also vary in its personality and physical state, meaning that each individual avatar should be able to behave autonomously, reacting to both the perception of the environment, and its own mental and physical state. In order to test the effectiveness of the virtual avatars, the project proposes three control methods, the first being the autonomous behaviour already outlined. The second control method will be a scripted method, in which the reference of the real audience filmed, will be replicated exactly for each of the avatars. The final control method will be one of random control, in which avatars will choose a completely random set of behaviours, in no way connected to their environment. Although this project is not yet complete, the audience design is certainly moving towards more realistic virtual humans, with dynamically generated behaviour based upon individual programmed personalities.

Finally, at the cutting edge of development of virtual healthcare agents, are projects creating virtual humans capable of understanding patients and detecting real-time indicators of physical or emotional health.

A collaboration between USC IT and the Defence Advanced Research Projects Agency (DARPA) aims to create a virtual agent capable of identifying indicators of psychological distress, in order to provide healthcare support for military personnel and their families (Figure 5). The system should eventually, be able to monitor users over weeks and months to monitor change, working closely with real life healthcare professionals if psychological disorders are seen to be present (DeVault et al., 2014; Morbini et al., 2014). The systems used within this research consist of two core components. The first is 'Multisense', which is used for the automatic tracking and analysis of facial expression, body posture, acoustic features, linguistic patterns and high level behaviour descriptors such as fidgeting. This system directly communicates with the virtual human, 'SimSensei'. This human platform is able to process all of the signals sent from the Multisense, providing real-time reactions to the user, through speech and gesture.

For the development of this virtual human, whom the researchers named 'Ellie', three design goals were set. The user should be comfortable talking to Ellie, the system should be sensitive to non-verbal behaviour displayed by the user, and should generate appropriate non-verbal behaviour in response. This should all take place within a 15-25 minute interview, in which the user should remain engaged and interested in communicating with Ellie. Data has been extracted from numerous interviews between a real psychologist and patient, and verbal and non-verbal behaviours identified and quantified. Based on this data, the team have begun to create a system that will enable Ellie to operate automatically and autonomously. In order to do this, specific modules were implemented into the mulitsense, which all communicate with one another, in order to dictate how Ellie should behave and react. These modules currently include; 3D head position and facial tracking, expression analysis, gaze direction and audio analysis. The multisense uses all of these signals, measuring factors such as smile intensity, in order to inform the other modules of the SimSensei, which are used to drive Ellie. These modules include the generation of non-verbal behaviour, realization and rendering.

In order to animate Ellie, reference footage of real clinicians was used to create a series of key frame animations for Ellie to draw from. These were placed into a program called Cerebella, which can be programmed to take into account factors such as characters personality, culture, gender and body type to drive animation.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. 'Ellie' is being developed to provide healthcare support to military personnel and their families.
© USC Institute for Creative Technologies. Used with permission.



Ellie currently has the ability to use 100 fixed utterances, with 60 top level interview questions. Elllie's speech is pre-recorded, rather than generated, which does seem to give a more realistic result then synthetic speech generation. In order to detect and process user speech, the system used a variety of different software implemented within the speech recognition module, similar to the processes used within the Ada and Grace study (Swartout et al., 2010), although more complex. These systems were not only able to process human speech, but able to detect the emotion within the speech, and select an appropriate response in real time. Full details of the dialogue processing systems used can be found in Morbini et al., (2014).

Although the majority of participants to date claimed to have felt comfortable with Ellie, most felt that she didn't understand their non-verbal behaviour, preferring speaking to the Ellie who was controlled by the researchers. Given the early stage of development, these results are unsurprising; however do give promise to the future development of this system. From viewing a video of Ellie speaking with participants (USC ICT, 2013), it can be observed that the system appears to display a high level of intelligence. However, the video gives no indication as to whether the system is running autonomously or whether the avatar is being controlled, limiting any conclusions to be drawn as to the effectiveness of the automated system. The animation of the avatar itself is somewhat lacking in fidelity, an issue which still plagues most real-time virtual humans, being a trade-off between realism and system performance. Nevertheless, the avatar is seen to accurately interpret body language and emotion of the two participants, reacting appropriately, in a similar manner to a real life therapist, with reassuring nodding motions, and knowing when to respond once the user had finished speaking. Ellie does demonstrate the potential for future deployment of caring virtual humans, with technological advances only serving to improve the system, and allow for fully autonomous behaviour.

A UK-based early stage project with a long-term vision is developing a virtual human which is designed to provide support, companionship and advocacy to older adults (RITA, n.d). The RITA avatar has been built using state-of-the art motion capture and animation techniques, with a design based on features which appear trustworthy and competent (Powell, Garner, Tonks & Lee, 2014) (Figure 6).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. RITA is designed to appear trustworthy and competent.
© 2014, University of Portsmouth. Used with permission.



The system is designed to store personal, social and medical information, and to monitor health and emotional state in real-time, in order to provide tailored companionship and support. It will link with body-worn and smart-home sensors and also monitor activity and mood over time, identifying early warnings of deteriorating physical or mental health, and accessing appropriate support and intervention.

Although extensive further development is needed to fully realise the vision of projects such as RITA and SimSensei, it is clear that virtual carer-givers are likely to play a part in our future healthcare. However, to create virtual humans which are fully accepted into our daily lives requires pushing the boundaries of technology, in order to achieve realism in both behaviour and appearance. The development of Artificial Intelligence (AI) is being driven by demands in a wide variety of application areas, and the need for the humanised interfaces to AI is increasingly apparent. The next section of this chapter will explore at the technical development of expressive virtual humans, from ancient history to the present day.

THE DAWN OF THE VIRTUAL HUMAN

The Study of Facial Expression

Facial expression is key to our communication with each other, and is an area that has been studied since as early as the 17th century. The earliest know writer on the subject of expression, as well as the use of expression for the deaf, was the medical practitioner John Bulwer, who was the first to record that communication could come by way of movement, rather than just speech, naming this form of gesture, the "Universall language of Humane nature" (Bulwer, 1649). However, it was Charles Darwin who was the first to detail why and how emotional expression is formed, and suggested that expression is universal, expressed the same regardless of culture or background, being a product of evolution and biology (Darwin, 1872). Darwin wrote at length about the basic principles of expression, including the muscles which are activated during expression of each emotion.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In Ekman and Friesen's book 'Unmasking the Face' (1975), it was revealed, through scientific study, that Darwin was in part correct, with six expressions found to be universal. These expressions are; surprise, fear, disgust, anger, happiness and sadness, and many modern studies directly use these six emotions as basis for creating facial animation. Darwin's work, as well as the research of Ekman & Friesan, provides a valuable theoretical underpinning to facial animation, particularly when building a muscle based facial system. The universal rules of which muscles work during each of the emotions, can be applied directly to an animated model. Ekman and Friesen later went on to create a 'Facial Action Coding System', as a way to categorise the physical expression of emotions (Ekman & Friesen, 1978). The Facial Action Coding System, or FACS, is an anatomically based system for the measurement of all visual facial movements. This facial activity is described upon the basis of 44 unique action units (AUs), as well as including categories of head and eye movements and positions. Each action unit has a designated numeric code, with the muscle groups involved in each action linked to each of the AUs. Whilst FACS is anatomically based, a one to one correspondence does not exist between the muscle groups and the AUs, with any given muscle acting in different ways or contracting in different regions to produce visibly different actions. For example, contraction of the medial position of the Frontalis muscle will result in a raise of the inner corners of the eyebrow, whilst a contraction of the lateral portion of this muscle will raise the outer brow. The use of the coding procedures as outlined in FACS also allows for the coding of the intensity of each facial action unit. This is of particular use in animation, whereby varying degrees of emotion will need to be portrayed dependent on the character or situation, and is a system widely used and referenced today.

Although the study of the human face with regards to expression and emotion has ancient beginnings, the application of this understanding to computer based facial modeling, expression and animation is fairly new, first appearing in the early 1970's (Parke, 1972). With the technology to create animations in its relative infancy, early attempts at facial animations were fairly basic by today's standards. However, the techniques and processes used have been adapted and improved, directly influencing the high quality animations that we see today.

Early Examples of Animation

The first method of animating, before the use of computers, was via the use of hand drawings, portrayed in such a way as to give the impression of motion. Perhaps the oldest animation was found on a 5,200 year old bowl in Iran. The bowl had drawings of a goat on it, which when spun, gave the impression of the goat jumping to retrieve leaves from a tree (Ball, 2008). Many centuries later, Georges Melies, who was a professional magician by training, was one of the first great visual effects artists, using tricks during filming to create animation on top of live action from 1896 to 1912. However, it was in 1928, when Disney released 'Steamboat Willie' the first ever hand drawn cartoon animation with synchronised sound, which saw the birth of modern day animation production (The Museum of Modern Art, n.d.).

Computer-based facial animation came somewhat later, with Parke creating one of the first examples in 1972, generating an animated sequence of a head changing expressions, using a 250 polygon basic mesh, and using dots on facial photographs as references for creating the animated expressions (Parke, 1972). By 1988, a more sophisticated approach was seen, with 'Mike the Talking Head' being animated in real-time by puppeteers (Robertson, 1988). 'Mike' was created using data scanned from a real actor to create a polygon mesh, with software algorithms created to interpolate between phonemes of speech.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In the same year that 'Mike the Talking Head' was created, Jim Henson Productions brought out their own real time animation solution in the form of 'Waldo C Graphic' (Walters 1989). Waldo got his name from the device used to control him, a Waldo. This was a type of input device which consisted of an upper and lower jaw, to control the characters mouth, and an arm, which controlled the movement, similar to the technology used to control Mike. Waldo was used by Jim Henson Productions to perform in real time, in a concert with real puppets, appearing in the TV series The Jim Henson Hour until 1992 (IMDb, n.d). The Waldo type of device used to control animation was a rather popular choice in the early days of animation, particularly for use in real time, with examples of this seen through until the 90's (SimGraphics, n.d; Robertson, 1992). This popularity is probably due to the ease of the device allowing an animator to have full control of the character without the need to be physically sat at a computer. The use of a Waldo is rare now, with computers now able to be drive real-time animation without the need for a puppeteer.

The early 1990's saw the introduction of 'Performance Driven Facial Animation' (Williams, 1990). Using a more advanced system of the earlier work of Parke (1972), Williams used motion tracking via the use of physical markers to track the motion of an actor's face making different expressions, using this data to drive animation of a 3D model. Williams was the first to use performance to directly drive animation in this way, and paved the way for motion captured facial animation, with the same marker based motion techniques still in use today. Parallel development saw the introduction of physically-based facial models in the early 1990's. These were based on Ekman's FACS system with an anatomically based muscle system used to control facial movement (Terzopoulos & Waters, 1990). To animate the facial model, 2D static images from a video of a real actor were taken, with an algorithm applied to track the facial features of each image, and apply these in 3D to the model.

During these early years of computerised facial animation, processing constraints restricted the complexity of the 3D models, resulting in characters which appeared blocky and unrealistic. In the early 70's, most standard computers came equipped with just 4Kbs of memory, using 8 bit graphics. Floppy Disks were invented in 1971, with these used for external storage of up to 250Mb of data up until 2000, when the 8Mb memory stick was introduced. In comparison, a standard modern desktop computer will come equipped with 500GB of memory and a graphics card supporting HD video at around 250 bits. These exponential improvements in computer processing power and dedicated graphics processors are able to support very complex digital meshes, enabling the creation of characters which at times are almost indistinguishable from real humans (Alexander, Rogers, Lambeth, Chiang & Debevec, 2009; Borshukov, Piponi, Larsen, Lewis & Tempelaar-Lietz, 2005). In parallel with this development, advances in Artificial Intelligence software techniques have enabled computers to learn, to make semi-autonomous decisions, and to communicate independently to humans, responding to speech and emotion (Rizzo et al., 2011; DeVault et al., 2014). These Virtual Humans are now being used in a variety of settings, from the entertainment industry through to marketing and health-care.

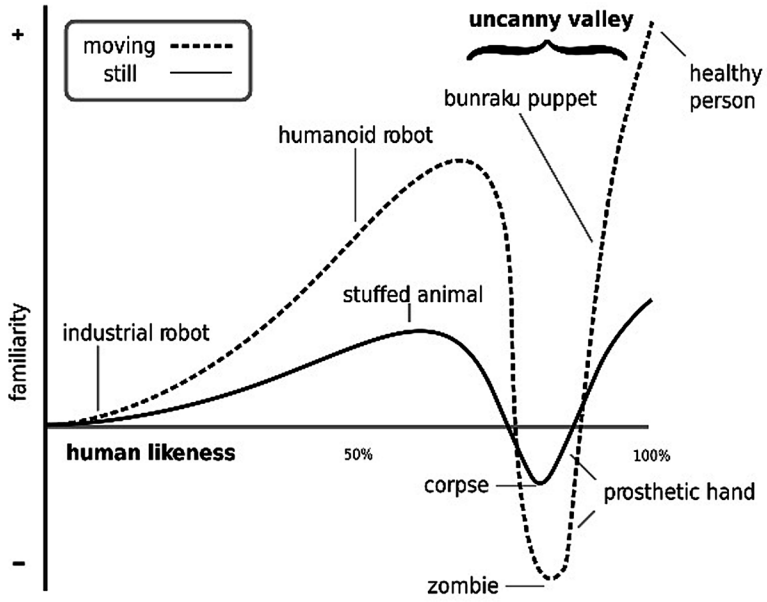
Whilst it is generally considered that the goal of virtual human design is for every-increasing realism, near-realistic avatars can actually be less acceptable than more stylised offerings, a phenomenon known as the 'Uncanny Valley' (Figure 7).

The Uncanny Valley

The 'Uncanny Valley' was first described by Mori (1970). His theory suggest that as a robot (or an avatar) becomes more realistic it is more acceptable, until it approaches a level of near-realism, at which point we begin to notice imperfections, finding these unsettling (MacDorman, Green, Ho & Koch, 2009).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. The Uncanny Valley
Mori, 1970.



This feeling is heightened when movement is added, as indicated by Mori. If a virtual human becomes realistic enough, it emerges from the 'Uncanny Valley' and can be fully accepted. In fact, our perception and acceptance of virtual humans is somewhat more complex than first postulated by Mori, with factors such as facial proportion, texture, detail and believability of movement all playing a role (MacDorman, et. al., 2009; Qiao & Eglin, 2011). The 'Uncanny Valley' graph (Figure 7) proposed by Mori and reinforced by others since, in fields ranging from robotics to CGI, should be considered a conceptual model only. It does not imply any precise values, predict any absolute positions of avatar characteristics, nor give a magnitude of the axes and does not define points on its distribution, nor has it been empirically validated. Nevertheless, conceptually it is generally accepted within its respected disciplines even though such a simplified model is unlikely to really convey the nuances of human emotional reaction to the range of possible integrated characteristics that make up a complex avatar that aspires to be acceptable and lifelike (Tinwell & Grimshaw, 2009). Any deviation or discrepancy in any given characteristic, especially when others are reasonable might be enough to generate an uncanny response. The use of real time motion capture technology for both body movement and capturing real facial expressions has the potential to reduce the discrepancy and discord often perceived in moving characters as reported by Mori. Nevertheless a near veridical detail of subtle movements, skin detail and lighting, as aspired towards by projects such as digital Emily (Alexander et. al. 2009) are as yet far from commonplace and highly demanding. Whilst incongruities remain, the human perception of virtual humans may continue to suffer moments of discord and disengagement, when some element or another does not quite match our expectations or experience and as such, the 'Uncanny Valley' whilst no longer an insurmountable barrier to virtual human acceptance, certainly remains a problematic challenge and a potential pitfall.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Whilst pre-rendered animation produces the highest fidelity results, it is the potential of real-time dynamic animation which is currently of great interest for individualised interactions across a wide range of applications, and this will be the main focus of this chapter. The following sections will address the state-of-the -art techniques for the creation of the virtual human, including building the facial mesh, methods of motion capture and animation, inclusion of speech and emotion, and interactivity.

BUILDING A VIRTUAL HUMAN

There are two main considerations when first creating a virtual human; firstly, how the mesh itself will be created, and secondly, how it should be rigged for animation. Within the creation of the mesh are such considerations as whether to create the mesh manually, or through the use of scans, and the implications of the "Uncanny Valley". For rigging, two main methods can be implemented, bone based rigging, blend shapes, or a combination of the two. Ultimately this initial stage will dictate how the final product will appear, with a better quality mesh, and more accurate rigging producing the highest fidelity of animation, thus this could be considered as the most important stage of development, as it provides the core foundation for the rest of the stages of the production pipeline.

Creating the Mesh

There are two main methods for creating a mesh. The mesh can either be created manually within the 3D space of the computer, or it can be created from scans of a real human face. With both of these techniques, a method of physically based modeling can also be implemented, in which the anatomy of the facial tissue and muscles can be used to influence the creation of the mesh (Figure 8).

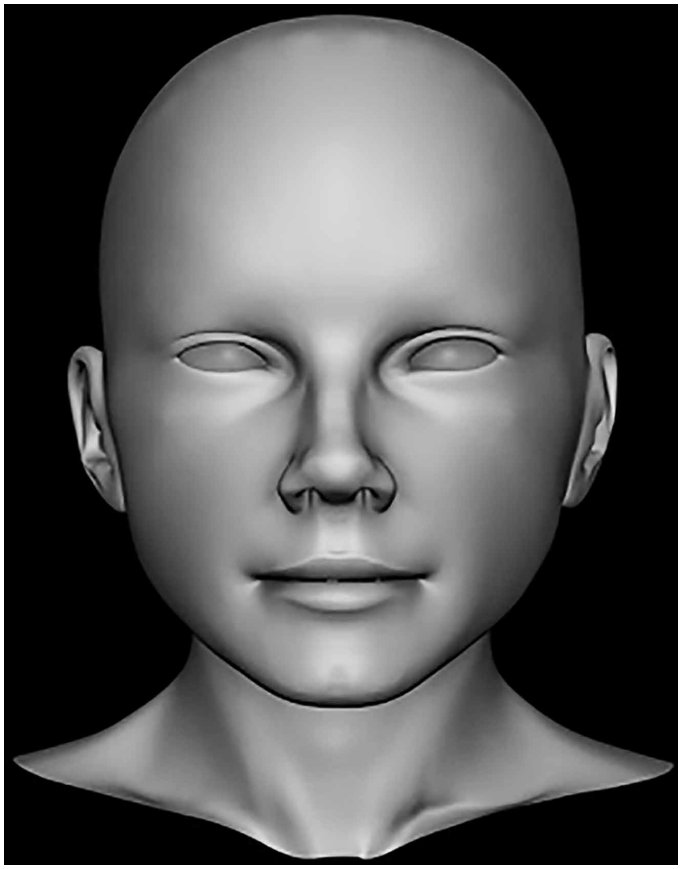
Manual creation of the mesh involves careful manipulation of numerous flat planes to create a 3-dimensional polygonal shape. Whilst time-consuming, it has the advantage of allowing the digital artist complete control of the process. It often results in smaller mesh size, better suited to real-time performance, and allowing complete customisation of the appearance of the face. However, some of the fine details of facial anatomy are difficult to reproduce using manual techniques, which can reduce the visual realism of the resultant avatar.

Alternatively, facial scans can be taken of a real human and the data used to reconstruct the face within the digital space. This method is faster and generally results in a much higher fidelity avatar. However, it tends to result in a much bigger mesh, less suited to real-time performance, and the final appearance of the avatar is constrained by the anatomy of the human used for data capture. In addition, the equipment required for high-definition facial capture may be prohibitively expensive, and produce more than a Gigabyte of data every second.

In order to create a virtual human which is not only realistic in its static form, but also in dynamic expression, the underlying anatomy of the face should be taken into account. Physically based modelling places the polygons of the facial mesh in such a way as to mimic the layout of the underlying facial musculature seen (Terzopoulos & Waters, 1990), allowing for more accurate deformation of the mesh. However, this level of detail can result in a polygon count of millions, making it less suited to real-time animation. Nevertheless, physically based modelling can result in a very high level of visual fidelity (e.g. Sifakis, Selle, Robinson-Mosher & Fedkiw, 2006).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. A base mesh is created in 3D modelling sofware prior to rigging and animation
© 2014, University of Portsmouth. Used with permission.



Rigging the Mesh

A 3D mesh is like a clay sculpture - it can be further moulded, but creating real-time animation would be extremely difficult. Therefore the sculpted mesh must be linked to a digital skeleton (rig), with a system of joints and controls in order to move the face from pose to pose. The two main systems used for rigging a virtual human are bone-based and blend-shape based.

Bones are generally used to control specific parts of the face such as eyelids and jaw. Bones are relatively simple to rig and animate, and one set of bone-based controls can be used to animate a number of different face types. However, facial animation using just bones can appear wooden and emotionless.

Blendshapes are animation tools which allow smooth morphing of one digital mesh into another, giving the appearance of natural movement. These blendshapes can be weighted, with more than one blendshape influencing facial geometry during movement. Although the use of blendshapes adds to both the polygon count and the processing load, it does allow for much more natural facial expression, and can approach a level nearly indistinguishable from a real human. For real-time animation, a combination of bone and blend-shape rigging gives the optimum results, with the blend-shapes being based around Ekman's FACS system (Ekman & Friesen, 1978).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

With the virtual human created and rigged, the next stage in the process is how the avatar is to be animated. The two main options for this stage are animating by hand, or to use video recording, either as reference, or to directly drive the animation through use of motion capture. Within these options are a plethora of different methods, which each yield slightly differing variations of results. These options range from type, and number of cameras to be used, as well as lighting set-up, use of markers, tracking, and application of the data to the virtual human.

Lights, Camera, Action

In order to animate a virtual human, after the initial building phase is complete, many options can be considered. The use of video reference is common practice, either acting as a guide for hand animation, or to directly drive the animation through use of motion capture. The use of video as reference can also be used in a method called 'Rotoscoping', in which an animator will draw over video footage frame by frame to create the animation. This method dates back to the early days of Disney films, with animators directly drawing over video footage for such films as Peter Pan, Alice in Wonderland and Snow White. Rotoscoping is still a method in use today, and can be adapted for use in various situations.

However, hand animating is extremely time-consuming in order to achieve realistic results, and modern motion capture techniques can accelerate this process by an order of magnitude. We suggest that for the level of animation detail necessary for realistic facial expression, motion capture is currently best option, and this is the approach we will address in this section.

The three key considerations one must make when considering facial motion capture are lighting, camera type and placement, and facial tracking techniques. In addition, the data needs to be transferred to the mesh (re-targeting).

Lighting

Whilst not essential for facial motion capture, lighting can support and enhance the capture process, leading to better quality data and a more realistic animation.

Structured light techniques use projected light to cast a structured pattern of light onto the face to provide a detailed surface texture which aids in the tracking of motion (Bradley, Heidrich, Popa & Sheffer, 2010). This can be used instead of placing markers on the face, being less intrusive for the actor and requiring shorter setup times. However, high powered light beams can be a distraction for the actor, and reconstruction done in this way can suffer from a drop in precision, due to interleaving between the ambient illumination and the structured light (Bradley et al., 2010).

Uniform lighting allows a clear view of the face, with no shadows influencing the capture. This reduces 'noise' in the capture data, and results in a better quality of motion data for the animation (e.g. Bradley et al., 2010).

Light stages allow a programmed approach to lighting setup, allowing extremely precise control of the lighting during facial capture. A light stage may have over 150 lights placed around the actor (Alexander, et.al., 2013.).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Camera Set Up

In order to record facial performance, one or more cameras are needed, and the quality, number and lay-out of these cameras can have a significant impact on the quality of the final animation data (Figure 9).

Single camera capture will generally result in 2D data (unless structured light is used). However, due to the low cost and ease of use of single-camera setup, this is the approach of choice for many commercially available systems (e.g. Faceware Tech, Faceshift, Mixamo). This form of motion capture can be effective (e.g. Weise, Bouaziz, Li & Pauly, 2011), although the post-processing require to achieve high quality animation can be complex, and 2-dimensional capture may miss subtle facial movements, leading to a drop in visual fidelity.

Multiple camera setup on the other hand, allows 3-dimensional data to be captured using multiple cameras in different positions. Whilst the quality of the data obtained in this way is much higher, the hardware and software costs may be prohibitive, and there is a much higher processing overhead due to both the quantity and complexity of captured data. Systems can range from just 2 cameras (Cudalbu, Anastasiu, Grecu & Buzuloiu, 2006), to 200 (Havaldar, 2006)

Stereo camera tracking is arguably the most accurate approach to facial tracking (Lewis & Pighin, 2006). This method uses pairs of cameras to track facial movement using triangulation of the data. However, as in other multi-camera approaches, the cost is high and the captured data demanding to store and process.

After lighting and camera setup, the next consideration is the type of tracking used to convert the facial performance into motion capture data in order to apply it to the rigged digital model.

Motion Capture Tracking

There are two main approaches to motion capture tracking. Physical markers may be placed directly onto the performer as reference points, or the key facial features such as eyes mouth and nose are used directly as reference from the video stream.

Figure 9. An actor's facial performance can be captured using an array of cameras.
© 2014, University of Portsmouth. Used with permission.



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Marker-based tracking involves the placing of multiple small markers on key anatomical landmarks of the face. Marker positioning varies, and may use simple visual reference points, or sophisticated reflective marker and camera systems, but the core principle remains the same. The relative positions of the markers are tracked over time, and the resultant data is processed in order to interpret the complex movements of the facial features during speech and emotional expression. Whilst this approach can produce high quality data which is fairly robust (Alexander, et.al., 2013; Sifakis et. al., 2006), it is intrusive and may affect an actor's facial performance. Furthermore, it interferes with the ability to acquire facial texture data from the capture cameras during live performance.

Markerless tracking does not require the placement of any reference points directly onto the actor's face, but relies on the ability of the software to translate the changing facial appearance into 3D motion data. This can be done by tracking key features of the face, such as eyes, nose and mouth (Terzopoulos & Waters, 1990; Lewis & Pighin, 2006; Bradley et.al., 2010). Although feature-based tracking is often considered to be a somewhat outdated approach, it's low cost and ease of use make it the preferred method within a number of commercial systems, particularly where the main goal is to drive bone-based animation. A more popular alterative in recent years is appearance-based tracking, which uses a training set of data based on variations of the face shape and skin texture during reference expressions, before capturing a full facial performance (Bradley et. al., 2010; Borshukov et. al., 2005).

Data Retargeting

Unless the model is an exact representation of the human facial actor, the captured data will need to undergo further adaptation before it can be applied to animate a rigged model. This processing is known as retargeting. The most common approach is to use blendshape weightings, with an estimation of the parameters for the target blendshape derived from the capture facial data at each frame (Figure 10). Alternatively, non-linear mapping produces a target expression for each source expression extracted from the captured data in order to replicate distinct facial expressions.

The final stage in the creation of a virtual human is in the implementation of speech and emotion. This is of particular importance when creating a real-time conversational avatar.

Figure 10. The retargeted data can be applied to the model to create a range of facial expressions.
© 2014, University of Portsmouth. Used with permission.



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

BRINGING A VIRTUAL HUMAN TO LIFE

If we are to create a believable avatar to remind us to take our medicine, teach our children, or chat to an elderly relative when they are lonely, we need them not only to have a realistic face, but also to be able to communicate, both verbally and non-verbally. Within pre-rendered animation, the addition of speech and facial expression is a relatively straightforward task, with facial capture methods used as the foundation for replication of speech and emotion from the performer to the 3D model. Within real-time animation, however, the method of the generation of speech and emotion is a challenging one, particularly if both are to be portrayed simultaneously. Within speech, challenges such as coarticulation and synchronisation of lip movements to audio (lip sync) must be addressed, as well as the replication of the shapes the mouth makes when annunciating, avoiding looking robotic and unrealistic. The consideration of synthetic or natural speech must also be accounted for, with various limitations in the use and creation of synthesised speech. Emotion is a slightly easier behaviour to replicate, with Ekman's FACS system a useful guide for replicating this both in real-time and pre-rendered examples, however it does present challenges when paired with speech. This section will focus on the speech requirements for real-time interaction.

Natural Speech

In order to support realistic interactions, the goal of most virtual human projects is to produce speech elements which are as natural as possible. The most straightforward way to achieve this is to pre-record all the required speech segments with their animations (Swartout et al., 2010 DeVault et al., 2014 Morbini et al., 2014 Rizzo et al., 2011). However, this does constrain the range of possible communication, and is better suited to applications with a limited set of required responses

Phoneme Based Speech

A phoneme is the smallest unit of sound within speech, for example the word 'cat' is made up of three phoneme sounds – 'c' 'a' and 't'. A grapheme is a letter or number of letters that make up the sounds in our speech, for example the word 'ship' is made up of two, two letter graphemes – 'sh' and 'ip'. A viseme is the shape the mouth makes when articulating a phoneme or grapheme (Ezzat & Poggio, 2000). However, the term 'phoneme' is often used to encompass all of these terms. A library of phonetic sounds and animations can be created for the animated character to draw on. This is particularly useful in real-time conversation or situations where the virtual human will need to have a wide variety of words and phrases. Manual animation or motion capture of a large speech library is costly and time-consuming, and by creating this core library, any speech content can be dynamically generated. The fidelity of this type of speech generation varies greatly, depending on the methods used to combine the discrete elements together. Every phonetic segment is influenced by the preceding and following segments (co-articulation), and if this is not adequately addressed then the generated speech will sound unrealistic and robotic. Phonemes can either be pre-recorded or synthetically generated.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Synthetic Speech

Unless speech can be fully pre-recorded, synthetic speech is often the only option, stringing together sets of phonemes to automatically simulate speech (Wan et al., 2013; Arellano, Spielmann & Helzle, 2013). Although natural speech is generally preferred by users over synthetic speech, some factors have been identified which improve the perception and understanding of synthetically generated speech. The first factor is the acoustic-phonetic properties of the speech. The clearer and more realistic the speech sounds, the better understanding will be. As quality of synthetic speech degrades, the listener will rely increasingly upon their own linguistic knowledge to identify the words spoken. This leads on to the second factor, being the previous experience and background of the listener. A listener with previous exposure to the synthetic voice will perform much better in comprehension tests than one that is hearing it for the first time. A person with higher linguistic knowledge is also likely to perform better. As well as this, voices that are familiar to the listener are better understood, with native English speakers performing better than non-native English speakers when listening to an English voice (Pisoni, 1997). Although synthetic speech generation allows far more conversational flexibility, the techniques are not yet able to generate fully realistic voices or emotional overlays, resulting in speech which is, at best, somewhat flat and robotic. Nevertheless, progress in this area is continuing, and in the future synthetic speech is likely to be the best option for interactive virtual humans.

Emotive Speech

In order to fully engage with a virtual human, it is necessary to achieve some level of emotional realism and empathy. However, applying speech and emotion simultaneously is a challenging task for application in real time, often not quite achieving realism (Arellano, Spielmann & Helzle, 2013). Effective emotional communication requires the selection (identification) of the appropriate emotion, accurate expression of the emotion, and blending of the expression onto speech elements with the correct weightings. At the most basic level, this can be achieved by predetermined emotions associated with set phrases which are stored in a pre-animated library. However, real-time integration of emotional expression with dynamically generated speech segments requires sophisticated algorithms which can assign weightings to different elements of the blendshapes according to both the linguistic and emotive demands of the speech elements. Whilst development in this area is still in early stages, the Cambridge University 'Zoe' project has had some success using Hidden Markov Model (HMM) based Cluster Adaptive Training (CAT) in order to allow for synthesis from the text written, to the speech, and to allow Zoe to produce new emotions. The system learns which phrases or emotions to use (Wan et al., 2013). However, in spite of good lip synching and the advanced speech algorithms, she still appears relatively lifeless and robotic in video footage (University of Cambridge, 2013). This could be, in part, to do with the fact that the animation has been applied to a fairly basic 3D model, rather than a high quality mesh with many polygons. It could also be due to the fact that, as with the other studies that have used synthetic voice (Ezzat & Poggio, 2000) (Arellano, Spielmann & Helzle, 2013), the voice does not sound human, somewhat detracting from the animation. The emotion does come through in the voice, which is an impressive feat, not seen before during real time synthetic voice projection, however the face still needs some work to also express this emotion, with the mouth animation doing most of the work, the eyes and eyebrows largely remaining in a neutral pose throughout the animations.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The RITA project (RITA, n.d) has taken a slightly different approach for its avatar prototype. An actress was used to record a library of speech segments and facial animations, including a range of emotional expressions, which were retargeted onto the facial mesh and deployed in real-time (Figure 8). Words and phrases from the library are selected and combined by the intelligent software according to the identified user and their facial and vocal input. Whilst this approach allows for a much more realistic encounter for short interactions, there are limits to the scalability of such a technique, and as improvements in natural speech synthesis continue, it would be more useful use this technique in the future. Nevertheless, the concept demonstration of the RITA system is impressive, and shows the level of real time animation performance which can already be achieved (RITA, n.d)..

Although there is clearly still much work to be done to achieve truly realistic virtual humans, they already have a significant presence in our digital encounters, and this is likely to increase as technology improves and the underlying intelligent systems develop.

WHERE NEXT?

This final section examines the rise of the virtual human in fact and fiction, and reflects on the narrowing gap between the two.

Media representations and perceptions of virtual humans, avatars and artificial intelligence (AI) take many forms and are typically used either to explore philosophical debates as to what it is to be human, or to express our concerns as to what form these might take. Can they be controlled, or will they attempt to control us? Are they to be welcomed, or feared?

In the entertainment industry there has been a notable shift from the depiction of detached cold automata in the original "Battlestar Galactica" (1978) and "Terminator" (Colla & Larson, 1978, Cameron & Hurd, 1984), to portraying AI and virtual humans as main characters in their own right with emphasis either on their seemingly ever present assistance and acceptance in "Ironman", "I Robot" and "Her"(Favreau, Theroux & Lee, 2002, Proyas & Goldsman, 2004, Jonze, 2013) and with more depth of character, demanding empathy from the audience: Automata, Ex Machina, I Robot, A.I., Battlestar Galactica (2004) (Ibanez & Legaretta, 2014, Garland, 2015, Proyas & Goldsman, 2004, Spielberg & Aldiss, 2001, Larson & Moore, 2004). Nevertheless the cautionary tale of too much power and autonomy leading to potentially omnipresent and almost omnipotent technology which judges humans to be inferior or a threat to be eradicated or controlled remains a persistent and popular narrative: Terminator: Genisys, I Robot, Battlestar Galactica, (Taylor et.al., 2015, Proyas & Goldsman, 2004, Larson & Moore, 2004).

Such ambivalence around both acceptance and fear of the virtual human in fiction is reflected in the perception of its adoption in mainstream society. Individual reactions to the concept of virtual human assistance can equally be met with enthusiasm or open hostility (Figure 11).

Films such as "I Robot" (Proyas & Goldsman, 2004) pitch these two perspectives against one another utilising a unique sentient machine with human emotion, creativity and compassion which realises the intent of Asimovs rules (Saygin et al 2000) is to protect humanity in its holistic form including physical and emotional state and its desire for life. The antagonist is portrayed as an almost omnipresent and omniscient purely logical system which concludes that the only way to comply with Asimovs rules is to effectively incarcerate mankind to protect it from itself.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 11. Typical comments on the RITA concept, showing the ambiguity felt towards virtual humans
©, 2014 University of Portsmouth. Used with permission.

> Wow. What a depressing idea – makes me want to cry! Maybe the 'strain of taking care of Britain's aging population' could be shouldered by human beings? with kindness, and beating hearts?

> Having had three elderly relatives die in the past five years, but being painfully conscious that however hard I tried the majority of their time was spent alone (one in Belfast, one in Norfolk, one in South England) I would not have seen RITA as a substitute for me, other family members and for the professional carers each had. Rather RITA would have augmented what we all did to make at least two of those elderly men a little less lonely and bewildered in their final years. The reality is that in a society that has grown away from living in extended family units, with increased younger generation mobility, and with an ageing population, I think we do need to look to technology for some of the answers. For myself, I'd pre-order a RITA for my own elder years !

The film 'Her' (Jonze, 2013), unlike the others discussed, uses a mobile phone as the main vehicle for the Artificial Intelligence, rather than a humanoid robot. This film is almost a love story between human and AI, with the two forming a relationship. The AI portrayed in this film has the ability to 'learn', something which is already being demonstrated in real life AI development, albeit not yet in such a human-like way (Baker, 2012). Nevertheless the ability of programs such as IBMs Watson (Baker, 2012) to draw upon information from an unstructured corpus of data, including social network feeds and online chats to create a reasonably credible persona tailored to the preferences of an individual is bringing such scenarios closer to reality. In addition the possibility that many online virtual interactions may utilise avatars to convey a personality that may bear no relation to the user, real or virtual (Cooper, 2007), has many implications for who or what we perceive as our social circle.

The film "Ex Machina" (Garland, 2015) poses a slightly different dilemma and looks at the creation of AI so sophisticated that it encompasses a simulacrum of emotion in a humanoid form. Whether sentient self awareness actually occurs or whether it is simply following its instruction to convince someone to help it escape remains deliberately ambiguous. Nevertheless the film develops a plot where its capacity to appear human not only exceeds the Turing test (Saygin et. al., 2000) where an individual can no longer discern it from a sentient person but its ability to emotively manipulate someone to empathise and collaborate with its wishes leads an individual to attach their loyalty to the AI as opposed to the AIs creator. In this role the virtual human 'AVA' even draws upon the protagonists own social media and internet searches to identify appearance and clothing styles that are likely to be attractive to him, a capability that is no longer entirely science fiction. The film also briefly explores a crisis of identity where the individual discovers that other individuals in the house are not human as first perceived but actually AI personas which leads the individual to dramatically question their own humanity as well as their sanity.

Artificial Intelligence (Spielberg, 2001) was based on the 1969 book 'Super-Toys Last all Summer Long' by Brian Aldiss. This film is based around a futuristic society in which Robots are designed to think and feel as humans do, centred on a robot designed to mimic a young boy who is placed to live with a human family as surrogate for the child they lost. The child robot of course needs to be designed

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

to believe in his role that he portrays, yet does not grow up and is even abandoned by those around him. The conceptual dilemma of 'When is a machine no longer just a machine that can be discarded?' is becoming a more common theme explored in popular media.

There is a sense here that virtual humans might be acceptable in the short term as companions but the prolonged lack of real human contact might further the sense of isolation and such considerations would suggest that such systems should be considered augmentation of and not replacements for real human company. In addition the concept of replacing a lost companion whilst in the short term might have a certain appeal for some would require an incredible suspension of disbelief and denial that might not easily be maintained and perhaps just postpone grieving and possibly exacerbate guilt at the acceptance of the virtual substitute.

Virtual humans for healthcare have been and continue to be explored in both fiction and reality. The Doctor (Emergency Medical Hologram) in Star Trek (Frakes, et.al., 1996) is represented as a virtual human as a means to be relatable when treating patients. Star Trek has previously experimented with virtual humans such as "Data" (Frakes, et.al., 1996) as a means of exploring the concept of human and both of these characters undertake this narrative element seeking to become "more human". The idea of virtual humans being responsible for our healthcare is potentially contentious in some aspects but with the current healthcare burden technological assistance to support with common tasks is now arguably a necessity and restoring a virtual human interface over seemingly impersonal touchscreens would appear to be well received for a number of duties freeing up qualified human staff for more vital roles. Virtual healthcare practitioners, such as SimSensei and RITA, arguably may have been inspired by the early iterations of such sci-fi concepts and certainly seem to aspire to that level of interaction and knowledge base.

Sci-Fi Games have also naturally adopted the concept of the virtual human as an A.I. interface or as an assistant or narrative director. 'Cortana' in the Halo series of games (Staten & Boren, 2001) is depicted as a female smart A.I. virtual assistant providing narrative and context to missions as well as tactical updates and guidance. Largely there is little in-game mechanics for her in the Halo series of games although she is depicted accessing computers etc in cut scenes. The role of an avatar as a digital assistant is already available in various guises such as "Denise" and her various alter egos which are the creation of a Brazillian company, Curitiba-based Guile 3D Studio. Their system uses artificial intelligence software, and voice recognition and synthesis in 46 languages as well as biometric systems. A more direct, albeit contrived, merging of science fiction with factual reality is to be represented in the debut of an integrated personal digital assistant "Cortana" for the Windows Phone 8.1 (Warren, 2014). Cortana is designed to continuously learn a user's behaviours, interests and preferences as well as organise their daily activities and schedules. The name is not a coincidence; one novelty of this virtual human assistant is the very deliberate use of the Cortana avatar from the Halo games, including using the same voice actress (Jen Taylor) who plays Cortana in the game. In this case the role of fiction inspiring reality is clear.

Virtual human personas as a means of technological assistance or enhancement are perhaps more readily acceptable for some. Concepts such as the seemingly omniscient Jarvis AI in Ironman (Favreau et. al., 2010) seem to operate like sophisticated personal assistants and information repositories for much of the time but are clearly capable of operating the systems in the home and even autonomously operating the "Ironman suit". Naturally enough it is this autonomous operation of weaponised technology that is often a source of concern both in films and the real world. Unmanned autonomous battlefield assistance vehicles and even weaponised drones are potentially helpful for the protection of troops on the ground and have even demonstrated rescue and evacuation capability for injured or vulnerable personnel (Osborne, 2009). Whilst many of them have AI functionality and are capable of making autonomous decisions,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

including identifying and firing upon threats, there has been a notable reluctance to portray them with a virtual human persona to the general public. Machines being able to carry out tasks, without risk to their human counterparts, are clearly an attractive proposition to the military. Yet the idea of machines being able to autonomously fire upon human beings is largely unpopular in the media and dysfunctional systems or unanticipated programming conclusions are the inspiration for many films where machines turn upon their makers, such as 'Terminator', 'Battlestar Galactica' or 'The Machine' (Cameron, 1984, Pate, 2012, James, 2013)

In the past the idea of the machine being a slave or a tool that toiled or made life easier for its creator seemed part of the utopian vision, and naturally a narrative that allowed that machine to throw off its shackles of slavery and rebel against its masters seemed an obvious counterpoint dystopian vision. The subtle shifts in the perception of A.I., and the philosophical dilemmas they raise, have complicated this vision. Popular media is exploring narratives where machines can appear so human that we cannot discern whether they are or not, and even machines that believe they are human or at least similar in free thought and complex emotion, including love and loss. The reality is perhaps not far behind, but for now, other, perhaps less anticipated interactions between man and machine, are emerging that are far from our utopian ideal.

Cyber security is now a pervasive issue with a number of artificial intelligence systems and programs designed to overcome the conventional security systems we often take for granted. However we rationally determine that there are some subtle differences in the perception capacity of these "bots" and those of a "real" human that we can utilise to our advantage. For example, the "Captcha" security system is a visual arrangement of letters and numbers that can be interpreted easily by humans but are a barrier to intelligent "bots". In a somewhat ironic twist some A.I. systems now identify these captcha and copy them in order to present them to real human operators who may believe they are entering in the correct details to legitimately access information on some other site (Egele et. al., 2010). In this manner A.I. is now effectively utilising humans to achieve its designed goals, and so the puppet guides the puppeteer.

However, the concept of A.I. determining our actions is not new, and is still being explored for its considerable benefits for relieving the cognitive load in our everyday lives, and in scenarios with extreme tedium and risk of poor error management from the human controller. This has been evident in automatic pilots, and the choice of the female human voice to calmly deliver flight control warnings to pilots and in the investment into autonomous self driving cars and transport systems with voice and gesture commands and audio responses. It is likely that in the near future such systems may use a human-like persona for mediating interactions.

We have a long way to go to fully comprehend both the extent of the impact and the moral and ethical dilemmas of the virtual human and its perception and acceptance in society. Fears, both rational and unfounded, will have to be addressed. Nevertheless technological assistance has been welcomed into many aspects of our lives, becoming commonplace to the extent where we wonder how we would mange without them. Moreover the future potential of such technology is hardly something to be dismissed lightly. Perhaps in time the virtual human that tends to your health and knows your every need and preference so well will become equally acceptable, as will the virtual PA who has all the information you need and can tend to those tedious reports. Maybe even the virtual companion who can share your memories and reminiscences with anecdotes and stories will be like one of the family. With ageing populations, greater demands and expectations of healthcare, more complex technology and the need to manage vast amounts of data, something needs to be done to relieve the burden. The virtual human in differing guises seems

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

well suited to the task and is realistically already among us. The fear of "what if it breaks" will perhaps always remain in our psyche, but humans too are fallible particularly under stress. If "to err is human" then perhaps "to err less often is virtual human" and perhaps that will suffice for now.

We would clearly be foolish to abandon all the potential benefits inherent in the virtual human but we must strive to focus our design towards the safe utopian vision of our digital future.

Apart from the actual representation of virtual humans, perhaps the key concerns which emerge as repeating themes in fiction and in fact, are centered around issues of AI autonomy and ethical decision making, and social isolation from human contact. In order to embrace the full potential of the virtual human, these issues must be addressed.

Even before the recent advances in virtual human technology, our increasingly technological society is creating concerns around social isolation and lack of human-human interactions. Some might argue that for those already naturally socially isolated (the elderly, disabled, social phobias etc), technology-enhanced interactions actually decrease social isolation by providing a medium and means to interact through the guise of an online alter ego. In Science fiction this has naturally been the subject of film plots such as "Surrogates" (Mostow, 2009). In the 'real' gaming world this is a natural expression of 'self', with users portraying themselves in a variety of customizable avatars that may or may not relate to their real persona. This can be empowering (allowing a physically disabled individual to portray himself in an able bodied avatar) or sometimes allow the freedom for individuals to portray themselves as they would like to be perceived by others, or to explore alter egos whose appearance and behaviour diverge from their normal life (Cooper, 2007). Alternatively however, it might also allow those with ulterior motives to approach unsuspecting individuals in a guise that intentionally deceives age, gender and intent with far more sinister implications.

Positive or negative, we cannot ignore the reality of increasing levels of human-computer interaction in all its forms (games, social media etc) and its potential impact on health, obesity, and social withdrawal. The research in this area is of prime importance, not just in documenting the concerns, but also in investigating solutions which don't exclude technology, but facilitate more positive interactions.

The concerns around AI having too much autonomy are extensively investigated in some aspects such as military and legal issues, yet largely neglected in other more social forms to date, yet these are often fears expressed by the general population as the technology closes the gap between reality and fiction. Although we don't yet have any clear failsafe mechanisms to address these theoretical future risks, for now perhaps the approach taken by the Watson AI (Baker, 2012), whereby the AI will offer an informed opinion, but also presents the evidence supporting the opinion, leaving the (human) user to make the final decision, offers one acceptable compromise.

This strategy of keeping a human "In the loop" is one commonly suggested for key decision making and is often advocated in military options, (and indeed the fears of its absence makes a key plotline in films such as "Terminator"), but maintaining fully informed human decision making as a safety override imposes its own limitations in terms of timely decisions and bias judgments and allocating effective man-hours which may negate or undermine some of the key benefits of virtual humans.

Ethical decision making is complex among 'real' humans relying on judgments and weightings that are dynamic and even personal. Translating this into a rule based AI system with theoretically universal weightings and priorities will be contentious at best. Asimov's rules, although a good conceptual starting point, are simply not enough here as their interpretation and implementation are as yet not fully defined and prone to conflicts whose consequences we as humans would have to live with.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Even the concept of realistic virtual humans has ethical conflicts. Blurring our ability to recognize human from virtual human has some notable advantages, for example as companions, care providers, advocates and assistants among others. However, there is a fine balance to maintain.

It's vital that we retain our humanity, which includes caring for the vulnerable in our society. If we leave this task entirely to an army of robots, we will in the process lose something of what makes us fully human. (Powell W, 2014)

At what point does convenient, comfortable deception become negligent, malicious or abusive? Should 'Virtual Humans' always be marked out as different somehow so that when the beguiling suspension of disbelief strays into misrepresentation we have the means for a reality check? Such issues have a form of precedent in the media with images of models that have been 'enhanced' with image manipulation software, often now carrying a disclaimer to that effect. Nevertheless the message conveyed by a static 2D image is unlikely to be as complex to mitigate, as that of a fully interactive avatar.

With our continuing changes in society it is increasingly likely that we may come to rely on these 'Virtual humans' for our healthcare and many other everyday interactions and services, but therein lies a dilemma: we need them to be like us for our sake, possibly better than us for societies sake, but we remain unsure how much like us is permissible and ultimately controllable. Wherever the future of virtual humans takes us, we should not be reactive after the event, but anticipate and plan to firmly define their role and place in society in order to fully realise their benefits without the realisation of our fears.

REFERENCES

Alexander, O., Fyffe, G., Busch, J., Yu, X., Ichikari, R., Jones, A., … Antionazzi, B. (2013) Digital Ira: Creating a real-time photoreal digital actor. In *ACM SIGGRAPH 2013 Posters 1*. ACM.

Alexander, O., Rogers, M., Lambeth, W., Chiang, M., & Debevec, P. (2009). *Creating a Media Production*. CVMP.

Arellano, D., Spielmann, S., & Helzle, V. (2013). The Muses of Poetry - In search of the poetic experience. In *Symposium on Artificial Intelligence and Poetry* (pp. 6–10). Academic Press.

Baker, S. (2012). *Final Jeopardy: The Story of Watson, the Computer That Will Transform Our World*. New York: Mariner Books.

Ball, R. (2008). Oldest Animation Discovered in Iran. *Animation Magazine.* Retrieved September 26, 2014, from http://www.animationmagazine.net/features/oldest-animation-discovered-in-iran/

Bickmore, T. W., Pfeifer, L. M., & Jack, B. W. (2009). Taking the time to care: empowering low health literacy hospital patients with virtual nurse agents. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1265-1274). ACM. 10.1145/1518701.1518891

Blomkamp, M., & Tatchell, T. (2015). *Chappie* [Film]. Columbia Pictures, Media Rights Capital, USA.

Borshukov, G., Piponi, D., Larsen, O., Lewis, J., & Tempelaar-Lietz, C. (2005). *Universal capture-image-based facial animation for The Matrix Reloaded*. Academic Press.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Bradley, D., Heidrich, W., Popa, T., & Sheffer, A. (2010). High resolution passive facial performance capture. *ACM Transactions on Graphics*, *29*(4), 41. doi:10.1145/1778765.1778778

Brave Heart Veterans. (n.d.). *BraveHeart: Welcome Back Veterans Southeast Initiative*. Retrieved August 28, 2014, from http://braveheartveterans.org/

Bulwer, J. (1649). *Pathomyotomia: Or a dissection of the significative muscles of the affections of the mind*. London: Humphrey Moseley.

Cameron, J., & Hurd, G. (1984) *Terminator* [Film]. Hemdale Film, Pacific Western, USA.

Chops. (2014). *Cartoon Characters Come Alive For Children with Autism*. Retrieved August 22, 2014, from http://www.chops.com/blog/cartoon-characters-come-alive-for-children-with-autism/

Colla, R., & Larson, G. (1978). *Battlestar Galactica* [Film]. Glen A. Larson Productions, Universal Television, USA.

Cooper, R. (2007). *Alter Ego: Avatars and their Creators*. London: Chris Boot Ltd.

Cudalbu, C., Anastasiu, B., Grecu, H., & Buzuloiu, V. (2006). Using stereo vision for real-time head-pose and gaze estimation. *University "Politehnica" of Bucharest Scientific Bulletin, Series C*: *Electrical Engineering*, *68*(2), 15–26.

Darwin, C. (1872). *The expression of the emotions in man and animals*. London: John Murray. doi:10.1037/10001-000

Debevec, P. (2012). The light stages and their applications to photoreal digital actors. In SIGGRAPH Asia Technical Briefs. ACM.

DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., . . . Morency, L. (2014) SimSensei kiosk: a virtual human interviewer for healthcare decision support. *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. Academic Press.

Ekman, P., & Friesen, E. (1978). *Facial action coding system: a technique for the measurement of facial movement*. Palo Alto, CA: Academic Press.

Ekman, P., & Friesen, W. (1975). *Unmasking the face* (1st ed.). Englewood Cliffs, NJ: Prentice-Hall.

Ezzat, T., & Poggio, T. (2000). Visual speech synthesis by morphing visemes. *International Journal of Computer Vision*, *38*(1), 45–57. doi:10.1023/A:1008166717597

Faceshift. (2012). *Faceshift face animation software: we put marker-less motion capture at every desk*. Retrieved August 18, 2014, from http://www.faceshift.com/

Faceware Tech. (2014). *Faceware Tech, Facial Motion Capture & Animation*. Retrieved August 11, 2014, from http://facewaretech.com/

Favreau, J., Theroux, J., & Lee, S. (2010). *Iron Man 2* [Film]. Paramount Pictures, Marvel Entertainment, Marvel Studios, USA.

Frakes, J., Roddenberry, G., Berman, R., Braga, B., & Moore, R. D. (1996). *Star Trek: First Contact* [Film] Paramount Pictures, USA.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Garland, A. (2015). *Ex Machina* [Film]. Pinewood Studios, UK.

Havaldar, P. (2006). Performance Driven Facial Animation. *Proceedings of SIGGRAPH*, *2006*, 23–42.

Ibanez, G., & Legaretta, I. (2014). *Automata* [Film]. Green Moon, Nu Boyana Viburno, New Boyana Film Studios, Sofia, Bulgaria.

IMDb. (n.d.). *Waldo C. Graphic (Character)*. Retrieved September 5, 2014, from http://www.imdb.com/character/ch0116570/

Itti, L., Dhavale, N., & Pighin, F. (2003). Realistic avatar eye and head animation using a neurobiological model of visual attention. *Proceedings of the Society for Photo-Instrumentation Engineers*, *5200*, 64–78. doi:10.1117/12.512618

James, C. (2013). *The Machine* [Film]. Red & Black Films, UK.

Johnson, W. L., & Rickel, J. (1997). Steve: An animated pedagogical agent for procedural training in virtual environments. *ACM SIGART Bulletin*, *8*(1-4), 16–21. doi:10.1145/272874.272877

Jonze, S. (2013). *Her* [Film]. Annapurna Pictures, USA.

Kandalaft, M., Didehbani, N., Krawczyk, D., Allen, T., & Chapman, S. (2013). Virtual reality social cognition training for young adults with high-functioning autism. *Journal of Autism and Developmental Disorders*, *43*(1), 34–44. doi:10.100710803-012-1544-6 PMID:22570145

Kang, N., Brinkman, W., van Riemsdijk, M., & Neerincx, M. (2011). Internet-delivered multipatient virtual reality exposure therapy system for the treatment of anxiety disorders. Canadian Human-Computer Communications Society.

Larson, G., & Moore, R. (2004). Battlestar Galactica [Television Programme]. British Sky Broadcasting (BSkyB), David Eick Productions, NBC Universal Television, USA.

Lewis, J. P., & Pighin, F. (2006). Performance-driven Facial Animation Introduction. *Proceedings of SIGGRAPH*, *2006*, 5–9.

MacDorman, K., Green, R., Ho, C., & Koch, C. (2009). Too real for comfort? Uncanny responses to computer generated faces. *Computers in Human Behavior*, *25*(3), 695–710. doi:10.1016/j.chb.2008.12.026 PMID:25506126

Mixamo. (2014). *Mixamo: Production-quality 3d character animation in seconds*. Retrieved 18 August 2014, from https://www.mixamo.com/faceplus

Morbini, F., DeVault, D., Georgila, K., Artstein, R., Traum, D., & Morency, L. (2014). A Demonstration of Dialogue Processing in SimSensei Kiosk. In *15th Annual Meeting of the Special Interest Group on Discourse and Dialogue*. 10.3115/v1/W14-4334

Mori, M. (1970). Bukimi no tani [the uncanny valley]. *Energy*, (7): 33–35.

Mostow, J., Ferris, M., & Brancato, J. (2009). *Surrogates* [Film]. Touchstone Pictures, Mandeville Films, Top Shelf Productions, Wintergreen Productions, USA.

Osborn, K. (2009, March 15). Army Robots: Will Humans Still Be in Control?. *Time Magazine*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Parke, F. (1972). Computer Generated Animation of Faces. In *Proceedings of the ACM annual conference* (vol. 1, pp. 451-457). ACM. 10.1145/800193.569955

Parsons, T., & Rizzo, A. (2008). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of Behavior Therapy and Experimental Psychiatry*, *39*(3), 250–261. doi:10.1016/j.jbtep.2007.07.007 PMID:17720136

Pisoni, D. B. (1997). Perception of synthetic speech. In J. P. H. V. Santen, R. W. Sproat, J. P. Olive, & J. Hirschberg (Eds.), *Progress in Speech Synthesis* (pp. 541–560). New York: Springer Verlag. doi:10.1007/978-1-4612-1894-4_43

Powell, W. (2014, September). *Time for a checkup: Health and Social Care*. *PC Pro Magazine*.

Powell, W., Garner, T., Tonks, D., & Lee, T. (2014). Evidence Based Facial Design of an Interactive Virtual Advocate. *Proceedings 10th Intl Conf. Disability, Virtual Reality & Associated Technologies*, (pp. 355-358). Academic Press.

Proyas, A., & Goldsman, A. (2004). *I Robot* [Film]. Twentieth Century Fox Film Corporation, Mediastream Vierte Film GmbH & Co. Vermarktungs KG, Davis Entertainment.

Qiao, S., & Eglin, R. (2011). Accurate behaviour and believability of computer generated images of human head. In *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry* (pp. 545-548). 10.1145/2087756.2087860

RITA. (2014). *RITA: Responsive Interactive Advocate.* Retrieved April 22, 2015, from http://rita.me.uk/

Rizzo, A., Sagae, K., Forbell, E., Kim, J., Lange, B., Buckwalter, J., & ... . (2011). SimCoach: An intelligent virtual human system for providing healthcare information and support. *Studies in Health Technology and Informatics*, (163): 503–509. PMID:21335847

Robertson, B. (1988). Mike, the Talking Head. *Computer Graphics World*, *11*(7), 15-17.

Robertson, B. (1992). Moving pictures. *Computer Graphics World*, *15*, 38–38.

Saygin, A. P., Cicekli, I., & Akman, V. (2000). Turing Test: 50 Years Later. *Minds and Machines*, *10*(4), 463–518. doi:10.1023/A:1011288000451

Sifakis, E., Selle, A., Robinson-Mosher, A., & Fedkiw, R. (2006). Simulating speech with a physics-based facial muscle model. In *Proceedings of the 2006 ACM SIGGRAPH/Eurographics Symposium on Computer Animation.* ACM.

SimGraphics. (n.d.). *SimGraphics VActor*. Retrieved August 11, 2014, from http://www.simg.com/

Slater, M., Pertaub, D., & Steed, A. (1999). Public speaking in virtual reality: Facing an audience of avatars. *Computer Graphics and Applications, IEEE*, *19*(2), 6–9. doi:10.1109/38.749116

Spielberg, S., & Aldiss, B. (2001). *A.I.* [Film]. Warner Bros., DreamWorks SKG, Amblin Entertainment, USA.

Staten, J., & Boren, B. (2001). *Halo* [video game]. Microsoft Game Studios, Gearbox Software, Bungie Software, Westlake Interactive.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Swartout, W., Traum, D., Artstein, R., Noren, D., Debevec, P., Bronnenkant, K., ... White, K. (2010). Ada and Grace: Toward Realistic and Engaging Virtual Museum Guides. In J. Allbeck, N. Badler, T. Bickmore, C. Pelachaud, & A. Safonova (Eds.), *Intelligent Virtual Agents (IVA)* (pp. 286–300). Heidelberg, Germany: Springer. doi:10.1007/978-3-642-15892-6_30

Taylor, A., Kalogridis, L., Lussier, P., Cameron, J., & Hurd, G. (2015). *Terminator Genisys* [Film]. Paramount Pictures, Skydance Productions, USA.

Terzopoulos, D., & Waters, K. (1990). Physically-based facial modelling, analysis, and animation. *The Journal of Visualization and Computer Animation*, *1*(2), 73–80. doi:10.1002/vis.4340010208

Museum of Modern Art. (2004). *Walt Disney, Steamboat Willie (1928).* Retrieved September 26, 2014, from http://www.moma.org/collection/object.php?object_id=89284

Tinwell, A., & Grimshaw, M. (2009). *Survival horror games - an uncanny modality*. Paper presented at the Thinking After Dark Conference, Montreal, Canada.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*(236), 433–460. doi:10.1093/mind/LIX.236.433

University of Cambridge. (2013). *Face of the future rears its head*. Retrieved August 1, 2014, from http://www.cam.ac.uk/research/news/face-of-the-future-rears-its-head

USC ICT. (2010). Talking with Ada and Grace. *YouTube*. Retrieved August 28, 2014, from https://www.youtube.com/watch?v=K6kcv3zwoo8

USC ICT. (2013). SimSensei & MultiSense: Virtual Human and Multimodal Perception for Healthcare Support. *YouTube*. Retrieved 2 September, 2014, from https://www.youtube.com/watch?v=ejczMs6b1Q4

Walters, G. (1989). The Story of Waldo C. Graphic: 3D Character Animation by Computer. In *Proceedings of ACM SIGGRAPH*. ACM.

Wan, V., Anderson, R., Blokland, A., Braunschweiler, N., Chen, L., Kolluru, B., ... Yanagisawa, K. (2013). *Photo-realistic expressive text to talking head synthesis* (pp. 2667–2669). INTERSPEECH.

Warren, T. (2014). *The story of Cortana, Microsoft's Siri killer.* Retrieved April 2, 2014, from http://www.theverge.com/2014/4/2/5570866/cortana-windows-phone-8-1-digital-assistant

Weise, T., Bouaziz, S., Li, H., & Pauly, M. (2011). Real-time performance-based facial animation. *ACM Transactions on Graphics*, *30*(4), 77. doi:10.1145/2010324.1964972

Williams, L. (1990). Performance-driven facial animation. *Computer Graphics*, *24*(4), 235–242. doi:10.1145/97880.97906

This research was previously published in Creative Technologies for Multidisciplinary Applications edited by Andy M. Connor and Stefan Marks; pages 99-129; copyright year 2016 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 71
How Games Improve Language in People With Language Dysfunctions

Robert Wahlstedt
Independent Researcher, USA

ABSTRACT

Many people as they age face a greater challenge of muscular dexterity around their facial muscles. This results in difficulty producing certain sounds, and sometimes the problem is so severe that they are unintelligible. People who could benefit from the methods in this chapter are those who are hard of hearing and do not have feedback readily accessible and people with ALS. This chapter describes a method that uses a computer learning algorithm that predicts what people are about to say based on earlier content and learns what the natural sound of their voice sounds like. This chapter illustrates speech trajectory and voice shaping. Clear Audio is a biologically inspired framework for studying natural language. Like the story behind Jurassic Park, Clear Audio attempts to make predictions about data from existing data, inspired by biological processes. Its main goal is to give feedback for speech pathology purposes.

INTRODUCTION

In order to comprehend meaning in a sentence the software must recognize a bit of language sound or phoneme meaning the software must then predict sentence trajectory of what is to come. In a word with more digital data than ever before, each word is a trap door to another set of meanings. Words are fossils of another time and another place brought back to evoke sentiment including thoughts or feelings of previous experiences. For example, in a paragraph that mentions "snow" denotes that the paragraph is about snow. Lieberman argues that the brain is always in the mode of social communications and that it does not turn off. Lieberman argues that although a person is not born with social mechanisms in place he or she quickly develops social capacity as the brain adapts to its environment. The human needs 10,000 hours' practice to become an expert in a specialty. The brain achieves this social learning

DOI: 10.4018/978-1-7998-0951-7.ch071

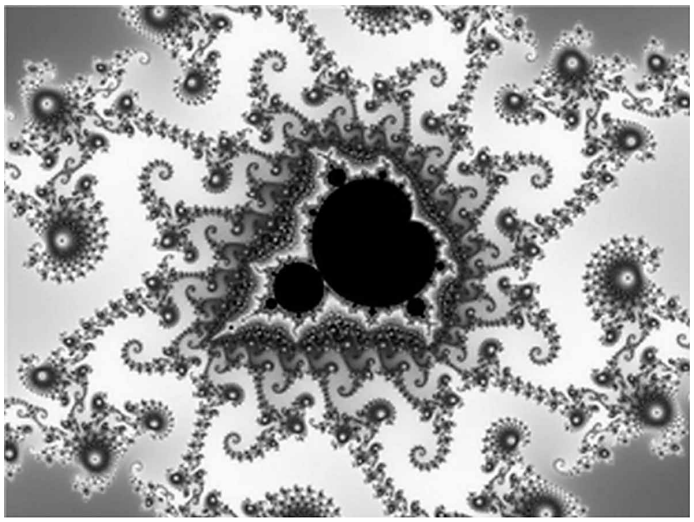
Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

by age ten (Lieberman, 2011). The Second Machine Age argues that while computers are not capable of doing anything other than what a human tells them to do, if they have an exact enough feature set, they are better than humans at determining how the proper order of events. One process is modeled by the Zachman framework of asking who, what, when, and why. For example, the Google Chafer project, a project to designed a computer to drive a car. The feature set was as good as of 2014 the car had only two accidents, one of which was rear ended at a stop light. Should the Zachman framework be used in natural sentence processors such as our Clear Audio Project can predict as some critics of natural language processing say, predict the surprising punch line of Grandpa's unpredictable jokes. Filters become more essential with the growing field of data mining and big data. Recently Google and Apple (Siri) released new versions of speech recognition in Google Now and IOS 6. These represent the best of today's big data. Traditionally data scientists compare the concept of big data to a big hard drive full of stuff that the scientists process. However, as this big pile of data grows the only people who are able to sort it have many hundreds of processor cores. Jeff Stibel points out that large internet-based companies have their servers based in cheap places to get electricity (Stibel, 2014). Cnet, in their review of IOS Maverick, claims that the catch or the storage of internet files on the macs is overbearing. Worse is the concern that having even anonymous data can have legal implications because people with unique conversations would stand out in a crowd. Aiden explains how almost in a decade worth of Google searches some queries stand out (Aiden, 2013). Particular phrases such as idioms from cultural minorities can distinguish a person's conversation. This predicament reveals a required feature set. Data without a feature set is useless. Mathematics is defined by The Oxford Dictionary as "the abstract science of number, measure, and space." Language and mathematics share the definition as being omnipresent, constant, without regards to feelings or sentiments. People argue that language is more chaotic.

Fortunately, mathematics offers fractals such as the Mandelbrot Fractal shown above to help people understand complexity and chaos in simple terms. Computers can find similarities in pictures without knowing what they are. Recently Peter Norvig of Google gave a talk to Brown University about memorization or a way that the computer can know what an image consists of although it may not know the

Figure 1. Shows a computer rendering of the Mandelbrot fractal (credits go to Wikimedia)



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

name of these items the computer found similar. In Surfaces and Essences, compares an object being placed in a box and then taken out of the box to words being conceptualized and formatted. Consider all the meanings of the word band. This can connote everything from a marching band to a wedding band to a Band-Aid that filters air to a wound so it is not infected. Which one of these is a metaphor? We can never tell which usage of the phrase is the original as opposed to the ones are comparable to the original fossil of the metaphor. With the Zachman framework, we come close to determining the usage of phrases. While metaphors are an important aspect of word origins, writers of newspapers write articles so people with a ten-year-old child's vocabulary and intellect can understand them. We cannot assume that many ten-year old children can understand metaphors such as a broken record or see an activist as something that acts upon an agenda. This network of vernacular fossils composes much of our language today. Therefore, something else is happening. Iacoboni asks why is it that a person gestures when they are talking on a phone and it the other person cannot see these gestures. It is because we use words as a vehicle to transport the other person to a thought or feeling we wish to convey (Iacobon, 2008). Using this method, I will refer to as capture and transport method, in chapter two of this paper, we discuss how educational games and by letting people have experiences using augmented reality to reinforce Hebbian learning of experiences. What we need to find out is how spelling constructs etymological concepts. If we were strictly, follow a formula for the English of yesterday of pragmatically tagging parts of speech (P.O.S.) we would never be able to keep up with new words as the as people use these words new to our corpus. In the later part of the last century, stock traders worked on algorithms to decide in which direction stocks will go. A key advantage was that stocks exist in numerical form. However, this does not relinquish the marvel of predicting what would happen next. Natural language processing researchers can take lessons from stock market predictions and apply them to both word and concept formation. Humans tend to repeat language behaviors if they were effective or energy-efficient or economical. The researcher's goal is to use the Maven Markup Modeling Language (MMML) based on predictive modeling markup language (PMML) to find ways of correlating words to integrate new words faster into the corpus. Text mining projects at Universitat Politecnica de Catalunya depend on lexical categories for their lexical analysis tagging. Yagoda suggests "According to grammarians, adjectives, nouns, verbs, and injunctions are 'open' parts of speech because they shift functions ... and because new words are continually added to their ranks. This proves that classification of words is not always necessary." (Yagoda, 2006). To make matters worse, a gerund is using a verb in its – ing form as a noun such as living. Although the English language has synoptically, there are two types of words, content words and structure words. The natural language processing engine can remove structure words such as "hope that" "clearly" "strangely"" indeed" "conceivably" "seriously" "ultimately" "theoretically" "naturally" "ironically" "fortunately" incidentally" do not containing meaning in themselves however they send signals to listeners about words that are coming later. The word "like" is overused because it is often used as filler words while the person thinks of something else to say. The National Science Foundation did a study that demonstrated when a storyteller used the word "this" instead of "a" or "an" the person demonstrated better retention. Another concern is that these systems do not account for pragmatically that is the study of context is it situational.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

BACKGROUND

There are two classes of people with speech challenges; the first class consists of those who had trouble speaking since birth. They are aware of the fact they are more challenging to understand. One hypothesis on speaking challenges is the disorder is more prevalent in children because the work of both conceptualizing and the formatting tasks of language overwhelms them. Bell Labs engineer Robert Lucky estimated the cortex cannot take more than 50 bits per second. Mihaly Csikszentmihalyi estimated one apperception to take 1/15th of a second or 105 bits per second. Regardless of the figure a person has limited bandwidth or computing power. Because it takes a lot of concentration to come up with metaphors and categorize ideas with words, the quality of speech declines. Along with this speech has limited prosody referred to in this paper as words per minute, pitch variance, and stress of syllables. With the knowledge that they are unintelligible, they compensate using repetition or choose words that do not contain the element that they have trouble with. They may do these substitutions or repetitions unconsciously. People with speech aphasia often have trouble with pronouns. A goal of Clear Audio is to find the use of pronouns and the repetition by using summarizing tools. Section 2 explains these methods. Section 3 describes other people who have difficulty speaking later in life due to a stroke are confident that their pronunciation is the same as it was before their decline and may not utilize either repetition or choice of words.

WHAT DO COCHLEAR IMPLANTS SUGGEST ABOUT SOUND?

Helping Those Who Repeat to Be Understood

In order for a contextual agent of software that recognizes speech, it must develop a cohesive hypothesis about what a speaker is saying. Denworth discusses the requirements of language for cochlear implant patients. She suggests that processing sound is a multifaceted event including both contextual clues for perception and the stemming. Denworth introduces David Pisoni's work on "visual speech" that Denworth describes as the act of a person to see objects without cognition of what they are. Densworth gave this as a reason her son has an implant starting at age three (Densworth, 2014). While experimenting with Audacity, we realize how redundant speech is as opposed to music. A more sporadic sampling of voice is enough for voice while for music every sound we must accommodate for. Summary generation for detecting repetitions A piece of literature is similar to a piece of music in that both contain repeated themes throughout the piece. While the BBC has a recording of Tchaikovsky's 1812 overture that lasts 17 minutes, most people remember only the theme. In this paper, we discuss the methods how literature themes stand out from the rest of the piece through both exact repetition and supporting variations of the theme. Some scientists who study the brain call this grouping "chunking." Many textbooks of text mining hold the author's beliefs that word categories or parts of speech are important. These papers create summaries by deleting irrelevant words until only the nuclear core remains. However, these authors overlook context. A new approach seeks to use structural knowledge of pragmatic through latent semantic analysis. For example, to get to the best diagnosis of patients, it is important to take into account their entire medical histories and not just a few quotes. This method does analysis on the entire document before concluding so metaphors and words with multiple meanings are not confused. This paper proposes if using metadata, or data about the data, to distinguish repetition of the theme from supporting variations,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

also known as satellites. This is achievable by finding the roots of words through their origins, translating the word prefix, root, and suffix into metadata. The brain constantly encounters more information than it can digest. Oliver Sacks discusses how Chinese speakers represent a higher than average percentage of people who have pure tone. Of these, the greatest were musicians who started learning their instrument before age six. This is because the tonal system of the Chinese language is a very important element and is necessary to distinguish words from each other. The brain of the Chinese speakers adapts so it is accustomed to listening to other Chinese speakers (Sacks, 2007). At the time of birth, a person is born they can absorb much information from their senses including hearing. Around 80 percent of the energy they get from food goes to developing their brain. Swaab discusses one theory that a person uses the equivalent of $1500 for their entire life or 15 watts per hour (Swaab, 2014). In order to come up with a computational model we need to ask the right questions. As they start building heuristics and finding out what is important and what is not then less energy goes to fuel for the brain. Dean Buonomano commented on how musicians dedicate larger than average part of their brain to muscular memory in their fingertips through reallocation. He goes on to say the brain does not consciously govern the real estate in the nervous system so this was a subconscious process. Why do English speakers who are not musicians have such a seeming disadvantage? One dominant theory is that of the Hebbian Synapse that neurons that fire together wire together. In other words, if two events are not linked by the human brain as related, the brain acts like a filter and events are never associated or learned with one another (Buonomano, 2011). Stanislas Dehaene takes this notion further and discusses the history of the mind and body model. In the ancient world, philosophers like Descartes often thought that the heart was the headquarters of human thought (Dehaene, 2014). The soul or spirit was a detachable element from the physical body. This is similar to today's notion between the difference between computer software and computer hardware. Later scientists developed interfaces for working with the injured brain to capture electrical signals. These in addition to fMRI have introduced new material under review as far as what a person can understand. Through the new methods, we can detect electrical signals of which the brain encodes information into neurons. We now have empirical evidence that people see other people through filters and this becomes imprinted on their brain based on prior experiences. For example, we know that when person has had a negative experience with big dogs, certain elements the perception system exaggerate such as teeth, when the person encountering future dogs. This conceptualization is utilized in visual speech.

In designing a program, it is necessary to understand what one has to work with. Imagine a machine with two cameras perceiving words. When a person reads a word like "unbuttoning" the human brain strips off the "un" and the "ing" and it is left with two morphemes, "but" and "ton." There is a priming effect where a person is more inclined to think about a certain flow of words. For example, after casa, there sometimes is the word Blanca like the movie. This is why we understand differing homograph meaning. A grapheme is a letter or a series of letters that map a phoneme in the target language. Imagine a web of words that links all words in existence. Pulverm introduces a word-related functional web. It is a synchronous firing chain or synfire chain (Pulverm 2002). At first glance, automata might be similar to this web. However, a computer differs from a brain because even computers with multiple processing cores are asynchronous in the sense that the computer can only process a few threads unified by a central process. The human brain is a democracy having processes happening simultaneously and the strongest impulse, usually by a collective number of neurons fire that is what the human brain thinks. Miguel Nicolelis argues there is not a separate part of the brain language support.

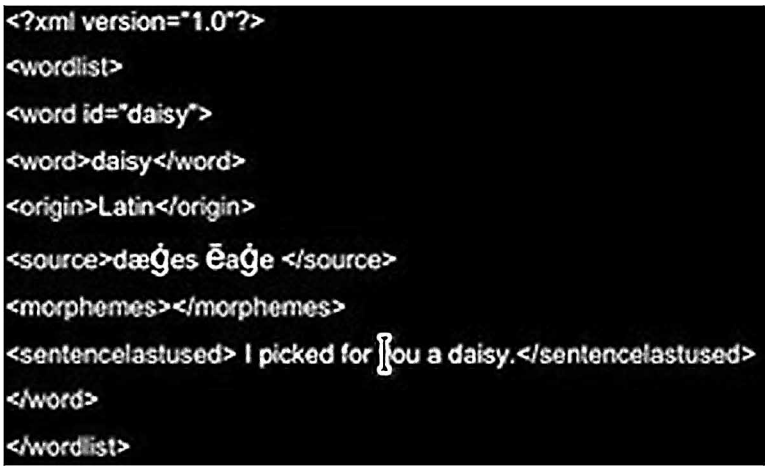
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

However, a certain sequence of neurons that light up when a person reads about an object because it invokes an emotion (Nicolelis, 2011). Kenneally describes how non-humans animals could use simple language. Instances of non-humans are territorial warning signs and primitive inner species communication (Kenneally, 2007). The article "How Dolphins Say Hello" says that dolphins only whistle when they are in a group of dolphins and find themselves through echolocation. Grammar is a set of rules in a language that allows people to communicate and understand. There is a debate between B.F. Skinner and Noam Chomsky. Skinner believes that a person learns language through association, the sight of things along with the sound of the word. There conformity of this association matching constructs a dialect. Chomsky believes that a Martian scientist observing children in a single-language community would see that language is very similar regardless of the culture and therefore innate. Although reading is unnatural, our brain wants to see patterns and group objects together. We recognize lines and shapes that we read as characters. We then group these characters into morphemes. We then put these morphemes into phrases, idioms, words, and sentences. These sentences form a paragraph. Similar to computers recognition is easy, recalling meaning is difficult. Henry Hitchings discusses how it is possible to use stemming while the human language evolves. He says it is amazing that people spell words uniformly. Before computers, he points out that even dictionaries were inconsistent (Hitchings, 2011). There was a survey where a computer used phonetics to spell out words and 50 percent of the words did not agree with the phonetic spelling. Steven Pinker tells us that 84 percent of the words have a spelling that conforms to patterns we can notice. As shown by these statistics, the ability to conform is possible (Pickner, 2007).

It is necessary to study of metadata to correlate etymology with words. Etymology is the study of finding word roots and morphology is the study of word making. Today advances in bioinformatics tools have yielded a complex study of the human genome and have made great strides in finding out mutations that can create a higher chance of cancer. We should do so because we can learn about the human race as well as expand our knowledge of words. Here are some examples: nickname originally came from nekename containing the compound eke and name. Eke came from ēac that means also. It means an extra name. The word "hobby" comes from the word hob and yn. Hob or hŏb means a threaded and fluted hardened steel cutter, resembling a tap, used in a lathe for forming the teeth of screw chasers, worm wheels (Webster's Revised Unabridged Dictionary) and yn means "to be." It means to spin time. The word "omelet" comes from the French word omelette. Omelette comes from alemette that comes from alemelle that as comes from la lemelle that means a thin plate like structure. We should know how to find the prefix, root, and suffix because we are able to understand why someone in history responded to situations. The written word often transcends a person's lifespan and the culture changes. A written manuscript is similar to a fossil. For example, rabbit comes from the words robète that means to steal. This explains why Mr. MacGregor thinks Peter Rabbit as such a nuisance. If you are a person that is a Christian, it might please you to know that Nicodemus comes from the word nincompoop. This comes from the Latin phrase non-compos mentis meaning not of sound mind. People regard him as a hero of the Christian faith because he walked by faith and not by knowledge. What does a computer have to do with this? Computers are able to carry out arithmetic or logical operations such as regular expressions or pattern matching. Given some rules and a lexicon file, they are able to study morphology. "A" as in rag or Prague is big and "I" as in pin am small, "e" is somewhere in between. There has been some big vowel shifting in England between 1350 and 1500 known as the big vowel change so words like big and small have the wrong letters. Little and large are right. Pimpf means "a little boy," pimple means a little swelling, and pampers means a lot. Servant is in between. Quick Summary uses Maven Metadata Markup Language or MMML. It has a structure similar to the one shown below.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Shows the XML layout of the word "daisy" in Maven markup modeling language (MMML)



It may seem everything the software needs are stemming or taking apart words to find their meaning. The book Found in Translation: How Language Shapes our Lives and Transforms the World, talks about how words are associated with different metaphors in different cultures. One example they give is intaxicato solar. This means food poisoning. It looks as though it means intoxicated as though someone had too much to drink. This is an example of what linguists call a "false friend" also known as a false cognate meaning that it looks as though they come from the same linguistic root but have different meaning. Another example is the word okoru in Japanese. Although (Truss, 2004). The word leaves can also be from a plant or it can mean to abandon an area. These point out in addition to stemming the software needs to run on a distributed blackboard approach to artificial intelligence. A blackboard is a comprehensive method that has several expert knowledge agents, they feed an aggregate stream with what the agent knows based on a limited perspective, and each agent fills in insufficiency of different agents. In our software, the stemming is one agent and context of the previously spoken topic are another. In addition, features should consist of tone or pitch, the tempo of words were spoken or speed, and vibrancy or liveliness. Pentland introduces a social-o meter that has a microphone for determining pitch and speaking duration. However, this microphone is unable to decide the real words (Pentland, 2014). A key distinguishing reason of the research presented from other software that can recognize speech such as Dragon Speaking Naturally Dictate has pride itself with how many sample voices have trained it. These speech recognition engines do not consider how important tonality is in making speech palatable. The brain needs some difference, dissonance, and variance to stay focused. When a robotic voice speaks, it often takes greater amounts of concentration to listen. In today's high-paced world, people might not always have energy enough to listen. Sounds constantly bombard with data and as we discussed before in this paper under the section A Human's Big Data Filtering Method, the brain does a good job of filtering out what are signals as opposed to noise it looks as though it may resemble the word to occur, it means to get angry. The book Found in Translation goes on to say that, certain brands like Mitsubishi have to rename their Pajero module and Honda Fitta module in particular countries. The countries have different connotations of the words. The word gift means poison in German. The Online Etymology Dictionary says that it was often associated with prescribed medicine. People used the word gift as in a potion by a doctor and it came to mean something of tangible from a knowing person. Found in Translation tells

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

how jokes in some languages are missed such as in the Harry Potter Series Lord Voldemort is "Tom Marvolo Riddle" is an anagram for I am Lord Voldemort. The Bulgarian translation is Mersvoluko whose anagram translates to "And here I am, Lord Voldemort". The word in Hebrew originally was tohu meant formless and vohu meant empty. In French today toha bohu meant chaos and confusion. These words became associated with chaos later in Hebrew (Kelly, 2012). Sometimes the environment makes it necessary to invent words to present the concepts behind them. For example, to explain to the Hmong people about cancer at UC Davis, they compiled English to Hmong dictionary for medical terms. Martin Luther when translating the scriptures invented words so he can explain to people Latin concepts such as Machtwort (authoritative guidance). Truss indicates this problem by having multiple definitions for the word shoots meaning to fire a fast projectile from a gun or shoots that are a part of a plant (Truss, 2004). The word leaves can also be from a plant or it can mean to abandon an area. These point out in addition to stemming the software needs to run on a distributed blackboard approach to artificial intelligence. A blackboard is comprehensive method that has several expert knowledge agents, they feed an aggregate stream with what the agent knows based on a limited perspective, and each agent fills in insufficiency of different agents. In our software the stemming is one agent and context of the previous spoken topic are another. In addition, features should consist of tone or pitch, tempo of words spoken or speed, and vibrancy or liveliness. Pentland introduces a social-o meter that has a microphone for determining pitch and speaking duration. However, this microphone is unable to decide the real words (Pentland, 2014). A key distinguishing reason of my research is that other software that can recognize speech such as Dragon Speaking Naturally Dictate has pride itself with how many sample voices have trained it. These speech recognition engines do not consider how important tonality is in making speech palatable. The brain needs some difference, dissonance, and variance to stay focused. When a robotic voice speaks, it is often takes greater amounts of concentration to listen. In today's high-paced world, people might not always have energy enough to listen. Sounds constantly bombard with data and as we discussed before in this paper under the section A Human's Big Data Filtering Method the brain does a good job at filtering out what are signals as opposed to noise. to noise. Those who hear a 60-hertz drone of a light fixture will likely ignore the drone because the brain has trained itself to ignore some sounds. Next, I hope I am able to find a feature set to annotate and make computer voices easier for people to listen.

In the Story of English in 100 Words, David Crystal discusses the origin of words. He gives the history of the word jail borrows from the French originating in the about the time of the Norman invasion of 1066 study we can tell because the word gaol is Latin. Crystal explains that languages borrowed this word twice, being what Crystal refers to as double borrowed. We can tell this because the word did not get its origin meaning. This is also true for the word convey originality coming from the French word convoy. Our approach is to take this research and combine with what is known about UML, I* and the Zachman framework to discuss the questions of what, where, when, why, who and how. With this data, it is possible to verify that our summary correctly reflects the emphasis of the paper. In Empires of the World: A Language History of the Word, Nicholas Ostler demonstrates beginning with ancient Babylonians carrying off the Hebrew slaves and forcing them to melt into the culture of the Babylonians how language can melt together. John McWhorter writes this is what happened during the Norman invasion. There was not a bloody war instead slowly, over time the Normans introduced French. English, Crystal says, is like a Our approach is to take this research and combine with what is known about UML, I* and the Zachman framework to discuss the questions of what, where, when, why, who and how. With this data, it is possible to verify that our summary correctly reflects the emphasis of the paper. In Empires of the World: A Language History of the Word, Nicholas Ostler demonstrates beginning with

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ancient Babylonians carrying off the Hebrew slaves and forcing them to melt into the culture of the Babylonians how language can melt together. John McWhorter writes this is what happened during the Norman invasion. This field of anthropology is transcultural transfusion (McCrum, 2010). This is an example of diffusion by choice. By studying hyperdiffusionism under the belief that man originated in one place, we can learn through cultural similarities. For instance, most civilized cultures today teach us that it is wrong to kill or commit adultery. Noam Chomsky writes about a universal grammar being indistinguishable by a Martian linguist. McWhorter claims that through DNA evidence anthropologists discovered the village where the anthropologists discovered the village where the Germanic dialect began. In tracing word etymologies, it is important to ask ourselves where features of a language began. Guy Deutscher makes the following observation: "Often, it is only the estrangement of foreign tongues, which their many exotic and outlandish features, that brings home the wonder of language's design." By incorporating origin into a NLP translator, we can see by parapraxis, images that speaker has in mind. Through a more careful of word choice, we have the advantage of seeing the speaker's words and come closer to making up for the disadvantage of not seeing him communicate those words to us. We can see how a language would transform and predict new words that might occur. By understanding how the introduction of new words and in what context it helps us predict what the likelihood that a speaker is, uttering these new words and this can greatly help the software that recognizes speech, sucking up words around it. McWhorter likens it to summer camp, where an exchange of ideas takes place. Robert McCrum writes that non-native speakers to surpass the limitations of English by borrowing from other languages use Globish (McCrum, 2010). This field of anthropology is transcultural transfusion. This is an example of diffusion by choice. By studying hyperdiffusionism under the belief that man originated in one place, we can learn through cultural similarities. For instance, most civilized cultures today teach us that it is wrong to kill or commit adultery. Noam Chomsky writes about a universal grammar being indistinguishable by a Martian linguist. McWhorter claims that through DNA evidence anthropologists discovered the village where the Germanic dialect began. In tracing word etymologies, it is important to ask ourselves where features of a language began. Guy Deutscher makes the following observation: "Often, it is only the estrangement of foreign tongues, which their many exotic and outlandish features, that brings home the wonder of language's design" (Deutscher, 2011). By incorporating origin into a NLP translator, we can see by parapraxis, images that speaker has in mind. Through a more careful of word choice, we have the advantage of seeing the speaker's words and come closer to making up for the disadvantage of not seeing him communicate those words to us. We can see how a language would transform and predict new words that might occur. By understanding formation of new words and in what context it helps us predict what the likelihood that a speaker is, uttering these new words and this can greatly help the software that recognizes speech.

How to Assist Those With Speech Issues That Do Not Overcompensate

In people with speech issues who obtained them later in life due to a stroke it is important to utilize elements of traditional speech recognition. These include language based game theory and eliminating what is not the topic. This paper proposes imitation of a biological system.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Language Based Game Theory

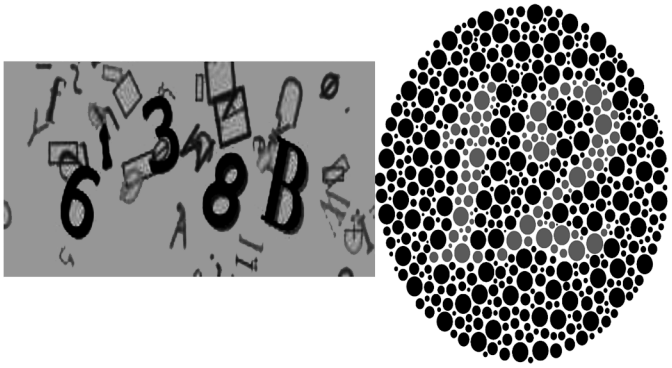
A project in the Netherlands led by Jan Dietz seeks to map words as they relate to events in the world. Their premise is that we are like parrots repeating ideas that come to us through movies and books. In a game of broken telephone, a group of people sits in a circle and whisper a message to the person next to them. The message tends to vary as the first person conveys the message to the next person. These researchers claim through recursion we repeat these events demonstrated to us in a particular order. They believe they can forecast behaviors based on the past. The DEMO research group claims that BPMN, UML-AD, EPC, IDEF-3 are method independent. They do not take into account the redundant nature of business. It captures the conversations between actors. Jan Dietz explains that the subject chose to share one of their thoughts called the locutor. It appears as <locutor>: <illocution>: <addressee>: <fact>: <time> For example, let us consider checking a book out from a library. Certain circumstances must satisfy the checkout conditions. First, to check out books the subject must be a member at the library. Then he must not have any pending fees. Essence, the attributes that make the subject what he or she is. He is a candidate who can pick books out of the library. Should there be problems, he cannot check out books from the library. The production act of picking books out of the library is the executioner. It is a repeatable act. David Bellos tells an account of how his father was able to check into a hotel in a non-English speaking community because both he and the hotel staff were familiar with the process (Bellos, 2011).

Another tool that can help with the speech is to find the topics not discussed. As the speaker is generating a narrative the chance of the speaker continuing to speak of a topic becomes greater. This directs the trajectory a certain way. Throughout the conversation, we use the Zachman framework to generate a possible persona of who the speaker is. If the speaker is a 16-year-old, we know there is a greater chance of them looking into colleges as opposed to retirement homes. While in certain situations, this can happen. The computer takes the conversation as a puzzle which has certain elements and can affirm or discredit the personas generated by the computer of who is speaking. To confirm accuracy, information can be gathered through abnormal changes in sediment. There is an emerging field of looking for statements and determining whether they are joking or using sarcasm. This technique is similar to those in visual saliency to help us deduct clues. Barash speaks of language like a stream of water (Barash, 2003). An adversary comes from the Latin word adversus meaning turned towards the flow. Language consists of hundreds or thousands of transactions and our program can learn what to expect. Language is much like the social behavior of ants. Can computers learn enough by watching transaction to find what stands out among the ordinary. The article Ant-Inspired Visual Saliency Detection, gives the image of ant behavior when they take the way that requires the least amount of effort to get towards their food. Saliency provides a method of processing of images for anything that stands out including edges. There is software for finding salience such as the Saliency Toolbox and the Neuromorphic Vision Toolbox. Consider the test for color blindness where a hidden object in an image or CAPTYA in figure one and two below.

Consider the book, Where's Waldo. In the Where's Waldo example the stimulus is when we Consider the book, Where's Waldo. In the Where's Waldo example the stimulus is when we see something that resembles Waldo. For example, the viewer considers the color red with Waldo. Then on to the next level, they consider red stripes. Then they associate glasses. The template starts to form that confirmed by people of different cultures is a grammar. It is possible by using a grammar to spot what the topic is not, much like a Bayesian spam filter. This monitors the flow of the conversation and when an unexpected sound appears in the recording it the software can quickly process the sound by assessing it and managing it. Artificial Neural Networks

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. Two hidden objects one in a CAPTYA and the other in a color blind test

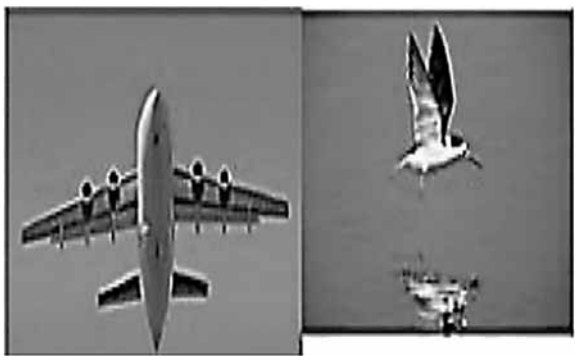


As discussed before in this section automata are not an adequate algorithm for defining morpheme parsing. A neural network is a statistical algorithm used in machine learning. It is simple yet effective. Biological neural networks inspire artificial neural networks. Neural networks are adaptable and can self-organize in real-time. They also give fault tolerance. The brain forms the basis of an artificial neural network. Consider the images shown in Figure 4. These are both images of a flying object.

The one on the right inspires the one on the left. They both have two wings and a tail and are both of the same air stream figures. When the inventors like Leonardo De Vinci was constructing a flying machine it very much resembled a bird. Similarly, a neural network is a natural model of a brain (Abu-Mostafa, 2012). It is self-organizing, fault tolerant, and adaptable. We need to build a software design that resembles this working model. Looking at the PyBrain code it looks as though we can and should adopt something like this for our sentence processing sequence. In How the Mind Works, Steven Pinker discusses the benefits of using a neural network; however, linguists need to improve multiple instances of the same object.

The one on the right inspires the one on the left. They both have two wings and a tail and are both of the same air stream figures. When the inventors like Leonardo De Vinci was constructing a flying machine it very much resembled a bird. Similarly, a neural network is a natural model of a brain (Abu-Mostafa, 2012). It is self-organizing, fault tolerant, and adaptable. We need to build a software design that resembles this working model. Looking at the PyBrain code it looks as though we can and should

Figure 4. A human invention (the airplane that is to the left) compared with a bird



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

adopt something like this for our sentence processing sequence. In How the Mind Works, Steven Pinker discusses the benefits of using a neural network; however, linguists need to improve multiple instances of the same object.

FUTURE RESEARCH DIRECTIONS

I hope to find a feature set to annotate and make computer's speech easier for people to understand. I am working on looking at integrating the Neuromorphic Vision C++ Toolkit, a vision toolkit, with GBBOpen, an open source blackboard system, and Pybrain, an open source artificial neural network framework. I hope I am able to understand the algorithms well enough to port them all to python and so it can run on an Android. Kivy is an open source virtual machine containing libraries on how to write crossplatform python applications. GBBOpen uses lisp and the Neuromorphic Vision C++ Toolkit uses C++. The requirements are an algorithm that processes events in parallel and learns from reinforced learning. A brain can be thought of as a multi-core or quantum computers that resemble neurons. Change blindness as described in the book The Invisible Gorilla, talks about how we long for continuity (Chabris, 2010). However, in this paper I argue that this pattern matching classification is done to overlook elements in sound. Naturally we each have a certain range of activities that go unnoticed. Sometimes this varies from person to person, a common example being tinnitus. Like light, sound is often measured in how it is different from one element to the next such as pitch. Sound appears unified until broken apart by a prism called a spectrometer. This helps us figure out what the light emitting object is constructed and composed of. We are unable to tell boundaries are twofold. First the mind likes unifying patterns into things it can understand and find patterns between two "corresponding" events. This is so we can communicate what it is we saw. This is despite the entropy and separations posed by natural limitations such as quantum mechanics in harmonic wave functions The second reason is pink noise also known as Brownian motion or noise that we disregard as background noise. This is what audiologists call masking and can dampen boundaries between two events in sound. However, I am uncertain that most pink noise is extraneous. It could provide useful clues as to the context of the speech by harmonic analysis and studying representational theory. This can be explored in the study of continuous-time Markov chain. This uses the analogy of how elements combined in a certain way makes a chemical reaction. Grammar is different and cannot be measured by finite states but rather provide a framework of rules for processing. Rather it can be seen as a horde of ants interacting and otherwise affecting their evolutionary walks. Noam Chomsky talked about how rules do not necessarily guarantee a sentence. For example the notion that "Green colorless ideas sleep furiously" is a sentence is simply unsound. David Ben-Zvi uses a concept he describes as the x-games to describe infinitesimal generator matrix by using differentials to study the elements in an evolution of sounds. In my papers I've discussed the benefits of starting to play musical instruments as a younger person (some would argue ideally preschool aged). This is because it raises the awareness and attention of the threshold between two different events musically.

The blackboard algorithm combined with a neural network is a good approach to the complexity of sound. To fully appreciate the significance of each node, we not only need what is preceding it but what is going on simultaneously that give clues about what might happen in the future. In my paper I suggest that one reason people have difficulty speaking is that their mind is not able to function and keep up with everything. In this area I believe computers can excel because as computers become more powerful, they can surpass human's attention spans. The ants marching in a row is a nice picture but seldom

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

happens because of the simultaneous complex world we live in. I am confident that with the emerging field of multiprocessor concurrency and abandoning the idea of having a master core and adopting a system more like the brain of adopting many neurons firing electrical signals would bear success in the years to come. There is a consensus and each core is aware of what is going on.

CONCLUSION

Much work is necessary to give a feature set to describe the spoken natural language. This paper intends to give a framework of how to think of big data as a filter analogous to a prism that a person can see more clearly through and to ask right questions. The paper discusses why examining why people talk the way they do is important. The paper proposes that the Zachman framework and through deducting what a topic is not aid the process. This paper also examines how a human understands language and sound.

SOLUTIONS AND RECOMMENDATIONS

Having a stroke is difficult because part of the brain may stop functioning. This affects the part of the brain that produces language. To counter this, engaging games that target the mind using neuro-gaming which may include augmented reality should be developed. It has been shown that video games and television can change people's perspectives about what words mean. This can be both used positively and negatively and some of the pioneers in the field include advertisers. Some examples of media which helps people learn words include Sesame Street while those who use this power negatively do so for selling products to enhance how people view their body. What I am proposing differs from competitors because does more immersive experiences similar to those played on PlayStation or GameCube. In order for a person to truly exercise they need to be caught up in the passion of the event and forget that they are exercising. It's like the book "See Spot Run" as compared to "Jurassic Park". You can base a movie on Jurassic Park but I think I'd save my money on the spot video. Athletes can get caught up in a sport such as tennis and forget they are working out The people who could benefit from this the most, the people with early signs of memory loss and mentally illness, are the ones who cost is a barrier to using other platforms. Learning should be based on applicable games based on learning. This differs from conventional mind games because feedback is given and expounded not only showing possible solutions but they take the user's attempts and building upon them so they are guided to a solution. The audio clarifier for the phone works during the assessment phase and does not give feedback. MIT Professor Eric Klopfer says a game that much like a spelling bee gives feedback on whether the answer is correct or not but frustrates the participant by making them feel incompetent. To feel competent the participant must be guided towards a solution. The game's purpose is functionality and how it works to assess whether the person knows about the concept and situation which the clear audio needs to be activated to shape the words. The goal is to become more comfortable taking risks.

The part of the brain lost caused by strokes make it necessary for other parts of the brain to compensate. The parts which compensate may be dormant but they still exist. Language is a foundation about how we communicate with others and often is a way of conveying thoughts that are within us. Studies demonstrate that find that people who have somniloquy or are sleep talkers. WebMD finds that 10% of children make utterances in their sleep or when they are in the process of waking up (Talking in Your

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sleep). When a person is asleep in the past they often thought that the brain activity waves would shut down, however recent studies in sleep found that a person uses the brain as much during the night as they do during the day. Eagleman tells of how a patient had hemispherectomy or part of their brain removed to prevent epileptic episodes. As a side effect, they were unable to create new memories but the memories that they had remained. Sometimes memories are activated during the night, however the part of the brain that determines whether the implications of the dream is not functional. As a result of some dreams perceptions change and sometimes people wake up irritated with another person because of an altered inaccurate perception. However, language can come up with ideas. During the reign of the communist leader Nicolae Ceauşescu of Romania and the results of the way he ran his state-run orphanages made clear our dependency on language to communicate ideas. He wanted to increase the number of workers in Romania to increase revenue into the country. Not only were abortions illegal but each family that did not produce at least five offspring were taxed heavy taxes. As the economy collapsed, the shortage of money caused many families to abandon their children in orphanages. In these state operated orphanages, when there should be three care takers for every 15 kids, there was only one. After the fall of the communists in 1989 the children that were in the care of the state grew from 47,405 orphans in the care of 25,870 orphanages in 1990 to 62,000 in the care of 19,000 orphanages. The caretakers were told not to pick up or show any affection to the kids that were in their cribs because of the fear that should the caretaker show affection to one, they all would want attention and draw the energy out of their limited number of caretakers. In 2009 BBC journalist Chris Rodgers was appalled by conditions and found that the orphans spoke a coarse language they made up themselves with the purpose of communicating among themselves to survive.

What does this have to do with patients who had a stroke? The answer is that human language is a foremost tool in comprehension and our survival. Suskind argues that language is the foremost building block of our education system (Suskind, 2015). Noam Chomsky argues that language is itself innate because people from other cultures communicate in much the same way universally. The book, American English compendium: a portable guide to the idiosyncrasies, subtleties, technical lingo, and nooks and crannies of American English by Marv Rubinstein makes the point that there once was a gap between people of different countries, even American English and British English were distinct while Canadian English were a blend between the two. Because of media this gap has somewhat closed, however many phrases are still based on historic culture. Dana Suskind cites Hart and Risley's study showing people who are more successful later in lives have more interaction with their parents that are memorable because they are positive. The people who did not do so well in life according to their study suppressed negative memories of interaction with their parents which were fewer (Suskind, 2015). Questions arise such as who has memories that are dormant that could benefit most from my study into games.

Games are beneficial because they involve multiple inputs of sensory information and can help people become acquainted with Hebbian learning that is that neurons which fire together wire together. Like many professional golfers who learn through muscle memory, because of the intense thought process behind processing language in both mechanically speaking as well as comprehending the conversation the brain often differs from a computer in that once a pattern is found, the brain goes into autopilot. This is why experts sometimes find it difficult to explain their decisions in a series of rules. While people may think that programmers are unequipped to know what an expert knows in topics on medicine, since they are outside the field they have the advantage of a recently acquired understanding of what they had to go through to get knowledgeable about the skill. Music provides a good example. To a person who began taking music lessons early in life, they are unable to, without some practice; describe to someone

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Shows a Tamagotchi pet (credits go to Wikimedia)



how to start making music. When the stroke occurs and kicks the stroke person out of autopilot mode acquired during Hebbian learning. Many of the tasks which seem so familiar must now have renewed thought and concentration to each event as though they are relearning how to become acquainted with human life at infancy. The goals are to take what motivates each gender and age group and helps them become motivated towards their goals. We are fortunate to live in an era of simulation games such as Opensimulator or Second Life. These are like Tamagotchi pets, which limit the input, and makes what is expected easier. For example, the Tamagotchi pet shown in figure 5 only has three buttons and from the context we can simulate the entire lifestyle of a dog or a cat.

Computer programmers are used to the concept of intelli-sense which hints given the context about what the person is about to say. However, it is imperative to make sure that it does not frustrate the person with the stroke.

REFERENCES

Abu-Mostafa, Y. (2012). *Lecture 10 - Neural Networks*. California Institute of Technology.

Adversus. (2011). *Wiktionary*.

Aiden, E., & Michel, J. (2013). Uncharted: big data as a lens on human culture. New York: Riverhead Books, a member of Penguin Group (USA).

Bellos, D. (2011). *Is that a fish in your ear? Translation and the meaning of everything*. New York: Faber and Faber.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. Academic Press.

Buonomano, D. (2011). *Brain bugs: How the brain's flaws shape our lives*. New York: W.W. Norton.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chabris, C. F., & Simons, D. J. (2010). *The invisible gorilla: And other ways our intuitions deceive us*. New York: Crown.

Crystal, D. (2012). *The Story of English in 100 Words*. New York: St. Martin's.

Dehaene, S. (2014). *Consciousness and the brain: Deciphering how the brain codes our thoughts*. Academic Press.

Denworth, L. (2014). *I can hear you whisper: An intimate journey through the science of sound and language*. Academic Press.

Deutscher, G. (2005). *The unfolding of language: An evolutionary tour of mankind's greatest invention*. New York: Metropolitan Books.

Deutscher, G. (2010). *Through the language glass: Why the world looks different in other languages*. New York: Metropolitan Books/Henry Holt and Co.

Eagleman, D. (2015). *The Brain: The Story of You*. Academic Press.

Hitchings, H. (2011). *The language wars: A history of proper English*. New York: Farrar, Straus and Giroux.

Hofstadter, D., & Sander, E. (2010). *Surfaces and essences*. Perseus Books Group.

Iacoboni, M. (2008). *Mirroring people: The new science of how we connect with others*. New York: Farrar, Straus and Giroux.

Kelly, N., & Zetzsche, J. O. (2012). *Found in translation: How language shapes our lives and transforms the world*. New York: Perigee.

Kenneally, C. (2007). *The first word: the search for the origins of language*. New York: Viking.

Kenneally, Kessler, & Topher. (2013). *Troubleshooting Enhanced Dictation in OS X Mavericks - CNET*. CNET.

Levitin, D. J. (2014). *The organized mind: Thinking straight in the age of information overload*. Academic Press.

Lieberman, M. D. (2013). *Social: Why our brains are wired to connect*. Academic Press.

Nicolelis, M. A. (2011). *Beyond boundaries: the new neuroscience of connecting brains with machines--and how it will change our lives*. New York: Times Books/Henry Holt and Co.

Norgvig, P. (2013, October 25). *2013 Family Weekend Keynote: "Live and Learn: How Big Data and Machine Learning Power the Internet"*. Retrieved January 07, 2017, from https://www.youtube.com/watch?v=vVprmVsCV_k

Sacks, O. (2007). *Musicophilia: Tales of music and the brain*. New York: Alfred A. Knopf.

Sleep Talking Causes and Treatments. (n.d.). Retrieved January 07, 2017, from http://www.webmd.com/sleep-disorders/talking-in-your-sleep#1

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Stibel, J. M. (2013). *Breakpoint: Why the web will implode, search will be obsolete, and everything else you need to know about technology is in your brain*. Academic Press.

Suskind, D., Suskind, B., & Lewinter-Suskind, L. (2015). *Thirty million words: Building a child's brain: tune in, talk more, take turns*. Academic Press.

Swaab, D. F. (2014). *We are our brains: A neurobiography of the brain, from the womb, Talking in Your Sleep*. Academic Press.

Truss, L. (2004). *Eats, shoots & leaves: The zero tolerance approach to punctuation*. New York: Gotham Books.

Yagoda, B. (2006). *When you catch an adjective, kill it: The parts of speech for better and/or worse*. New York: Broadway Books.

KEY TERMS AND DEFINITIONS

Artificial Neural Networks: A data processing method in which a representation of the information processing of the brain is mimicked.

Natural Language Processing: A branch subject of artificial intelligence dealing with words as they are spoken and gathering linguistic qualities for a computerized system.

Neural-Gaming: A branch of augmented and simulated reality using a computerized platform meant to benefit the advancement of a human's brain.

Verbal Apraxia: A speech impediment caused by a motor disorder resulting from possible damage in the brain.

Verbal Dyspraxia: A speech impediment caused by a developmental delay.

Voice Recognition: A branch of artificial intelligence and natural language processing interpreting human speech into structured data usable by the computer.

This research was previously published in Ubiquitous Machine Learning and Its Applications edited by Pradeep Kumar and Arvind Tiwari; pages 195-216; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Section 6
Critical Issues and Challenges

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 72
Research on Human Cognition for Biologically Inspired Developments:
Human-Robot Interaction by Biomimetic AI

Marko Wehle
Technische Universität Berlin, Germany

Alexandra Weidemann
Technische Universität Berlin, Germany

Ivo Wilhelm Boblan
Technische Universität Berlin, Germany

ABSTRACT

Robotic developments are seen as a next level in technology with intelligent machines, which automate tedious tasks and serve our needs without complaints. But nevertheless, they have to be fair and smart enough to be intuitively of use and safe to handle. But how to implement this kind of intelligence, does it need feelings and emotions, should robots perceive the world as we do as a human role model, how far should the implementation of synthetic consciousness lead and actually, what is needed for consciousness in that context? Additionally in Human-Robot-Interaction research, science mainly makes use of the tool phenomenography, which is exclusively subjective, so how to make it qualify for Artificial Intelligence? These are the heading aspects of this chapter for conducting research in the field of social robotics and suggesting a conscious and cognitive model for smart and intuitive interacting robots, guided by biomimetics.

DOI: 10.4018/978-1-7998-0951-7.ch072

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

Humans adapted in the former eras of technological progress often to the use of technology, in spite developed by humans, one had to learn how to handle technology via many steps and realise them with caution. Because of technology's pioneering complexity it was not always satisfying the users needs and the premise of safe handling, special training and expert knowledge was often required to a large extend. In the new era control and feedback control systems, also called Cybernetics, take over for handling technology with ease and comfort. The focus on the practical usage of technology shifted in an evolutionary process towards human centred interfaces, e.g. from first computers with command line via mouse input towards touch-screens. For instance, considering people, and even kids, on using touch-screens one can recognise an intuitive usage, familiar and intuitive movements for interactions are easier to accomplish. In addition, sophisticated countries and states have to face the challenge of a demographic transition in their societies, which conditions the vision of service robotics to develop machines to be adaptive in their social behaviour. Meaning, reflecting the environment and being available to different users or user groups and their pleomorphism. A carer or patient interacting with the same device will have perhaps miscellaneous requirements and emphasis for it. Therefore, the research on biological cognitive abilities and confident behaviour, as well machine learning with information seeking aspects, should be considered for developing artificial intelligence (AI) systems. Thus far, it is not a new idea and the subject of mind, consciousness and models of thinking and self-awareness are discussed already since the 16th century, but often in rather philosophical or mathematical semantic ways, because the inside perspective is hard to grasp. Notably the beginning in understanding of mind led to impressive sociological impact, for example by the "Age of Enlightenment". Hence resulting in innovative thinking, which still influences current research and this research is a remaining challenge to science and engineering with impressive progress about the last six decades. Despite the different nature of former scientific contributions, all of the work has one major aspect in common: to achieve the goal of an implementation of artificial thought, mind and consciousness – only an interdisciplinary perspective can provide the necessities and insights from biological role models towards intuitive artificial intelligent systems. With the aim to provide input to the development of intuitive user interfaces and artificial perception, leading to artificial cognition, only joint competence from fields like biology (physiology, cognition), physics/mathematics (cognition, information processing and artificial intelligence), engineering (cybernetics, biomimetics, robotics), sociology (cooperation, synergy of human and technology, as well sociological impact and consequences, ethics), design (materials, aesthetics, appearance), philosophy and arts can identify and display what humans are capable of and what is requested for feasible, useful and intuitive interaction with technology. Research is not aiming only to understand the basic mechanisms of mind by modelling it, also for generating computer models for robot artificial intelligence. So, that machines can interact (Human-Robot-Interaction: HRI) according to the gold standard of intuitive interaction, which is the role model of Human-Human-Interaction (HHI). Thus, empowering the robot with consciousness to be willing to learn, for adaptivity and its gain for autonomy and persistence as a service tool, able to tackle unknown tasks, changes of plans and at least mimicking HHI in HRI.

Reciting that "Biomimetics combines biology and technology with the goal of solving technical problems through the abstraction, transfer, and application of knowledge gained in interdisciplinary co-operation from biological models" (Bionik, 2012), the human could be the quasi-role-model for the robot. This is correct insofar as it is assumed that biological structures have been developed in the course of 3.8 billion years of evolution and are optimised under the given conditions in mutual dependency

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

with the environment. Now synthetic consciousness can be a useful concept for implementing intuitive HRI, despite the term consciousness somehow evades precise definition and the debate about it should be left to others, it is a concept for combining different ideas of modelling AI in robotic systems to gel intuitive and adaptive behaviour in AI. The question that rises therefore beforehand is, how can a human being, due to its biological and cognitive abilities, interact intuitive and modest? Given the biological role-model – the human – and its failsafe operation as an outline for biomimetic developments, it is important to study the processes of subjective perception and their application, to exploit the comfort zone of people for optimal interactions.

Consequently, robot developments are driven by the similarity of how we perceive the world through our senses to solve all the different tasks we are able to accomplish – so robots should receive the same information with their sensors and process it via adaptive artificial intelligence. The sensors absorb signals and transmit them to the robots controller. Sensors are available for a variety of measurements of changes in the environment in a variety of forms. The sensors of the robot can perceive even a wider range than the physiological senses of nature, but robots are limited at seeing the natural environment by their processing. Because the "brain" of the robot would be faced with the complex task to reconstruct images from the jumble of lines, shadows, and flitting spots of colour that are recorded by cameras and giving the input a meaning. The tactile sense as another example is a vast challenge for developing robotic hands which can tackle a variety of tasks as human hands do. The options in technology are more comprehensive as in the physiology of a human, but millions of years of evolution and adaption lead to a filtering process, which would remain to the engineers and programmers for the developments. In addition to give robots the necessary autonomy for operating, planning tasks and mission accomplishment, they have to be aware and their artificial intelligence has to lead to an embodiment for the interest of managing resources and given priorities in tasks. Nevertheless, there is also a catalogue of interests in AI developments for intuitive Human-Robot-Interactions aiming for social robotics, in summary, setting up clear and humble standards for common Robot-Human communication and interaction aspects, which can help to guide future developments. The acceptance of social technology and its impact in society – to be secure, fair and in the favour of most, should be improved by following a biomimetic approach and some basic questions:

- How can a robot act accordingly and adapt to a human being to achieve a simple and convenient interaction?
- What are cognitive capabilities, skills and limits of humans, theoretically, to mimic HHI in HRI?
- How can anticipation, support and help for preventing frustration in HRI?
- What are technical options for sensor fusion, artificial intelligence and perception modelling?
- How can privacy and as well data security be assured?

In order to elaborate these questions, HRI is in research a flourishing field, because the challenge is to overcome certain barriers for intuitive interacting AI by setting humble standards and a baseline for training an AI machine, the rest is, in an ideal case, left to adaptivity. The reasoning behind this is, that programming an open world and complex social behaviour may take infinite time and may never converge due to changes, caused by mutual interaction with the environment. Especially the process of decision making in unknown terrain and in adaptive interaction situations is a progressive problem.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

BIOLOGICAL ASPECTS: ROLE MODEL FOR ARTIFICIAL INTELLIGENT HRI

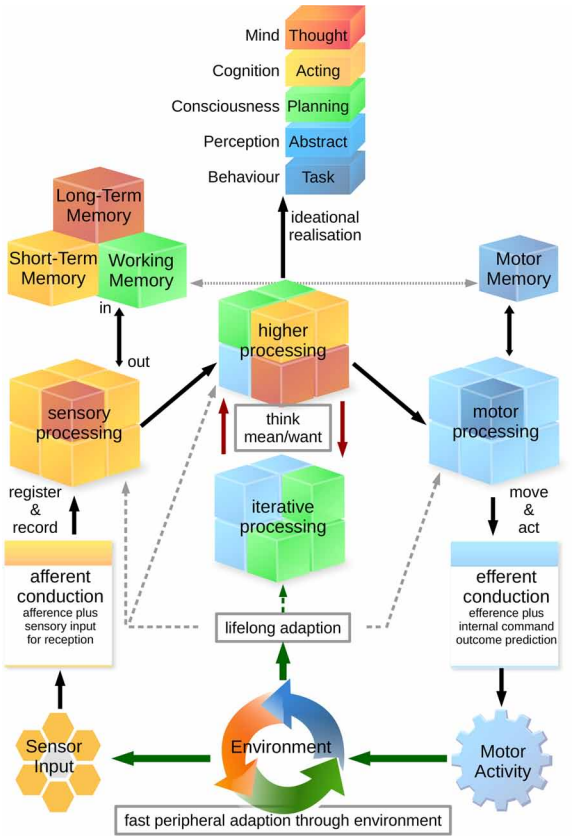
New technologies will to be developed in participative manner, eventually: through intensive HRI research by the advantages of video analyses, as well evaluation of user behaviour and comfort via social cues. Innovative non invasive measurements of physiological aspects, like monitoring skin conductance, gaze tracking and heart rate, blushing, speech analyses and analyses of questionnaires can help in order to establish a baseline. There are manifold biomimetic aspects for artificial intelligence to be considered and phenomenography is a powerful tool to classify most aspects, which will become goals in technology developments. As a starting point to layout an implementation, rather physiological aspects provide key aspects, for instance sensor-fusion and neural network dynamics lead for information integration, prerequisites for error correcting approaches, conscious information processing and weighting for output, in the end this progresses to decision making and interaction. But in the first instance, the window to our world is provided and in the same time limited by the physiology of our senses. In order to perceive and interact with our environment, we need constant feedback. Humans receive information in the order of magnitude of a 10^9 bit/s through our senses; only 10^1 – 10^2 bit/s are added and processed consciously. The remaining information will be processed subconsciously or filtered to be not used. Conversely, we generate, for example by speaking and motor information about 10^7 bit/s towards the environment. Figure 1 displays the interconnection between perceiving of and interacting with the environment for an embodied intelligence.

The neurobiological hypothesis behind these schematics in Figure 1 is the model for explanations of higher order mental activities. Assuming that humans perceive their environment and register it via multiple over the organism distributed sensory cells. The sensory processing within the cerebral cortex delivers information to the working memory and as well the long-term memory (LTM). The higher order processing and the motor processing can convert signals in fast movements, while the motor memory is delivering therefore already manifested excitation patterns towards the motor active neurons. Additionally, before higher level processing, contents can be iteratively processed in cycles, for the purpose of temporal extended thought processes. Motor active aiming and thinking corresponds here, with what is considered as will. Intensive, sufficiently long, vehement thinking results in ideational realisation or consciousness. Appropriately induced actions will be recognised as consciously triggered. The actual consciousness will stay without repercussion, in the sense of physical compatibility.

Generally we perceive our world with 5 senses, but for HRI especially the 3 senses for vision, acoustic and tactility are considered to be important, so to be focussed for developments. The stimuli of the environment affect the body in various forms of energy, like mechanical and electromagnetic waves mainly. In the sensory organs, there are specific receptors and each sensory cell has its adequate stimulus which causes for each stimulus a specific sensory impression. Within one sensory modality (vision, acoustic, tactility) the different qualities of a stimulus can be distinguished by quantities like intensity and frequency, e.g. sound or light. For instance, the optical sense provides alone and very fast about 80% of information of the surrounding, thus constitutes the most of our perception. This is accompanied by the acoustic and haptic perception. While in principle perception is indirect, meaning the stimulus information is passed to the appropriate areas of the brain via the thalamus as a hub, thus filtering the amount and distributing the information. The specific regions communicate with each other, they decode the previously encoded information from the sensory organs and discern the information. But of course it is the composition and conscious reconstruction of information that provides us an overall picture and most important gives rise to a meaning for information. Also a human being is confronted with a lot of

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Organisation scheme for AI in operational architectonics mimicking biological paragons



information that cannot be recorded and processed. Therefore, the selection of information passed is of utter importance and already begins in the sensory organs, e.g. target selection cells in vision, additionally filtering processes occur in further certain regions of the brain. Remarkably, the brain can shift individual perception fractions on the timeline, which in turn assembles perception performances and fills gaps in perception. For continuous, gapless perception the brain has to manipulate our perception of time in the performance. Perception is also a memorising process, by storing patterns of perception, and depends on our attention.

Perception leads to neurobiological consciousness and cognition, the localisation of consciousness is thought to be shown in large cortical areas of the left brain half, additionally support for the consciousness is given by input from the hindbrain, anterior to the cerebellum, seen as the stem of the brain and towards thalamic structures. A signal is retrieved as conscious input by signal propagation time of more then 0.5s and so to say is conversely a processing which is extensively carried out in the human brain, also called cognitive processes (Pfützner, 2014). Nevertheless, a healthy human is always seen as conscious while being not in a coma or deep sleep and therefore consciousness is a central aspect of our life and in interactions as well. Consciousness is fuelled with perceptions and memories giving meanings to our cognitive processing. As several definitions for consciousness arise, they all share the activeness of thought (Koch, 2013):

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Respective to our common sense, consciousness is equal to our inner mental life, consciousness appears if we wake up in the morning and is present till we fall asleep (not dreaming), it is present if we dream and disappears if we are in deep sleep, coma or being under general anaesthetic, death makes it disappear forever.
- On the basis of behaviour is consciousness a checklist of actions or psychology, which rates every organism as conscious that can achieve at least one of them (Glasgow Coma Scale).
- A neuronal definition describes minimal psychological mechanisms, which are necessary for conscious perception, an active and functioning cortico-thalamic (or thalamo-cortical) complex (neocortex and thalamus appended by hippocampus and amygdala, basal ganglia and claustrum).
- In philosophy, the general definition is that consciousness is how it is to be like something. How does an experience feel like if it is perceived by an organism and only this particular organism can know how it feels, if it has this experience (phenomenological consciousness).

In cognitive sciences, consciousness is also linked with information process theories, meaning that information exchange between regions of higher order in the visual cortex and the planning stages of the prefrontal cortex are perceived as conscious. Supported by the model of "global radiation", which states that data in a workspace is available to a number of underlying or subordinate processes (working memory, speech, planning module, etc.) and emanating information makes them conscious. Hereby is also worth noting the consideration that the workspace is rather limited and information is rivalling, so information can get expelled. However, conscious information is therefore accessible as a unity and globally for the whole system. Physiologically this is achieved by weeping pyramidal cells in the prefrontal cortex and from this follows the creation of the global workspace. Information theory is often used in this context as a mathematical formalism, which can quantify the causal interactions of each part of the neuronal system. It describes how much a status of each part influences the status of the other parts and how this influence develops with time. Also, it categorises and characterises therefore the interrelation between all components of the system, which is seen as a unity comprised by its components. Consciousness is not an accumulation of bits, it is not about culmination of data, it is rather the relation between those bits. The architecture of the neuronal system, its inner organisation, is crucial for consciousness. Another refinement of the model is the so called "double aspect theory" retrieved from philosophical monist theories, but still considers a distinction between mental and physical aspects. Which concludes that information of two originated states with inherent and elementary attributes exists divided into extrinsic and intrinsic states, whilst the number of discrete states positively correlates with the repertoire of conscious experiences/apperceptions. Centre point is as well the integration of information, each apperception is especially informative and specially differentiated and conscious states share the attribute of integration of information. Each conscious state is a monad, but it cannot be disassembled into its single components anymore, which could be independently experienced (Tononi, 2004). The quantity of conscious experiences, which is created from any physical system in a certain state, is equal to the amount of integrated information of these states, plus the information generated by its sub-entities. As integrative step, the system has to discriminate between the states within the repertoire of states, acting as a unit – the system cannot split the information into a collection of causal independent parts anymore. For instance, in a split-brain patient live two independent conscious units, each with information independent from the other hemisphere – the integration of information is for the brain as a whole unit unachievable.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Consciousness is assumed to be equal to the content of short-term memory (STM) and long-term memory (LTM) (Thompson, 2010):

- Consciousness plays a key role as necessity for cognition and cogitation.
- Conscious experiences appear as apperception – if attention is paid to the senses.
- Self-consciousness includes the moment/situation.
- Consciousness spreads with perception over time (past) for aspects of experiences and knowledge.
- Memory of knowledge, experience and skills in the long-term memory can be retrieved from the subconscious mind.

Recalling consciousness as a process which is applying numerous feedback loops regarding different parameters (e.g. temperature, space, time and in relation to each other) to create a model of an environment, mainly to attain aims (e.g. procreation, survival, food, etc). Biologically there is a classification of consciousness in its peculiarity (Kaku, 2014):

1. **Stream of Consciousness:**
    a. Mobile organisms with a central nervous system have a so called stage-1-consciousness, because they require a set of parameters to register their changing environment (e.g. reptiles, they possess hundreds or more feed-back loops which regulate their senses of smell, balance, tactility, hearing and vision, blood pressure etc., with individual feed back loops in the specific brain areas as well). The human brain roams sensory information through the brain stem (Truncus encephali), passes the thalamus with pre-processing and filtering, reaching eventually several areas of the cortex for further processing and will be passed to the prefrontal cortex. This processing results in our physical position in space.
2. **Sociological Rooting:**
    a. Organisms which not only possess a model of their location, also a relation to other organisms (social beings with emotions), have an exponential increase of feedback loops with a correlating complexity of behaviour. Therefore, creating a model of a position in a "society", which is depending on the establishment of brain structures in the limbic system, whereas the limbic system consists of the hippocampus (long-term memory), amygdala (centre for emotions) and the thalamus (relay station/hub for sensory input). This system issues new parameters to refine models for the relationship to others. In the human brain, emotions will be created in the limbic system and processed there, in the framework of the stage-2-consciousness the brain is confronted with a constant stream of information from sensory input, but emotion are the fast reaction of the limbic system on a situation, for instance the fastest in emergency situations, which even neglect an authorisation from the prefrontal cortex. The hippocampus is crucial for the contribution and the digestion of memories. Mirror neurons will therefore also play a role for emulation and for empathy and raise the ability not only to mimic complex processes, also to empathise with others.
3. **Simulation of Future/Prediction of Outcome:**
    a. If an organism reaches stage-3-consciousness, the number of feed-back loops is tremendous, so that no supervision is needed for filtering incoming data, in order to simulate future outcomes of situations and actions, so that predictions can be made in order to make decisions. With this model of the world, simulating future, causalities are connected and form a tree of

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

causalities (or a decision tree). In the human brain, the key aspect of stage-3-consciousness is formed in the dorsolateral prefrontal cortex, which is assumed as the most recently evolved part of the brain, it conveys between the orbitofrontal cortex (decision centre) and nucleus accumbens (centre of desire) and serves as a working space of mind.

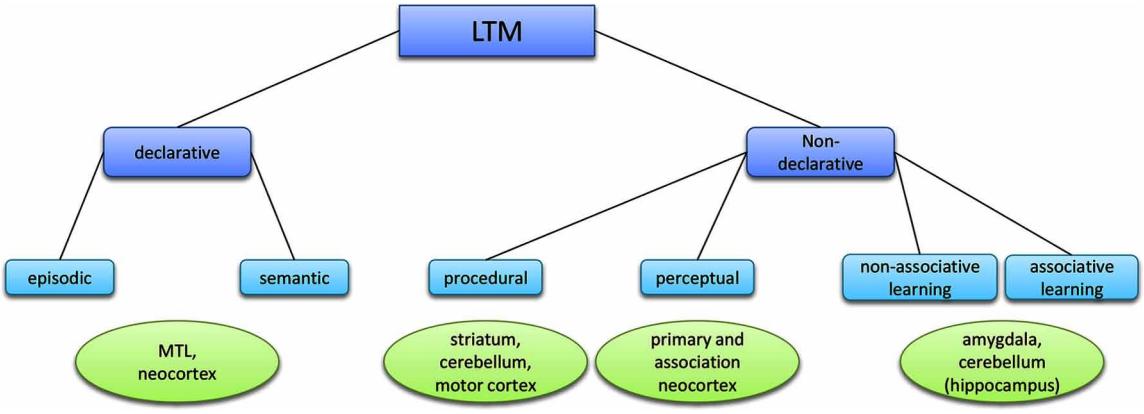
Deducing that in order to feel experiences; consciousness and cognition, the presence of information in short-term memory as a workspace and stored information from long-term memory is needed. Because the short-term memory is limited in capacity and memory, the persistence ranges from a few minutes towards 24 hours, new content is extruding older information. The long-term memory is described as a selection of information, which are filtered and consolidated from short-term memory by repetition and learning. Whilst all information, once attained in the long-term memory, are permanent and pattern regular memory trails, so called engrams, which facilitate remembering contents by biochemical connections/processes. Figure 2 summarises how the information storage is classified in the human brain in terms of memory aspects.

The Papez-circuit, see Figure 3, is thought to play a major role for information exchange between short-term memory and long-term memory, since experiments and observations have shown that damage of some of the involved parts can result in memory and language impairment. Repeated passing of the circuit can lead to information transfer from short-term memory to long-term memory, while this is also thought to be a cause of dreaming for long-term memory consolidation.

How the memory consolidation can be depicted in a schematic way is shown in Figure 4. Additionally, the instance of the iconic memory is a mainly visual sensory memory and part of the short-term memory, it is thought to be a domain which quickly records apperceptions from senses, e.g. a face or a scene, and is able to transfer the information more or less directly to the long-term memory.

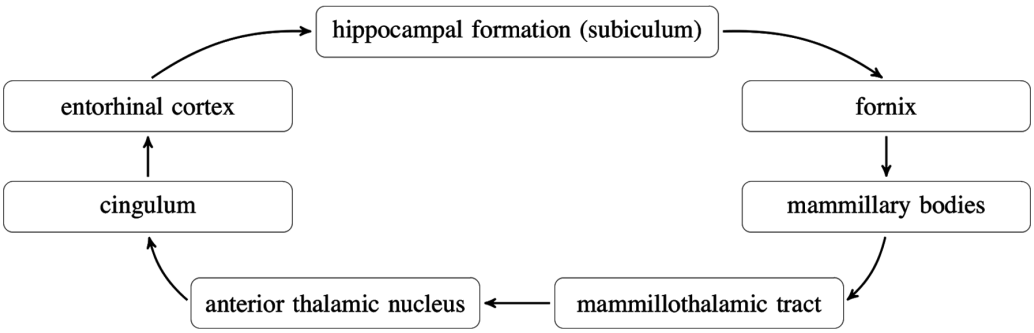
This organisational scheme from Figure 2 depicts the idealisation for interconnections between networks forming network complexes, additionally it shows an idealised hierarchical assumption for the long-term memory (LTM) as a top level memory entity, which is a naturally inspired assumption for supporting a global workspace theory and instantaneous memory access for all processing brain structures. Cross-connections are not drawn but biologically evident. (MTL: medial temporal lobe).

Figure 2. Hypothetical memory organisation scheme in a human brain, different types and aspects of memory (in blue) are connected to their associated brain structures (in green)



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 3. The Papez circuit with information flow in circular manner between the brain entities involved



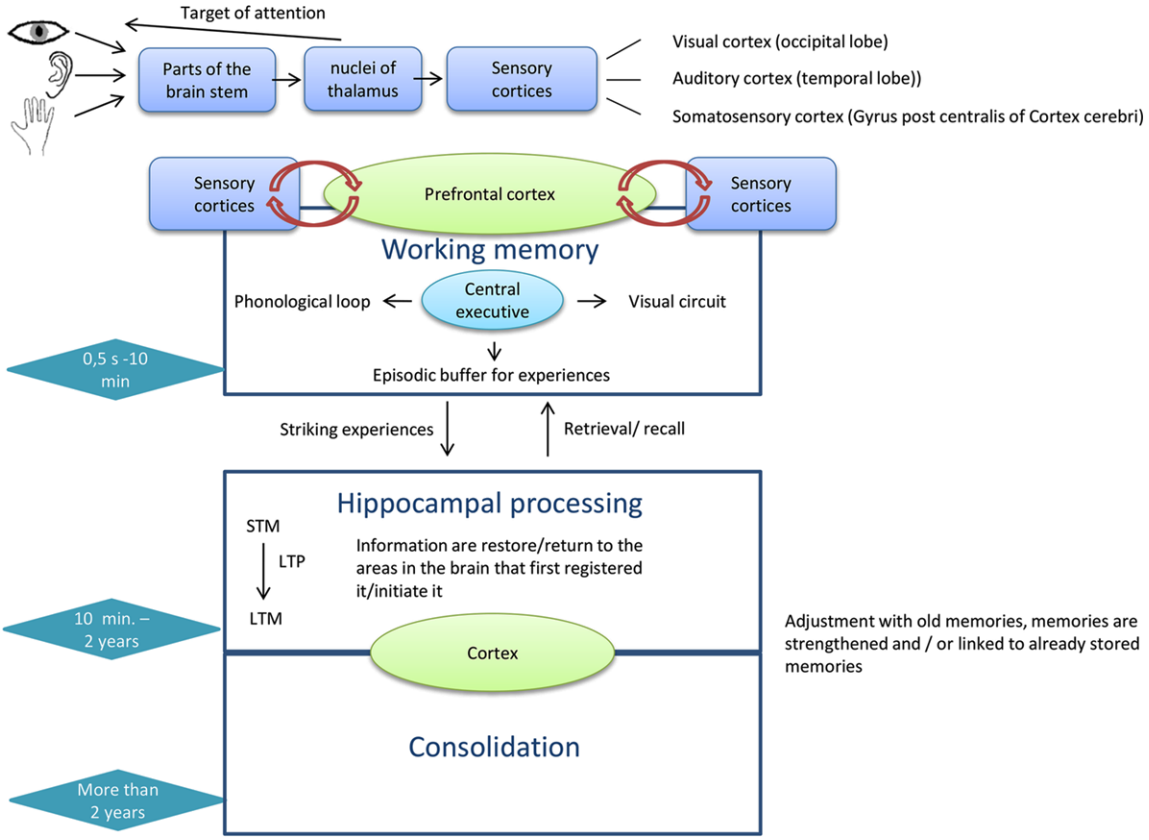
Some fragments of information still stay in the short-term memory and are available for the working memory. Are these information new or respectively motor skills and they are not practised, only a part is transferred to the long-term memory and most of it will be lost.

The schematic memory layout from in Figure 4 shos sensory information, which also includes procedural knowledge (like learned motor activity), is taken up, filtered and moved through parts of the brain stem and kept temporarily for short time with full detail in the working memory (seen as sensory or iconic memory), for processing the information. Further processing and association of a part of these information reaches the short term-memory (STM), for instance by paying attention to it, which gives it importance for the brain after a certain timespan forming spatiotemporal patterns and when these patterns are seen as relevant, the hippocampus repeats and loops those patterns in the consolidation process to combine these information containing patterns with associated memories for long-term storing. (LTP: long-term potentiation).

Long-term memories are conserved in a formation of groups of neurons, which fire in the same pattern as the original experience was represented. "Complete memories" are divided in components (emotions, sentiments, thoughts, etc.) and in each case distributed to be stored in the competent or relevant areas of the brain. The simultaneous firing or activation of these groups of neurons represents these memories at large (Carter et al., 2010). Attention ($t \sim 0.2s$) arranges neurons, which register the event to fire more frequent, this increases the intensity of experiences and in addition increases the probability to memorise this event. The thalamus regulates therefore the brain activity regarding the aim of attention, whereas the frontal lobe filters information to increase the attention span by averting possible diversions. Sentiments or feelings ($t \sim 0.2–0.5s$) are perceived through combination in associative areas. This especially happens within the memory consolidation process in deep sleep (delta-sleep), while information gathered throughout a day are absorbed, assigned and partially presorted, during the sleep phase it will be finally sorted and partially deleted, otherwise consolidated. Storing learned contents and consolidation of knowledge happens during these deep sleep phases, the brain combines subconscious and parenthetically recorded information to new knowledge, at the same time learning is conducted, the brain also subconsciously solves problems and combines knowledge to ideas. In the wake phase recorded information will be stored temporarily and is so recalled for long-term memory consolidation. Facts, which are connoted with emotions, remain easier in the memorising process and are known to last longer, because their activation patterns are much stronger. Presorting and filtering of importance of information is thought to be carried out through a process called "tagging", this process can be influenced and emphasised by emotions (Born & Wilhelm, 2012). Important information is tagged through gathering and perception,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Schematic layout for memory consolidation processes



while standing in a context, these tags support memory consolidation during sleep phases. The paradigm is also seen to be valid for premeditatedly recorded information (motivation aspects), which are correspondingly tagged. Since the brains memory consists out of many different systems, the central hub hippocampus processes information, for instance information is filtered, valued and compared and passed on to the long-term memory in the cerebral cortex. Different kind of memories can be stored in different systems, for instance is in the cerebellum the procedural memory, perceptual memory in the cerebral cortex and factual knowledge is located in the left hemisphere of the brain.

The human brain is subject of massive stimulation in every second being conscious and it can not only rely on filter processes in “bottom-up” approach, this information processing also has to be carried out “top-down”, in order to select relevant information from the massive amount of stimuli and experiences. Events can be even predicted from experience, touching, scent, taste, the brain constantly and continuously processes information and makes predictions from this. Consequently, if the predictions become reality, the brain can discard the information processing and dismiss the information, because the implicit knowledge about the event is already present. If the event causes a “better than” signal in perception, learning will be triggered and only this can cause an organism to adapt its behaviour over time for optimisation. Learning is by this means not everything what is perceived, rather what is perceived with positive repercussions (Spitzer, 2002).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Individual perception of self is also a result of a distinction between objectivity and subjectivity. Objective results require review by measurements without human influences like emotions or feelings, whereas subjective results are influenced by involvement and estimates like taste, memories and associations. Objective reality engages with life, how humans perceive reality and with the construal of coalesced personal experience. While subjective perception is different for each individual (senses work different, matching in standing by memories is different etc.), but still comparable, because individuals have comparable apperceptions and use learned expressions from extrinsic input. Consequence is, that an objective reality which is perceived remains objective, apart from the individual independent character, because humans interact with objective reality and explore it as a subjective extract, while this is than subjective reality.

SILICON DEDICATED TO CONSCIOUS HRI

Now, rather than to simulate a brain in order to conduct research on its processing, the aim is to emulate attributes of higher order processing for an implementation in an AI structure. Therefore, not the entirety of physiological and biochemical attributes has to be re-created, but individual aspects of those can be implemented for benefit, e.g. for a reward system for learning or achieving an aim, or even empathy in HRI. Before starting to model any ideas about AI architectures, the division of tasks concerning the aspects of filtering information and processing information on different levels has to be tackled. The term consciousness is here a vehicle to operate with subjective entities to address questions about information processing.

In cognitive sciences and philosophy, the ability to abstract from information and changing the level of representation of information is an alternative way to distinguish between unconscious and conscious states. From hereby inspired modelling attempts are preferably symbolic representations of artificial intelligence and therefore efforts to refer to knowledge representations which are suitable to express human expertise. Whereas in neuronal networks these representations are covered within patterns in the network activity, here the knowledge is depicted in data structures following rules of causality and association, which is emphasising multiple levels of representation in perceptual information processing.

In terms of a better scientific understanding, the questions what is the nature of consciousness and if a machine can be conscious will remain idealistic to be answered and may in present research not even converge to unfeigned machine consciousness. But imitative consciousness and essentially approaching machine autonomy by awareness can be tackled; despite the lacking definition of the terms (Rosenthal, 2009). Therefore, neurobiological findings will serve as a layout and biomimetic inspiration for technological developments. Neurophysiology provides the insights of neural architectures and the operational partitioning thereof, these insights are often obtained by electric field and imaging experiments, which obey an analysis in contrastive manner. The brain activity is monitored as patterns during events in the experiment, a distinction between conscious and unconscious states is possible, for instance given under laboratory conditions making a test person aware and conscious about a briefly seen written word or not, obtaining the outcome that patterns of conscious brain activity are monitored only for subjects perceiving the written input, called the “attentional blink” (Dehaene & Naccache, 2001).

Consciousness may be the product of carefully balanced chaos.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

This statement is displayed in the context of a biological study to reveal the importance of "cortical integration" in maintaining consciousness (Tagliazucchi et al., 2015). This means, in the current context of AI modelling, the processing and combination of multiple inputs from different "senses" at once. Whereas the fact if consciousness by itself is subjective or objective is rather philosophical, the division between conscious and unconscious processing can be made, and is even measurable (Tagliazucchi et al., 2015). Most of the experiences for humans are created by the complex intertwining of several sensations like vision, smell, touch and the reminiscence of previous experiences with the same object, for instance of a fruit. For this, the brain is merging all the inputs into a subjective experience at once, whilst consciousness is creating a meaning for the sensory patterns conceived by exploring the spatial patterns in the brain. It has been concluded that, in the brain, a well-balanced and optimised connectivity between neurons creates a large number of possible pathways and within this network of neurons consciousness might appear by exploring the network meticulously. These experimental results agree with the theoretical lead of information process theories, that a physical basis for awareness is achieved in a complex active system with conscious memory, which can be seen only as an integrated and unified system of specialised compartments for processing. In this respect division of labour, distribution of tasks and parallel processing is implemented and complex network dynamics provide an overview for pattern formation, to preserve integrity of procedural information processing as well representation of declarative knowledge. But also like suggested in Information Integration Theory (Tononi, 2004; Koch, 2009), the largest number of pattern forming arrangements is not necessarily implying the maximum connectivity between neurons. Because the signals need to be distinguishable and the cortical circuits also flexible, the arrangement of the patterns is what is thought of as driving consciousness.

Despite facing a line of philosophical implications within consciousness research models and each one driven by different objectives, a categorisation of model systems can be undertaken by some central attributes of concepts. In summary there is a distinction of models with global workspace, or integrated information, or internal model of self, or higher order representation within the abstraction of represented contents, or attention seeking and exploring mechanisms, these fundamental criteria can help to classify different models, as long as the particular criteria is central for the modelling approach. Of course, the listed features can overlap and applying criteria like qualitative-ness, subjectivity, sense and notion of self, unity, etc. and behaviour measurements on humans like the Glasgow-coma-scale, do not necessarily contribute to overcome the problems assigned to machine consciousness (Reggia, 2013). Several attempts to model synthetic consciousness are carried out and they were carefully realised in the context of information processing aspects of functional consciousness, not as any kind of subjective experience. Because the emulated aspects of consciousness within its assumed neuronal correlates are modelled by the help of computational methods, like it is done in many fields, for example enzyme-ligand binding or even activity and reaction transition states and nobody would expect this reaction even to happen in a simulation, as we would not put the substrate into the computer and expect the product of the reaction as physical entity coming out. The expectations are indeed not so high that machines develop phenomenological consciousness to be present, but simulations are evaluated in which sense they can reflect experimental correlates and in particular how they can help to improve artificial intelligent systems in functionality.

Substantial for a consideration in modelling a neurobiological layout are a few ideas, for instance is the biological role model here seen by the raise of global distributed neural activity in the brain via global patterns. Where the brain is seen as organised in a global network of specialised processing units organised in subnetworks like sensory input processing, motion control, learning and mainly working

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

on an unconscious level, but giving raise to consciousness as a unit by the means of a global information processing. This high level of interconnection and cross linking enables spatiotemporal patterns from modules to be spread globally and accessible for any mental state, if conscious (Baars, 2002; Baars et al., 2003; Massimini, 2005). The ideas for those theories are vastly inspirited by the "global workspace theory" of Bernard Baars (Baars, 2002, 2005). Several modelling attempts have been carried out in computational manner, for example with "IDA" (Intelligent Distributed Agent), which is an assembly of Java "codelets", a later version even is able to learn ("LIDA") and given a self-model too (Franklin & Patterson Jr, 2006; Snaider et al., 2011). Rather than a neuronal network, these program fragments are specialised on their tasks and executed in parallel as modules, but communicate via a global workspace. Other models show a stricter structure of neuronal networks, even incorporating biological relevant features like models of spiking neurons and explicit oscillatory behaviour. For instance, the "stoop" task model has been showing to simulate the "attentional blink" (Dehaene et al., 2006; Raffone & Pantani, 2010). Also robotic implementations have been tackled for a global workspace model to control a simulated robot with approximations of the concepts of consciousness, imagination, and emotion (Shanahan, 2006; Shanahan & Baars, 2005). While some of the models are made for simulating consciousness, they all show the significance of a global workspace approach, inspired by and taken from biology, to create instantaneous consciousness based on experience and empirical knowledge. The global workspace processing is collocating the subconscious processing done in specialised parts of the brain where local processes are carried out. To premise up on the global workspace model and additionally getting an idea what the global workspace is, one might assume that consciousness is correlated to the utilisation of information integrated in a system. For instance, realised in the concept of a neurocontroller for simulated robot vision (in order to make it "attracted to positive stimuli"), information integration based predictions about the conscious states of a spiking neural network (Gamez, 2010) were achieved. The evaluation of this trial was carried out by a measure of φ. Information integration theory (IIT), in contrast to others, already enables a quantification or measure for consciousness of a system with its number φ, which can be useful as a concept. Unfortunately, to compute this aspect is prohibitively computational expensive, so that for the experimental implementation only an approximate and more efficient algorithm could be applied (Gamez, 2010; Tononi, 2004, 2008; Tononi & Koch, 2014).

Further progress has been made for a first person perspective with self-awareness, in the frame of an internal self-model. As a necessity of conscious mind, the basis for the first person perspective is a neuronal representation of body properties in the brain, even if somewhat scaled and referred to as cortical homunculus which is a physical representation of the human body and this neural mapping can be even distinguished between a sensory and a motor representation. This aims the so called "body image": a modelled idea of the sensory-motor cortex (Bongard et al., 2006). A self-model in artificial intelligence, or even one step further as self-awareness, is more than helpful for a machine to support its way towards a first person perspective. So the aspect, which is central for a biomimetic approach, is the embodiment and the robots awareness of its extremities, to learn motions and to adopt within executing physical tasks. Moving away from a central controlling system upon which all sensors, actuators and limbs depend, provides the opportunities to outsource computation to the body and the environment in order to help to abstract information for higher integrity and to take advantage of ontologies within AI systems. This so called *morphological computation* is a notion of soft robotics and uses the capabilities of materials to effectively carry out computations and sensorial tasks through their properties of elasticity, viscosity and compliance (e.g. electric sense, magnetic fields, shape change). In physical robots this can be achieved by sensory data processing and morphological computing in the limbs, which can be

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

used to classify the input towards concepts of data and the concepts will be passed on to a higher level of processing in symbolic representations of causality (Chella et al., 2008; Chella & Macaluso, 2009). Additionally, soft robotics actuation and persistent autonomy can pay off together in energy harvest, simplified control, safer HRI and extending the robots endurance (Lane, 2012). While in addition, the use of special distributed neural network architectures helps to minimise computational effort and to ensure maximum efficiency. The virtual representation of a machines physique as approach for embodiment is ingeniously for autonomy and mobility of robots nevertheless and therefore a must have in an artificial intelligent robot.

Other vital functions, for instance plasticity in recovery states of a machine for memory consolidation can be even essential rather than plainly advantageous. A necessity for this idea is the phenomenological abstraction of perception and a hierarchy of operations for spatiotemporal pattern representation, which will also serve the purpose of instantaneous information and memory access. This is quite significant because the implementation of artificial intelligence, and therein contained aspects of perception and memory processing of information should lead to a synthetic knowledge that is useful for the perception of the environment and, accordingly, for integration into it via embodiment. In this respect an implementation does not mean seemingly autonomy by image processing, classification and conversion of observations into coordinates, the aim is to achieve interaction and communication with the user. In addition the technology should be participative to achieve a maximal acceptance, even if a convergence towards 100% is unrealistic. But exploration, guessability and willingness to learn should be integrated into the system, so that the robot adaptively follows the users behaviour. By stages, the exploration and attention of an artificially intelligent system can also be driven intrinsically by conscious awareness to fraction input streams of visual and acoustic data, while filtering of information is mainly thought to happen in subconscious areas of the brain. So, consciousness and attention are not synonym, but definitely linked and highly correlated. A valuable side effect is that, attention mechanisms are useful for exploration and learning in machine behaviour, driven by information gain to base decision making on machine experience, whilst learning can be reinforced due to positive reward, like modelling neurotransmitter distribution etc (Distefano et al., 2015).

In the often studied role model for human cognitive developments, children development, it is observed how an organism can learn to discover its own physical role in the environment, interact with objects and establish social behaviour. These studies show the basic mechanisms of organisation and progressive learning by increasing complexity of processes and interaction, while the connection between brain and physical body is networking with respect to the physical and social environment. Aiming the reproduction of such role models in robotic developments is helping to understand the basic mechanism not only further, also it raises the opportunity to mimic the progressive learning curve for social interaction in application for HRI research (Nguyen & Oudeyer, 2013; Oudeyer, 2013).

The perpetual dialogue between social studies and engineering can build up research to robots which are learning by watching and exploring through curiosity and can make use of their developmental trajectories to self-organise and discover the physical world by their physical attributes, estimate the functionality of an interactive object by visual recognition and as well the ability to communicate. This technology foresight ideally enables AI structures to be primed by limited training for social common sense on an established baseline dataset and afterwards to structure its own learning experiences, creating its own sets of skills, for instance manipulation and handling of objects, by specialised application and depending directly on its environment towards increasing complexity of its sensorimotor capabilities.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Modular Artificial Intelligence on a Pseudo-Biological Layout

In order to link biological input and developments of robotic AI technology in a bottom-up approach, following the course of

Kinematics → Dynamics → Phenomenology (Taketani, 1971)

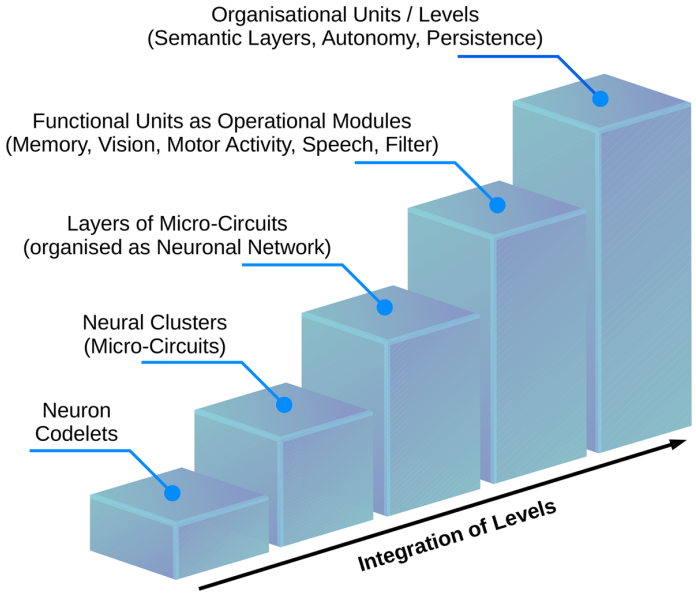
This course can help to avoid first hand implications by aiming for synthetic consciousness. In the current context this hierarchy of physical principles is motivating to model a structure of causal connected small entities, which will form an AI connected to the physical world in the end. In order to follow the lead of biological role models and applying biomimetic layouts, but not mimicking low-level neurophysiology, the modelling approach is summarised by the guidance of the Operational Architectonics Theory (OA) (Fingelkurts et al., 2012), and it describes the higher level features of machine consciousness in a hierarchical operational framework. In contrast (Fingelkurts et al., 2012), the term operational modules includes in this chapters context, additionally to spatially and temporal organised patterns of brain activity, the organisation of higher level features abstraction of information processing and perception. To start from the bottom, again a fundamental biological fact supports the idea: the brain is divided into functional units (modules) which each developed evolutionary a specialised functional architecture, but it is always to be considered as a "whole" for the preservation of "integrity of procedural information processing". Thus, the revised idea of OA can be extended with modularity in AI aspects by implementing functional units for task distribution and parallel information processing capabilities to aim for autonomy and persistence in robotics. Necessary phenomenological abstraction and a hierarchy of operations are thought to represent spatiotemporal patterns, which ensure the immediate exploration of information for their instantaneous access. This does not only offer new functionality in the processing of sensory data, it is in many ways connected with feedback about changes in the environment and even the impact of actions by the robot itself. In principle, this allows also for predictions of an outcome of an action through internal simulation of action sequences. The basic idea is illustrated in Figure 5 as integrated hierarchy of operational units.

Nevertheless, a neurological or anatomically inspired path still has to be trodden, in order to embed all phenomenological contents in a unified three-dimensional spatial coordinate system. A possible implementation aims consequently for flexibility by using different types of neurons in the network formation, e.g. an opportunity is provided in codelets (Franklin & Patterson Jr, 2006; Snaider et al., 2011), which can form an artificial neural network (ANN) by a bottom-up hierarchy via small recurrent networks organised in microcircuits over local networks of microcircuits (Local Area Networks (LAN)) as intra-cortical layout, which organise functional units and level up towards corticocortical connections as interconnected LAN's forming organisational units.

Specifically, for an implementation, one can consider the small recurrent neuronal circuits inspired by the neocortex, the so called microcircuits, see Figure 6. This layout is suggested for an implementation, since in the neocortex the encoding for nerve impulses and its circuitry is one of the key aspects for the brains processing capabilities. Basis for the biological computing is the recurrent connectivity and the major aspect of inhibition in the connections, which leads to the formation of dynamical spatial networks already in the smallest scale. Biologically not all the mechanisms in inhibition are well understood, additionally the classification of inhibitory cells follows a non-standard terminology and this is complicating the situation also for modelling these aspects. Because there is no clear distinction between the differ-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Organisation scheme for implementing an AI following the idea of an operational architectonics hierarchy by mimicking biological paragons



ent types of inhibitory cells and their connective specificity on subregions of their target neurons, most modelling attempts usually use only general inhibitory neurons for their tasks. But what is known can be helpful and used for modelling inhibition in cortical circuits. Of all cerebral cortical neurons about 15% are inhibitory and there exist about a dozen morphological distinct types (Douglas & Martin, 2004). It is inviting to assume the existence of specific spatial grids formed of inhibitory neurons, embedded in the complex lattice of excitatory neurons. But the actual distribution is mainly organised in boutons[1] of inhibitory neurons and is rather dense and local. The lateral axon extension is usually not more than 0.5mm, in comparison the excitatory neurons form less dense local bouton clusters with more exceeding lateral projections (Binzegger et al., 2007; Douglas & Martin, 2009). Observations for different types of inhibitory neurons forming specific synapses with distinct parts of targeted neurons promote the idea of functional specificity, but unfortunately here biological access and evidence for the mechanisms is rare and simulations show no biophysical effects either, but have shown that inhibition is not acting as a switch, rather that effectiveness of inhibition is graded depending by the firing rate and the arriving excitatory current at the soma of pyramidal cells for example. Taken together, this view supports that some specific inhibitory cells control the output while others rather control the input to neurons (Douglas & Martin, 1990). To summarise the main functions of inhibition by neurons, they moderate excitation of course and they as well inhibit inhibitory cells, in principle to de-allocate circuits from inhibition. A further explanatory thought for the roles of inhibition in cortical circuits, and for instance the generation of non-linearity in linear excitatory projections, is the normalisation by division of the linear firing rate of each neurone by growing input of an aggregation of cortical cells. The dividing process is proposed as a mechanism of “shunting inhibition”, which decreases the resistance of target cells proportional to global network activity. Nevertheless, the so called division is indeed rather a subtractive mechanism by changing of firing rates. The excitatory and inhibitory systems together form therefore the algebra of the nervous systems and enable processing by gain control (Douglas & Martin, 2009).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

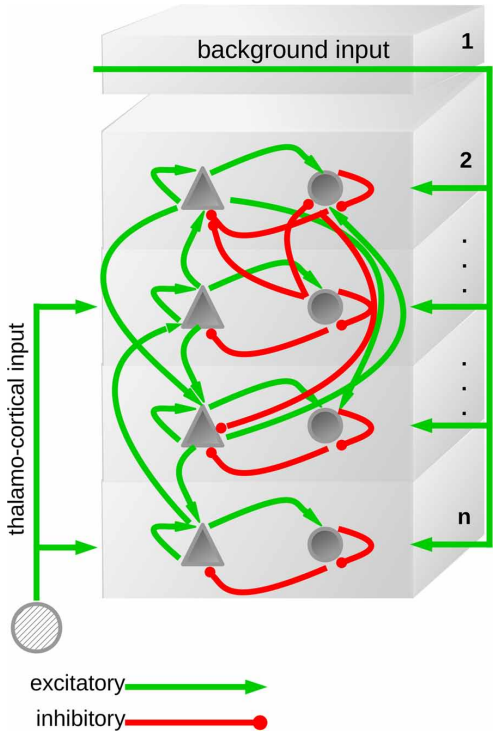
From the observations it seems clear that the reflex effect of concurrent stimulation of excitatory afferent nerve with inhibitory afferent nerve [...] is an algebraic summation of the effects obtainable from the two nerves singly [...]. The net change which results there when the two areas are stimulated concurrently is an algebraic sum of the plus and minus effects producible separately by stimulating singly the two antagonistic nerves. (Sherrington, 1908)

Emphasised is here the spike threshold importance and the fact of neurons being below threshold not contributing to the current circuitry. Inhibitory cells can keep the membrane potential under control and therefore actively shape the dynamical configuration of the circuits.

Considering inhibition as an important flexibility factor in the model, one can assume a dynamic network without explicitly changing the connectivity, but one step further, more flexibility and particularly a high depth of a network can be provided by recurrence, which aims for the so called deep learning. But the biological background is also provided from observations in the cortex. Estimates assume that the human cortex contains ~ 10^9 neurons and the architecture is laminar, while the functional diversity of the neurons can be broken down to about 100 different types in two major classes: excitatory and inhibitory neurons.

The model in Figure 6 is derived from cortical layers, represented via populations of exciting (triangles and green connections) and inhibiting neurons (circles and red connections). All Layers depicted here receive potential input from other microcircuits and external modules. The input can be seen as thalamo-cortical projections to corresponding layers, plus output in corticothalamic projections as recursion loops.

Figure 6. The definition of a microcircuit model



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Connections between those neurons are estimated from animal physiological data, mainly visual cortex of cat and monkey, as well from mouse and rat somatic cortex. Riveting circuit properties were revealed with this anatomical data, concerning the physiological connection of neurons in cortex' laminar and inter-laminar structures and as a consequence thereof, providing intriguing insights about fundamental properties of cortical processing. Important as an aspect for modelling is the dominance of local cortical synapses over those afferent to other cortical areas, which can be understood as clustering neurons to local networks and those connect to each other as bundles. The vast majority of cortical excitatory synapses and also most of the inhibitory synapses originate from neurons within local cortical areas. Concluding in afferent projections of cortex from and to thalamus and other individual cortical areas comprise a rather small percentage ($5 - 10\%$) of all excitatory synapses in the target area. In the same way the local neocortical areas are interconnect sparsely as well. Nevertheless, these afferents are sufficient to drive the excitation of the cortex (Douglas & Martin, 2004, 2009). Figure 6 captures these mentioned aspects in forming a microcircuit out of neuron codelets in a model idea. For enabling recurrence already in these small scale model clusters, one can assume that same types of neurons prefer connections with similar types and locally. In fact, superficial layers of the cortex are mainly connected excitatory by pyramidal cells with extensive arborisations about 70% of its own type. Thus, first order recurrent connections between cortical cells are likely between own types, in which a target neuron projects back to its source neuron in a tight positive feedback loop. Meaning, that cortical computation is carried out by positive feedback of relatively small afferent signals, so appearing as selection and re-combination of signals. An additional control parameter is given by spiking, individual neurons operating on small rate modes, because generally lots of input spikes are needed to produce a spike output event of a neuron. Positive feedback potentiation may be adverse for the balance, but the amplification of positive feedback enhanced neuronal activity is stabilised if the sum of excitatory positive current is less than the total of dissipated negative current through membrane leak conductances, action potential and synaptic conductances. This states, that positive feedback is a way to amplify small input signals to stand out of the noise in the cortex. This signal-noise-balancing is crucial for the network circuits, so that magnitudes of causally related signals are always controlled. However, the distinction between noise (spontaneous activity) and small signal input is not clear yet and can't be easily covered in artificial neuronal networks either (Douglas & Martin, 2004). For a better understanding of cortical circuits in research the focus was on models with linear threshold neurons, which comprise continuous positive output. This output is directly proportional to the positive difference between excitation and inhibition those neurons receive, if the difference is not positive or zero they do not contribute to the output. ANNs modelled with these linear threshold neurons (from excitatory and a smaller number of inhibitory types), interestingly, produce non-linear behaviour for the output, despite the amplification of the input is linear. The result of this behaviour is important, because it rises the capability to actively impose interpretations on incomplete or noisy input by restoring basal activity distributions from the connectivity of excitatory pathways in the network. For instance, the weight matrix of synaptic interactions describes the circuitry between various neurons in a network. If a neuron is now inactive, it contributes to the network in a temporary degenerative way and is so decoupled from the circuit, so to say not expressing its interactions. The full weight matrix can be now depicted as a reduced matrix, called the "effective weight matrix", whilst "silent" neurons are zeroed. The changes in network activity due to rise and fall of neurons are directly reflected in the weight matrix, which becomes a timed dynamical entity, the so called "permitted set" of neurons. Specifically, for local recurrent circuitry, the control is illustrating that patterns of thalamic input can be selected and input is mapped therefore also topographically across cortical areas, effectively by recurrent connections

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

being mainly local. These variations in magnitudes of a response due to changes in attention to a cortical stimulus, or simply neuromodulation, can be viewed as changes controlled by external attentional input. The mediation is simplified as an extension of the "permitted set" concept, which expresses the recurrent activity not monosynaptically between local clustering of connected neurons representing members of the cortical maps, single neuronal connections and activity would have to defeat failure as well (versus flexibility). In conclusion, the neuronal cortical computing or processing and organisation is done by local circuits or subcortical nuclei with a proportionate distribution of neuronal interconnections, surely evolutionary driven by limits in connectivity and energy consumption, as well robustness and reliability.

On the assumption that artificial neuronal networks can be seen as spatial memory, e.g. Hopfield networks or bidirectional associative memory networks or self-organising maps, they display stable patterns of activation and can be considered as solution to constraint satisfaction problems approximating complex functions, while reducing the parameter space efficiently. The major restriction or limitation in this case lies in the division between training and reproduction phase, so to state the rigidity of the network organisation after the training period. In order to achieve the best results for flexibility and to implement a type of "intuition" in an ANN, learning and training has to become a process of plasticity. Information seeking and learning through exploration, and driven by curiosity, novelty, surprise and positive reinforcement (plus "emotional tagging"), ensures a steeper learning curve and provides success rating (Baranes et al., 2015). With saturation, the learned and processed information will be owned as an interest and skill. Even thinkable is the possibility to enable mechanisms which can drive persistence in learning (for a selection of skills) which can lead learning to perfection, for instance like in the learning process when acquiring motor skills for playing an instrument. More precisely, robots must be able to act autonomously over longer periods of time, not only within interaction tasks, learning and to achieve goals, also to manage resources and to recover within time constraints. This requires greater flexibility in the way the robot chooses to sense, interact and remember in achieving goals, including an ability to identify when it has not been successful executing an action (Spranger et al., 2010). State of the art autonomous robots seek for human assistance if they recognise that they are in situations beyond their logic or sensing abilities. Typically, this is the case for robots with pre-programmed/planned tasks, this often imposes unwanted operational constraints and is limiting the use cases. Task planning adds sophistication to simple switching of behaviour in traditional autonomous architecture, so to say adaptive planning has to respond to execution conditions and re-plan when failures or opportunities occur. Most recent attempts to process complex decision making resulted in an impressive demonstration of network capabilities, where the ANN was playing "intuitive" in the game "Go" and able to beat a professional human player. The success of "AlphaGo" is based on an efficient combination of machine learning methods to train a pattern recognising deep learning network, plus Monte Carlo tree search to estimate the value of each game state in the production phase (Silver et al., 2016). Here the advantage was played out by the use of several ANN's for the purpose of training each other. This clearly demonstrated the technical feasibility of combined machine learning and reproduction phase, but the shortcomings lie in the tree search and the requirements of dedicated hardware for the sophistication of processing. If this can be improved with efficiency and exploration mechanisms for robots via functional network architectures and synthetic consciousness through a global workspace, a big progress step can be made. For these considered aspects, learning has to provide flexibility and robustness in execution of actions, as well teaching abilities, so that it allows for consideration of spatial and temporal variabilities. Further extensions to reinforcement learning, state-action maps (policies) and noise filtering can be applied as well. Artificial Intelligence offers the prospect of realising cognition desired for persistent autonomy

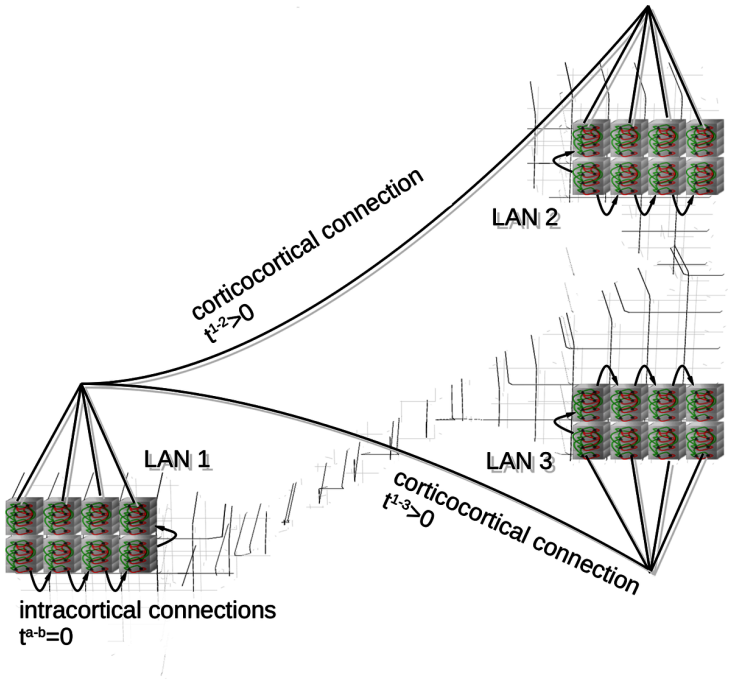
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

and the control of the robot can be stabilised and tuned to robot motion and contact conditions during task execution, whilst skill learning provides the means to respond to disturbances and errors within behaviour, without a recourse in planning. Intrinsic physical world modelling in AI processes sensor data to maintain a logical description of the environment, while it is used for directing the sensors accordingly to task execution and information seeking. AI status assessment monitors the performance of action execution and uses live sensor data for interpretation of the state of the environment to detect and diagnose failures and success of the task execution. Nevertheless, intelligence is as well devoted to spend time and resources to explore and obtain information, but the intrinsic motivation for such behaviour is not well discovered yet. Developments in machine learning and eye movement behaviour have shown, there is a connection between curiosity and information gain for learning progress mechanism in neuronal aspects (Baranes et al., 2015). A starting point in this research field of curiosity driven learning is the exploration of environment by its observers and actors, while they estimate their own epistemic states and how they specify their own uncertainties in order to find strategies to reduce theses uncertainties. Considerations show that possible motivations, besides gain control in elementary network aspects, can be derived from intrinsic and as well extrinsic factors. For extrinsic motivation the example of monitoring an obstacle while movements, with the goal of successfully moving towards a specific aim. This is a natural behaviour while pursuing a task-related sampling of information seeking, driven by maximising progress towards the goal. More complex is the goal setting of learning complex strategies, like in playing chess, over a long period of time. Intrinsically motivated information seeking in contrast is described as pursuing the goal as a goal in itself, like gathering information on reading for understanding, driven by positive reward mechanisms in the brain itself (Gottlieb et al., 2013). This falls back to the information sampling during solving a task by strategy progression and includes generating informational actions, as well learning. Machine learning approaches tackle this exploration or strategy searching for solving a task as an optimisation problem, where the task is modelled as a cost function and the strategy is simply to minimise this cost function.

Implementations take advantage of reinforced learning, stochastic optimisation, evolutionary algorithms and Bayesian optimisation. Supported is this procedure by division of duties, biological speaking e.g. eye movements and information filtering by target selection via cell activity. The agent monitors the uncertainty or predictability of action consequences and employs this quantity for the deployment of resources, like gaze. Finally, the environment becomes predictable by decreasing the uncertainty of future states through minimising the scattering or maximise the clustering of hidden states in the model. Another factor which can play a role is the trapping in unlearnable situations, because sampling can happen in a very large space. Contrary to the example used above, the real world is full of unsolvable tasks, maybe due to given limitations of the agent or irreducible uncertainty. Attention focus factors can be given by time dependencies for exploration based on heuristics and also task planning has to be constructive – one cannot walk before learning how to stand up. An attractive factor here can be provided by sensory novelty of information and uncertainty, while it is known biologically to enhance neuronal response, so that novelty can act as an intrinsic reward by dopamine release. Contextual novelty, like surprise, can combine sensory novelty with high prediction error and results in high attentional attraction. Additionally, the factor of curiosity is rising with awareness of differences for uncertainty between current state and reference state, as the goal or task. So this deprivation phenomenon is regarded to propel learning oriented tasks, which is driving the agent to fill a gap in knowledge for reward, e.g. minimising uncertainty about the state of the environment (Gottlieb et al., 2013).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Organisation scheme for AI in operational networks

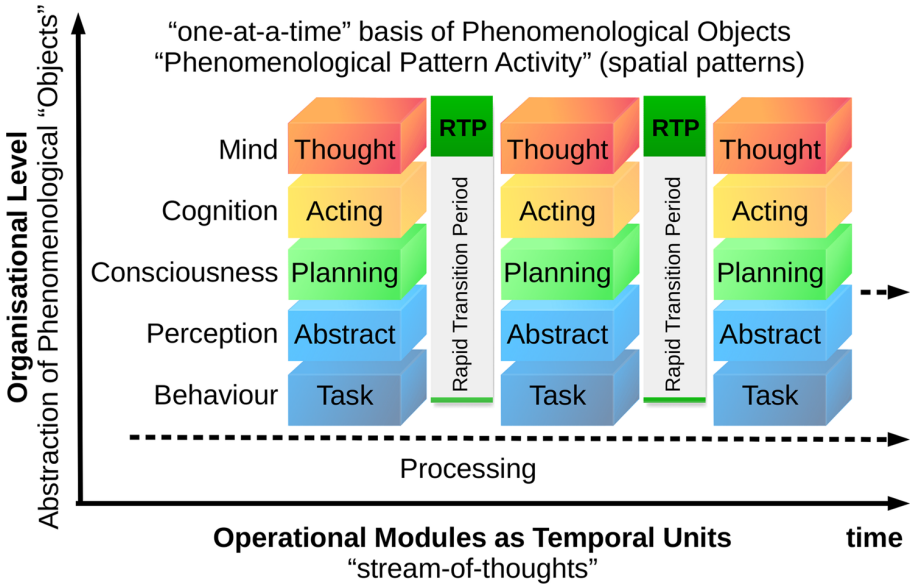


The specific modules shown in Figure 7 are artificial neural networks (ANN) in their own right and organised in Local Area Networks (LAN's), they can posses very different structures from each other, are interchangeable and can posses even specialised modelled neurone types. The overall global organisation is done via the interconnection and cross-linking of these modules, and especially by sharing a common workspace and memory. The connections are timed by distance to achieve spatiotemporal dynamics.

These arguments definitely speak again for a modular approach in implementing robotic AI and support as well the approach to combine machine learning with deep learning networks, this leads to a network of networks. Starting with the neuron codelets of specific types forming microcircuits and connecting these microcircuits as LANs can meet this idea, which is depicted in Figure 7. From a network of networks arise further aspects which are advantageous, already simplest network structures coupled via symmetric nearest neighbour connectivity matrices display a rich variety of oscillatory behaviour by applying time delays. These time delays arise from the network topologies where time t(i-j) is a function of distance between nodes |i-j| of network interconnections. While still in small local networks (microcircuits) the time delay for propagation is allowed to be neglected. The implied assumption of a homogeneous connection topology in the network, which is symmetric and translational invariant, is fulfilled by the fact that the recursion and inhibition in microcircuits mimics the heterogeneous connectivity observed in the biological structures of the cortex (Jirsa, 2004). The modular structure and the fact that neural information processing is not localised in one small brain area, plus a global workspace idea, the achievement of implementing several network classes and their hierarchy of connections leads beneficially to an interplay between connectivity and spatiotemporal dynamics. This interplay is eventually achieving the connection between time and space via geometry. This network-network connections can be organised in "global weight & connectivity matrices", again these are limited in flexibility, but in "sleep mode" plasticity

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 8. Scheme for an AI stream of "thoughts"



can be achieved through machine learning from the short-term memory. But more crucial at this point is the reproduction of "parent information processing" by abstraction, for example as semantic events by summation in local "Field Potentials" or potential fields for the production of higher-level patterns of neural activity. Biomimetic modelling, connectivity and functioning of significant parts of the brain to determine an architecture as of a series of nested sensorimotor loops are fundamental for this ideas. The networks have to consist out of tightly coupled layers for behavioural control; respectively the reactive, the adaptive and the contextual layers, and each one increasingly performing more memory dependent mappings from sensory states to action. The reactive layer performs reflex actions, including exploration, collision avoidance, search and homing, using sensory modalities and predefined sensory-motor mappings. The adaptive layer performs landmark recognition, feature extraction and memory segment construction, comprising heading direction and landmark information, for instance like implemented in self local area mapping ideas. These memory segments are sequenced in the short-term memory of the contextual layer until a goal state occurs. If this is recognised as a success state, the contents of the short-term memory are retained into long-term memory, like in hippocampus memory training. This architectures are already successfully implemented with modest general purpose computational resources e.g. for mapless navigation.

The illustration in Figure 8 relates neuronal assemblies and their dynamics with an operation and to neurophysiological abstract levels of operational modules, if cognitive phenomenal and behavioural operations coincide in time. Operational modules, as a result of synchronised operations caused by distributed transient neuronal assemblies, can be further synchronised between each other on shifted timescales as well and form so a more abstract and complex operational module, constituting an integrated experience. Consequently, functionally distinct operational modules can co-exist on different timescales.

Exemplarily implemented features of OA have shown promising results in decision making processes, in comparison to central state modelled decision making algorithms (Burmakin et al., 2009). Taken further is the idea of OA by a stream of "thoughts" implementation via operational modules in an abstraction

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

level hierarchy, see Figure 8, adapted from (Fingelkurts et al., 2012). Respectively, OA assumes, besides functional layers, the implementation of operational modules as time dependent connected abstractions of network activity. But the rapid transition periods, which basically reset parts of the working memory, are mutually independent between operational modules and coincide with self-organised stability in the network (Fingelkurts et al., 2012). This is to idealise semantic abstraction towards spatiotemporal pattern interpretations, which is an approach for generating sense in phenomenological space and time events. Apart from automatically resulting in a sense of time in artificial intelligence machines, the integration of information as a "stream-of-thoughts" means "spatial operational modules as temporal units", which means a sequential "stream of thought", but at the same time processing has from the global workspace all the information available and therefore unifies into a parallel processing of information, which promises to result in time dependent and predictive "thinking" in robots.

CONCLUSION

In contrast to a central processing unit (CPU), the performance of the human brain is clearly due to it's fundamentally different way of processing and its different state of matter. While computer chips follow mainly the paradigm of precision and fast calculation in sequential manner, the neuronal structures process massively parallel and abstract, while being less precise, but due to abstraction being more efficient and focusing on the real world representation. The global workspace idea gives rise to instantaneous access to contexts of information for processing and unifies the spatial networking and co-existence of time wise shifted functional spatio-temporal patterns. An artificial intelligence model, able to cover the latter aspects, should comprise to be intuitive, self-aware and even synthetically conscious for the means of an interactive robot, taught by biology. Joint aspects from different research disciplines and experiments lead to models of biomimetic approaches for artificial intelligence. To create robots, which should be able to learn, to recognise the environment, and to adjust to the change of situations respectively, research on cognitive abilities and AI is pulling on one string. By the biological role model, biomimetic aspects follow the modular organisation of the human brain as paragon and technologies like machine learning combined with deep learning networks will provide sufficient and efficient platforms to achieve these goals in robots. Therefore, the prediction is clear, that AI is able to mimic intuition and empathy in robots for intuitive user interactions.

REFERENCES

Baars, B. J. (2002, January). The conscious access hypothesis: Origins and recent evidence. *Trends in Cognitive Sciences*, *6*(1), 47–52. doi:10.1016/S1364-6613(00)01819-2 PMID:11849615

Baars, B. J. (2005). Global workspace theory of consciousness: Toward a cognitive neuroscience of human experience. *Progress in Brain Research*, *150*, 45–53. doi:10.1016/S0079-6123(05)50004-9 PMID:16186014

Baars, B. J., Ramsøy, T. Z., & Laureys, S. (2003, December). Brain, conscious experience and the observing self. *Trends in Neurosciences*, *26*(12), 671–675. doi:10.1016/j.tins.2003.09.015 PMID:14624851

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Baranes, A., Oudeyer, P. Y., & Gottlieb, J. (2015). Eye movements reveal epistemic curiosity in human observers. *Vision Research*, *117*, 81–90. doi:10.1016/j.visres.2015.10.009 PMID:26518743

Binzegger, T., Douglas, R. J., & Martin, K. A. C. (2007, November). Stereotypical Bouton Clustering of Individual Neurons in Cat Primary Visual Cortex. *The Journal of Neuroscience*, *27*(45), 12242–12254. doi:10.1523/JNEUROSCI.3753-07.2007 PMID:17989290

Bionik, V.-F. (2012). Biomimetics – Conception and strategy – Differences between bio- mimetic and conventional methods/products. *VDI-Handbuch Bionik, 1*, 1-1. Retrieved from https://www.vdi.de/richtlinie/vdi_6220_blatt_1-bionik_konzeption_und_strategie_abgrenzung_zwischen_bionischen_und_konventionellen/

Bongard, J., Zykov, V., & Lipson, H. (2006, November). Resilient Machines Through Continuous Self-Modeling. *Science*, *314*(5802), 1118–1121. doi:10.1126cience.1133687 PMID:17110570

Born, J., & Wilhelm, I. (2012, March). System consolidation of memory during sleep. *Psychological Research*, *76*(2), 192–203. doi:10.100700426-011-0335-6 PMID:21541757

Burmakin, E., Fingelkurts, A. A., & Fingelkurts, A. A. (2009). Self-organization of dynamic distributed computational systems applying principles of integrative activity of brain neuronal assemblies. *Algorithms*, *2*(1), 247–258. doi:10.3390/a2010247

Carter, R., Aldridge, S., Page, M., Parker, S., Frith, C., & Frith, U. (2010). *Das Gehirn: Anatomie, Sinneswahrnehmung, Gedaechtnis, Bewusstsein, Stoerugen* (K. Hofmann & J. Wissmann, Trans.). Dorling Kindersley.

Chella, A., Frixione, M., & Gaglio, S. (2008, October). A cognitive architecture for robot self-consciousness. *Artificial Intelligence in Medicine*, *44*(2), 147–154. doi:10.1016/j.artmed.2008.07.003 PMID:18715770

Chella, A., & Macaluso, I. (2009, January). The perception loop in CiceRobot, a museum guide robot. *Neurocomputing*, *72*(4-6), 760–766. doi:10.1016/j.neucom.2008.07.011

Dehaene, S., Changeux, J. P., Naccache, L., Sackur, J., & Sergent, C. (2006). Conscious, preconscious, and subliminal processing: A testable taxonomy. *Trends in Cognitive Sciences*, *10*(5), 204–211. doi:10.1016/j.tics.2006.03.007 PMID:16603406

Dehaene, S., & Naccache, L. (2001, April). Towards a cognitive neuroscience of consciousness: Basic evidence and a workspace framework. *Cognition*, *79*(1–2), 1–37. doi:10.1016/S0010-0277(00)00123-2 PMID:11164022

Distefano, S., Milano, P., & Mazzara, M. (2015). Neuromodulating Cognitive Architecture. *Conference Paper*.

Douglas, R. J., & Martin, K. A. (2004). Neuronal Circuits of the Neocortex. *Annual Review of Neuroscience*, *27*(1), 419–451. doi:10.1146/annurev.neuro.27.070203.144152 PMID:15217339

Douglas, R. J., & Martin, K. A. C. (1990, September). Control of Neuronal Output by Inhibition at the Axon Initial Segment. *Neural Computation*, *2*(3), 283–292. doi:10.1162/neco.1990.2.3.283

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Douglas, R. J., & Martin, K. A. C. (2009). Inhibition in cortical circuits. *Current Biology*, *19*(10), 398–402. doi:10.1016/j.cub.2009.03.003 PMID:19467204

Fingelkurts, A. A., Fingelkurts, A. A., & Neves, C. F. H. (2012). Machine consciousness and artificial thought: An operational architectonics model guided approach. *Brain Research*, *1428*, 80–92. doi:10.1016/j.brainres.2010.11.079 PMID:21130079

Franklin, S. & Patterson, F. (2006). The lida architecture: Adding new modes of learning to an intelligent, autonomous, software agent. *pat, 703*, 764–1004.

Gamez, D. (2010). Information integration based predictions about the conscious states of a spiking neural network. *Consciousness and Cognition*, *19*(1), 294–310. doi:10.1016/j.concog.2009.11.001 PMID:20018526

Gottlieb, J., Oudeyer, P. Y., Lopes, M., & Baranes, A. (2013). Information-seeking, curiosity, and attention: Computational and neural mechanisms. *Trends in Cognitive Sciences*, *17*(11), 585–593. doi:10.1016/j.tics.2013.09.001 PMID:24126129

Jirsa, V. K. (2004). Connectivity and dynamics of neural information processing. *Neuroinformatics*, *2*(2), 183–204. doi:10.1385/NI:2:2:183 PMID:15319516

Kaku, M. (2014). *Die Physik des Bewusstseins: Uber die Zukunft des Geistes*. Rowohlt Verlag GmbH.

Koch, C. (2009). *A "Complex" Theory of Consciousness.* Retrieved 2015-10-24, from http://www.scientificamerican.com/article/a-theory-of-consciousness/

Koch, C. (2013). *Bewusstsein*. Berlin: Springer Berlin Heidelberg. doi:10.1007/978-3-642-34771-9

Lane, D. M. (2012). Persistent autonomy artificial intelligence or biomimesis? In Autonomous underwater vehicles (auv), 2012 ieee/oes (pp. 1–8). doi:10.1109/AUV.2012.6380719

Massimini, M. (2005, September). Breakdown of Cortical Effective Connectivity During Sleep. *Science*, *309*(5744), 2228–2232. doi:10.1126cience.1117256 PMID:16195466

Nguyen, S. M., & Oudeyer, P.-Y. (2013). Socially guided intrinsic motivation for robot learning of motor skills. *Autonomous Robots*, *36*(3), 273–294. doi:10.100710514-013-9339-y

Oudeyer, O. (2013). *Object learning through active exploration*. Hal.Archives-Ouvertes.Fr.

Pfutzner, H. (2014). *Bewusstsein und optimierter Wille*. Berlin: Springer Berlin Heidelberg.

Raffone, A., & Pantani, M. (2010, June). A global workspace model for phenomenal and access consciousness. *Consciousness and Cognition*, *19*(2), 580–596. doi:10.1016/j.concog.2010.03.013 PMID:20382038

Reggia, J. A. (2013). The rise of machine consciousness: Studying consciousness with computational models. *Neural Networks*, *44*, 112–131. doi:10.1016/j.neunet.2013.03.011 PMID:23597599

Rosenthal, D. (2009). *Concepts and definitions of consciousness. In Encyclopedia of Consciousness*. Amsterdam: Elsevier.

Shanahan, M. (2006). A cognitive architecture that combines internal simulation with a global workspace. *Consciousness and Cognition*, *15*(2), 433–449. doi:10.1016/j.concog.2005.11.005 PMID:16384715

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Shanahan, M., & Baars, B. (2005, December). Applying global workspace theory to the frame problem. *Cognition*, *98*(2), 157–176. doi:10.1016/j.cognition.2004.11.007 PMID:16307957

Sherrington, C. S. (1908). Reciprocal innervation of antagonistic muscles. thirteenth note.-on the antagonism between reflex inhibition and reflex excitation. *Proceedings of the Royal Society of London. Series B, Containing Papers of a Biological Character*, *80*(544), 565–578. doi:10.1098/rspb.1908.0053

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., van den Driessche, G., & Hassabis, D. (2016, January). Mastering the game of Go with deep neural networks and tree search. *Nature*, *529*(7587), 484–489. doi:10.1038/nature16961 PMID:26819042

Snaider, J., McCall, R., & Franklin, S. (2011). The LIDA Framework as a General Tool for AGI. In D. Hutchison & ... (Eds.), *Artificial General Intelligence* (Vol. 6830, pp. 133–142). Berlin: Springer Berlin Heidelberg. doi:10.1007/978-3-642-22887-2_14

Spranger, M., Thiele, C., & Hild, M. (2010). Integrating high-level cognitive systems with sensorimotor control. *Advanced Engineering Informatics*, *24*(1), 76–83. doi:10.1016/j.aei.2009.08.008

Tagliazucchi, E., Chialvo, D. R., Siniatchkin, M., Brichant, J.-F., Bonhomme, V., … Laureys, S. (2015, September). *Large-scale signatures of unconsciousness are consistent with a departure from critical dynamics.* arXiv:1509.04304

Taketani, M. (1971). On formation of the Newton Mechanics. *Progress of Theoretical Physics*, *50*(Supplement), 53–64. doi:10.1143/PTPS.50.53

Thompson, R. (2010). Das Gehirn: Von der Nervenzelle zur Verhaltenssteuerung (A. Held, Trans.). Spektrum Akademischer Verlag.

Tononi, G. (2004). An information integration theory of consciousness. *BMC Neuroscience*, *5*(1), 42. doi:10.1186/1471-2202-5-42 PMID:15522121

Tononi, G. (2008). Consciousness as integrated information: A provisional manifesto. *The Biological Bulletin*, *215*(3), 216–242. doi:10.2307/25470707 PMID:19098144

Tononi, G. & Koch, C. (2014). *Consciousness: Here, There but Not Everywhere.* Arxiv, 15.

ENDNOTE

[1] Bouton: Knoblike enlargement at the end of an axon, where it forms a synapse with other neurons.

This research was previously published in Advanced Research on Biologically Inspired Cognitive Architectures edited by Jordi Vallverdú, Manuel Mazzara, Max Talanov, Salvatore Distefano, and Robert Lowe; pages 83-116; copyright year 2017 by Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 73
How to Identify Rheumatic Diseases by General Physicians

Eduardo C. Contreras
Autonomous University of Coahuila, Mexico

Gustavo J. Puente
Autonomous University of Coahuila, Mexico

ABSTRACT

A large part of the population in countries in process of development ignores what Rheumatic Diseases are, and general practitioners are in most cases unaware of enough information to identify them and the treatments to successfully control them. A proposal to help those general practitioners to detect if an articular condition belongs to a Rheumatic Disease case is to present them the clinical semiology that should lead them to redirect the given conditions to a specialist on the subject, a rheumatologist. The clinical semiology is presented by an automated algorithm inside a goal-based software agent, containing all the necessary information to identify the seven most common inflammatory Rheumatic Diseases, and fourteen of the non-inflammatory ones. The purpose of this tool is to provide the general practitioner with the correct information to redirect the patient with a rheumatologist, in order for it to receive the appropriate medication to be controlled.

INTRODUCTION

Project

This project is a field investigation, applied to medical area on the Rheumatology specialty related with Clinical Semiology and the Information Technology. It is in itself, an automated medical tool to identify joints affections, based on an Intelligent Agent from Artificial Intelligence specialty, based on the heuristic compilation of the tacit knowledge of rheumatologists. This study tries to associate the factors than relate the support process on Information Technology on a General Doctor, towards a modern tool as didactic software, which facilitates the most common rheumatic's affections identification.

DOI: 10.4018/978-1-7998-0951-7.ch073

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

This identify tool has a section of rheumatic inflammatory conditions and a section of rheumatic non-inflammatory conditions. Inflammatory conditions referred to joint. In the section of inflammatory conditions the user will find windows that include the different aspects of the joint disease that must be considered in diagnosis such as kind of onset, number and size of involved joints, topography of the arthropathy and symmetry of involvement. The section of non-inflammatory conditions considers only localized, regional and generalized involvement. Both sections follow a reasonable and logic algorism that allows reaching diagnosis of the rheumatic conditions included in the diagnostic tool.

OBJECTIVE

To value the utility of the algorism for the identification of Rheumatic Affections" with a general practitioner, so that he will be able to provide an accurate diagnosis of joint diseases and redirect the patient with a rheumatologist for its early evaluation.

The proposed clinic semiology is based upon the human experience that is used to build a software agent didactic tool related to the Information Technologies; the software agent is a tool that allows the doctor to offer an appropriate evaluation of the rheumatic diseases.

BACKGROUND

Rheumatic Diseases

The study and management of rheumatic diseases in a rigorously scientific way has been in charge of the medical specialty of Rheumatology, a branch of Internal Medicine, since the middle of the 19th century, making it an medical area relatively new and this explains why, even in the beginning of the 21st century, rheumatic diseases are largely unknown both by the population in general and by doctors, who are not dedicated to this area of health.

Rheumatic diseases are heterogeneous clinical entities which affect musculoskeletal (Schumacher et al, 1998) system and other systems and organs, whose diagnosis often comes late due to little knowledge of these diseases and scarcity of rheumatologists, among other reasons (Taranta & Markowitz, 1981). Delay in diagnosis of these diseases may lead to damage of joints and related tissues, organs and systems with subsequent functional disability and increase in morbidity and mortality.

A person who suffers from a rheumatic disease that has affected the hip joint is facing a severe pain when attempting to walk. The first three or five steps walks with great difficulty, as it decreases the pain, once that it lessens can walk with less pain and greater agility. When attempting to stand the pain is so acute that it does not balance and when sitting drops the body waiting for that pain to decreases, spending time it dulls, but knowing that when get up again will face that terrible acute pain. When the joints of the hand are concerned it is very difficult to remove the lid to a jar, as pain does not allow you to force. If the person suffers from a knee, usually the cartilage that covers the joint decreases its thickness, staying in touch with the bone part, decreasing joint mobility and increasing pain to move it. Usually the patient will live with the pain articulates for a lifetime, with the probability that the joint will be merged and significantly decrease their movement, in such degree that it turns into a person with a

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

disability. Most of the people must be in constant motion, as it permits the disease, otherwise it would be in for life in a wheelchair or bed.

When a rheumatic disease affects joints, they are damaged in an irreversible process; therefore it is substantial to identify them as soon as possible. If general practitioners supply inappropriate medication, they will be at great risk of affecting the patient severely. The rheumatologist is the only appropriate person to administer the correct medication, due to the existence of the called "antirheumatics".

Distinctive features of this disease are pain, stiffness and articular inflammation that would only decrease after the consumption of the rightful medication.

Definition of the Common Ailments

- **Rheumatoid Arthritis:** A chronic inflammatory disease of unknown etiology which involves multiple diartrodial joints leading to hyperplasia of synovial tissue and destruction of cartilage and subchondral bone with subsequent deformity and functional limitation. (Stone, 2009)
- **Osteoarthritis or Degenerative Joint Disease:** A no inflammatory disorder of cartilage of the movable joints which suffers deterioration due to biochemical and structural changes in its matrix and in the subchondral bone.(Cañete,2008)
- **Gout:** A metabolic disease characterized by hyperuricemia and episodes of acute arthritis and progression to a generalized arthropaty with accumulation of monosodium urate (Wright & Harvey, 1988).
- **Spondiloarthritides:** A heterogeneous group of diseases characterized by inflammation of the joints of the spine, sacroiliac, limbs joints and in variable frequency affects also skin, mucosal of the digestive tract, urogenital tract and eyes.(Kats,1977)
- **Systemic Lupus Erythematous:** An autoimmune disease characterized for the appearance of multiple autoantibodies which participate in inflammatory disease of diverse organs and systems.
- **Juvenile Arthritis:** A term which refers to a group of inflammatory diseases of the joints in pediatric populations until the age of 16 years.
- **Infectious or Septic Arthritis:** A disease caused by invasion of the joints by bacteria and other microorganisms.
- **Fibromyalgia Syndrome:** The name of a condition characterized by chronic widespread musculoskeletal neuropathic pain in specific sites. (Clauw, 2005)
- **Carpal Tunnel Syndrome:** A clinical entity characterized by entrapment of the median nerve when it passes through a tunnel formed by the carpal bones and joints and the transverse carpal ligament.
- **Digital Stenosis Tenosynovitis:** (Of the fingers) is a constriction of the lumen of the fibrous sheath of flexor tendons of the fingers which impedes its free gliding.(Salter, 1998)
- **D'Quervain Tendinitis:** A condition that appears due to excessive friction between the tendon abductor policis longus and extensor pollicis brevis and its common fibrous sheath.
- **Epicondylitis or Tennis Elbow:** A degenerative tendon disease located in the origin of the forearm extensor muscles.
- **Rotator Cuff Tendinitis:** An inflammatory condition located at the insertion of the rotator muscles of the shoulders in the humeral head.
- **Dupuytren's Contracture:** A progressive retraction of the fibrous tissues of the palmar fascia.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Biological Therapies

Biological therapy (also called immunotherapy or biotherapy) often employs substances called biological response modifiers (BRMs). Biological therapy is thus any form of treatment that uses to stimulate or restore the ability of the immune (defense) system to fight infection and disease. Biologic therapy to block the action of a messenger of inflammation called tumor necrosis factor (TNF) is being used to treat conditions such as Crohn's disease and rheumatoid arthritis.

Biologics are genetically engineered proteins derived from human genes. They are designed to inhibit specific components of the immune system that play pivotal roles in fueling inflammation, which is a central feature of rheumatoid arthritis.

Antirheumatic drugs are drugs used to treat rheumatic diseases. The major classes of antirheumatic drugs include Nonsteroidal Anti-Inflammatory Drugs (NSAIDs), Corticosteroids, Disease Modifying Anti-Rheumatic Drugs (DMARDs), Slow-Acting Antirheumatic Drugs (SAARDs).

- **Nonsteroidal Anti-Inflammatory Drugs (NSAIDs):** Drugs belonging to this class bring symptomatic relief of both inflammation and pain, but have a limited effect on the progressive bone and cartilage loss associated with rheumatoid arthritis. They act by slowing the body's production of prostaglandins.
- **Corticosteroids:** These drugs are very powerful anti-inflammatory agents. They are the synthetic analogs of cortisone, produced by the body. Corticosteroids are used to reduce inflammation and suppress activity of the immune system.
- **Disease Modifying Anti-Rheumatic Drugs (DMARDs):** Influence the disease process itself and do not only treat symptoms, hence their name. DMARDs also have anti-inflammatory effects, and most were borrowed from the treatment of other diseases, such as cancer and malaria.
- **Slow-Acting Antirheumatic Drugs (SAARDs):** Are a special class of DMARDs and the effect of these drugs is slow acting and not so quickly apparent as that of the NSAIDs. (Medical Dictionary)

The treatment in a patient with anti-rheumatic drugs active ankilosing spondylitis (AS) and spondyloarthritides (PsA) whose condition is not sufficiently controlled with non-steroidal anti-inflammatory drugs (NSAIDs) in the case of axial disease, and sulphasalazine or methotrexate in the case of peripheral arthritis. Anti-tumor necrosis factor (Anti-TNF) therapy seems to be a powerful tool for the treatment of AS and other SpA.

Biological therapies for the treatment of rheumatoid arthritis, currently approved, are available: Etanercept, Infliximab, Adalimumab, Anakinra, Tocilizumab, Rituximab, Abatacept. Biologic agents most frequently used as first option are infliximab and etanercept. The progress in manufacturing biotechnology has contributed to the development of several other prospective agents that may form the basis for the therapy of rheumatoid arthritis in the near future. (Šenolta et al, 2009)

On a study in the hospital La Fé located in Valencia Spain, descripts the registry of adverse events of biological therapies in rheumatic diseases. It contains the description of: frequency of adverse events by groups, frequency of serious adverse events, frequency of fatal adverse events, incidence rate of adverse events and incidence rate of serious adverse events. (Research Unit of the Spanish Society of Rheumatology, 2013.)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The treatment in a patient with anti-rheumatic drugs active ankilosing spondylitis (AS) and spondyloarthritides (PsA) whose condition is not sufficiently controlled with non-steroidal anti-inflammatory drugs (NSAIDs) in the case of axial disease, and sulphasalazine or methotrexate in the case of peripheral arthritis. Anti-tumor necrosis factor (Anti-TNF) therapy seems to be a powerful tool for the treatment of AS and other SpA. (Braun & Sieper, 2004)

Biosimilars is the term used for generic versions of biologic agents, with identical primary sequences in which changes in the manufacturing processes have led to important differences in safety and efficacy. Introduction of biosimilars requires a specifically designed pharmacovigilance plan. (Kolkata, 2012)

Home Remedies for Arthritis and Joint Pain

In countries in process of developing most of the people relies on home remedies when they suffer a condition, given that they have little information, because of their culture, economic situation or the fact that they don't have access to specialized medical attention. In the case of suffering a joint pain without receiving a professional diagnosis, people often believe this pain is derived from a rheumatic origin, thus applying to themselves popular remedies such as: putting a potato, a magnet or a piece of copper or silver in their pockets; applying a poultice of ingredients like lettuce, smashed garlic, nutmeg, figs or cabbage; ingesting sage, leek or poppy seed infusions; and finally, drinking Turmeric and Ginger Tea, White Willow Tea, or Juniper Berry Tea.

Another home remedies are: Lubricate joints With Extra Virgin Olive Oil, Peppermint Eucalyptus Oil Blend, Blackstrap Molasses Drink and Pectin and Grape Juice.

These remedies don't change the course of the arthritis itself, but they do have an analgesic, or pain-relieving effect.

Clinical Semiology

In medicine, the clinical semiology is the group of knowledge that deals with the identification of the various pathological manifestations (signs and symptoms), from people and animals (in the veterinary field), in both cases with the objective of how to collect data.

Semiotics. With the objective of compile this data in syndromes and how to interpret them (clinical semiology). This discipline allows the physician not only directed at the diagnosis but have a prognostic assessment.

Unless other diseases, the traditional semiology teaching-learning process is focused on isolated concepts and practices. Rheumatic diseases, for example, need a systemic approach for their diagnosis.

This approach allows the study of the whole thing, it's not enough to study an only part or region of the human body because the rheumatic diseases come all from a mismatch on the immunological system. If a general practitioner applies immunosuppressers to a patient with a rheumatoid arthritis condition, with glaucoma, diabetes or an infectious condition, without considering the concomitant effect, this would affect the patient severely. For this reason, only a specialized practitioner, a rheumatologist, can prescribe the adequate medication without the risk of secondary effects.

For all the reason previously mentioned, the semiology must have a systematic approach (Bertalanffy, 2011). Semiology is the mainstay of clinical medicine. It is an art and a science. It presents a method of sorting knowledge (clinical method) and the goal is the diagnosis of health problems.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The methods applied on Semiology are complex and differ one from another, mainly because their use as a method and recognition as a science have stopped in the present time.

According to the systemic approach, questionnaire, observation, physical examination and lab exams are not enough, because it involves historical methods, sociology, genetics (biological inheritance), psychology, psychiatry, etc., among other sciences.

- Rheumatology Method,
- Questionnaire,
- Physical Exploration on Joints,
- Inspection,
- Palpation,
- Mobility,
- Arthrocentesis,
- Lab exams,
- Reactive C Protein,
- Rheumatoid factor,
- HLA-B27 HLA-DRB1 HLA- DR4,
- Erythrocyte sedimentation rate (ESR),
- Antinuclear antibodies,
- Radiological examination,
- Anteroposterior,
- Lateral,

Radiographic semiology to detect:

- Ankylosing hyperostosis,
- Ankylosing spondylitis,
- Rheumatoid Arthritis,
- Paget's Disease,
- Neoplasms.

Software Agent Definition

In artificial intelligence, a software agent (IA) is an autonomous entity which observes and acts upon an environment and directs its activity towards achieving goals. (Rich & Knight, 1991) An agent is everything that can be considered that perceives its atmosphere by means of sensors and that response or acts in such atmosphere by means of effectors. (Russell & Norvig, 2003) A sensor is a device that measures a physical quantity and converts it into a signal, which can be read by an observer or by an instrument. For example, voltmeter, thermometer, thermocouple, seismometer, parking sensor, carbon monoxide detector. Effectors are agents or structures that cause an activity, like a nerve causing a muscle to flex, or a cell triggering an immune response to a foreign substance.

In robotics, an end effector is a device or tool that's connected to the end of a robot arm where the hand would be. The end effector is the part of the robot that interacts with the environment. The structure

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

of an end effector and the nature of the programming and hardware that drives it depend on the task the robot will be performing. (Resconi, 2004)

Intelligent agents are often described schematically as an abstract functional system similar to a computer program. On the Internet, an intelligent agent (or simply an *agent*) is a program that gathers information or performs some other service without your immediate presence and on some regular schedule.

Models to Construct Agent's Programs

- **Agents from Simple Reflex:** (Russell & Norvig, 2003) The rules condition-action allows the agent to establish the connection between perceptions and actions.
- **Agents Good Informed into Everything What Happens:** It updates the information based in its own internal state.
- **Agents Based on Goals:** It is a combination of the previous ones, with the information relative to the result than will produce the possible actions than will undertake. It handles the information that use to update the internal state and this way to select those actions that allow reaching the goal.
- **Agents Based on Utility:** It allows to differentiate a state from another and to take the decision that offers major utility to the agent.

Of these types we will be based only on Agents Based on Goals. Besides the states, agents need some kind of information about its goals. Goals will to detail the situations to that it is desired to arrive. The agent program could to combine goals with actions and in this way, to be able to choose those actions that allow reaching the goal.

Design of an Ideal Agent

To specify which kind of action should undertake an agent, as an answer to a certain sequence of perceptions constitutes the design of an ideal agent.

Agent Type: Medical Identification System

- **General Doctor's Perceptions:**
  - Symptoms, clinical signs, physical examination, heredity, personal data
- **Actions:**
  - Identifies: joint state, anatomical region, number of joints.
  - How the patient evolves.
  - Differentiate diseases.
- **Goals:**
  - Affection identification, time reduction, canalization of patient in case of such another disease.
  - Atmosphere.
  - Patient, general doctor.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Medical Identification System for Rheumatic Diseases (MISRheD)

The actions than this agent undertakes, are based exclusively in an integrated knowledge (provided by the rheumatologist), and so, its own perceptions are ignored (it does not count with sensors towards the environment), therefore the agent has not autonomy. Its conduct is defined by the human agent experience. It is reasonable to equip an agent of artificial intelligence with certain initial knowledge and capacity to learn (the rheumatology heuristic).

Similar Studies

Rejón established relationship between the characteristics of clinical judgment, medical semiology, and psychiatric semiology. Abductive inferences are especially useful to balance universal and singular information. (Rejón, 2013)

Noachtar and Peters mentioned that the Semiology has served for a long time as the major tool to classify epilepsy syndromes. They gives an introduction into the semiological seizure classification. (Noachtar & Peters, 2009)

Sadly most physicians depend on imaging services and laboratory tests far more frequently that a carefully executed history and physical examination for establishing a diagnosis or monitoring a therapeutic response. (Saravolats, 2007)

Monteagudo on the doctoral thesis presents didactic software for teaching the clinical semiology of the respiratory system. (Monteagudo, 2003)

In the game "Camino Cardíaco", it's goal was to promote a new educational resource in teaching the topic cardiovascular nursing assessment, the subject of semiotics, in addition to encouraging greater student interaction in the teaching-learning process. (Martínez, 2010)

De León proposes the use of hypertext in medical students for learning rheumatic diseases by using images to identify and interpret the signs and symptoms higher frequency. (De León, 2012)

Rheumatic Diseases: New Insights for the Healthcare Professional: 2013 Edition is a ScholarlyBrief that delivers timely, authoritative, comprehensive, and specialized information about Additional Research in a concise format.

A descriptive and prospective observational study was conducted "Ocular manifestations of rheumatic diseases in children" with the objective of characterizing the most frequently ocular manifestations in the course of the rheumatic diseases. (Parapar, 2007)

An Ontology-based intelligent agent for respiratory waveform classification. In this paper presents an ontology-based intelligent agent for respiratory waveform classification to help the medical staff with the judging the respiratory wave form from the ventilator.(Lee &Wang 2006)

Was as applied Utility-based Agents for Hospital Scheduling: A Conjoint-Analytical Approach for the scheduling of centralized operating theatres in large hospitals can be regarded as an archetypal co-operative decision problem.

The University of Aberdeen researchers are interested to collaborate with India on development of trusted mobile technologies for rural healthcare. The aim of the project is to explore the potential of trusted mobile technologies and to improve the management of chronic diseases in rural areas of the UK and India.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

A Novel Approach for Cardiac Disease Prediction and Classification Using Intelligent Agents. In this paper the goal is to develop a novel approach for cardiac disease prediction and diagnosis using intelligent agents.

Exploiting social reasoning of open multi-agent systems to enhance cooperation in hospitals.

This article explains how agents can exploit their capacity in order to support the medical staff in a hospital complex specialized in the treatment of burn patients, either by cooperating with other agents, or with human actors that they represent in the virtual system.(Horn,1999)

MAIN FOCUS OF THE CHAPTER

Problem Definition

Problem Identification

Most of the Mexican population, including the medical people, does not have the necessary information on rheumatic diseases to enable them to understand the importance of them and their socio-economic and psychological impact.

The Rheumatology is a relatively new medical specialty. In Mexico it was until the 1960s when it began with the training of specialists in this area of health. Most of the schools and faculties of medicine, both the public and private sector lack, even today, of the field of Rheumatology as part of their curricula and there is a limited number of specialists in our country. According to data collected and provided by the National Institute of statistics and geography until the year 1995 there were 59 faculties of medicine in our country; and it was until 1990 when it included the field of Rheumatology in these powers, so that medical graduates prior to this date were not given adequate information on this area, and only obtained some knowledge about the pathophysiology Traumatology, orthopedics and dermatology; matters that are related to the Rheumatology but do not have full knowledge about the same. For this reason there are currently a very small number of specialists in this area in our country.

Even though the Mexican College of Rheumatology Association Civil, has carried out information campaigns through the specialists affiliated to it and has also established the *National Day of the Rheumatic Patient* and other institutions have been involved in campaigns such as the Ministry of public health and private health institutions, as well as schools and faculties of medicine in our country have not achieved the objectives to inform the population in general on this condition.

All of these factors are contributors to the very low and erroneous information on rheumatic diseases in our country. Despite efforts to disseminate information on rheumatic diseases, are still insufficient the established programs for this purpose until the moment.

The population taken care by general doctors, or those from the population that has no access to the health services, due to its low economical condition or extreme poverty, this due to in Mexico is an indispensable requirement to access this health service system to be employed by some company, institution or government. Most of the population in Mexico is served by the main clinics hospitals of Social Security (Social Security Mexican Institute, Social Security Institute to Serve the State Workers, and specifically in Saltillo's Municipality by Magisterial Clinic Hospital) the rightful claimants must request its appointment in first instance with the general doctor, and this as well depending on the disease and also with some luck, it will alternate to the specialist doctor: however, the patient waits till three

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

months to obtain an appointment with the specialist doctor. This without a doubt has consequences that it can worsen the rheumatic disease of patients due to they are not served by a specialist doctor with the knowledge to identify the affection and to treat it suitably.

Medical specialists in the field estimated that 10% of the population develops over the course of his life any rheumatic disease and because general practitioners not diagnosed in time and correctly the disease is progressive and irreversible can subsequently cause a disability which may have been avoided if they had received the proper attention of a specialist.

Problem Quantification

In Other Countries

Rheumatic diseases in Spain according from data obtained from the National Institute of Social Security were in 2007 spending of 1.678 billion euros in temporary work disability. It is estimated that in Spain, approximately ten million people suffer from a rheumatic disease, 22.6% of population and cause between 10 and 15% of the primary care.

According to Statistics in USA: nearly 70 million Americans suffer from some form of arthritis or symptoms of joints chronic, rheumatic diseases are the leading cause of disability in people of 65 years of age and older, rheumatoid arthritis, how more disabling of arthritis affects approximately 2,100,000 Americans and attacks on women are two to three times more than males. These data come from Centers for the Disease Control and Prevention, National Institute of Arthritis and Musculoskeletal and Skin Diseases, which are part of the National Institutes of Health and Arthritis Foundation.

According to the statistics managed by the National Center for rheumatic diseases (CNER), in Venezuela, 20 of every 100 Venezuelans have been diagnosed with a disease of this kind.

In Our Country

In Mexico is not satisfied with the recommendation of the World Health Organization (WHO), that there is a rheumatologist per hundred thousand inhabitants, because there are only 560 Rheumatologist certificates, this is, a specialist for every 200 thousand inhabitants, it is not enough.

Rheumatic diseases are a major cause of morbidity and mortality. The available information indicates that the Systemic Lupus Erythematous has a mortality rate of 10% for the five years of its inception and 20% after 10 years; the rheumatoid arthritis reduces the survival in a range of 3 to 15 years compared with the general population and are diseases that on course have significant medical complications. Both diseases often affect people in productive age (18 to 60 years).

Statistics of the Mexican Institute of Social Security (IMSS) in our country report that until 1974 25% of disability pensions were granted by rheumatic ailments, and in 1978, 34% of the dictums of disability were granted for patients with diseases of the muscular-skeletal system. Rheumatic diseases are among the top ten causes of demand for health care in the family doctor system.

In Our Region

A number of questionnaires applied in the city of, Saltillo, Coahuila. To a based sample of 190 people older than 13 years (except for medical specialists in rheumatology) gave the following results:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

From the sample of the population of Saltillo the 18.42% have information about rheumatic diseases, the 62,63% they have never been to talks or known information about these diseases, the 50.52% don't know that there is a medical specialty dealing with rheumatic diseases, the 27.36% believe that rheumatic diseases affect only to people of the third age, the 55.78% believes that rheumatic diseases affect little in economic or social, to the 61.57% almost aren't interested in further information on rheumatic diseases..

In the city of Saltillo with a population of 578 046 inhabitants only has 4 medical specialists in the medical industry.

Development Method of the Software Agents

The proposal to solve the problem is to identify the type of software agents (intelligent agent) suitable, according to the following elements:

Human agent: to identify a rheumatic disease, knowledge and heuristics of the specialist physician in rheumatology, which will provide the basis of knowledge, rules, diagnostic approach in rheumatology, interpretation of radiological studies, laboratory and physical examination is necessary. The experience of the physician rheumatologist for issuing a diagnosis relies 70% on interrogation, 20% in physical exam and 10% in laboratory and radiological tests.

- **Environment:** Specify the environment in which the agent will run, in this case the main actors are: the general practitioner and the patient.
- **Percepts:** Determine perceptions, entry of data to the agent, identify the types of sensors that the agent will communicate with the environment; in this case the sensors are replaced by the perceptions of the general practitioner.
- **Effectors:** Determine the effectors, mechanisms of response of the agent, in our case is the information displayed on the monitor.

When the above elements have been defined it identifies the type of appropriate agent that will better solve the problem on the basis of the needs and resources available.

Agent Type will be based only on Agents based on goals. Besides the states, agents need some kind of information about its goals. The goals will detail the situations which it is desired to arrive. The agent program could combine goals with actions and in this way, to be able to choose those actions that allow reaching the goal.

The actions than this agent undertakes, are based exclusively in an integrated knowledge (provided by the rheumatologist), and so, its own perceptions are ignored (it does not count with sensors towards the environment), therefore the agent has not autonomy. Its conduct is defined by the human agent experience. It is reasonable to equip an agent of artificial intelligence with certain initial knowledge and capacity to learn (the rheumatology heuristic).

Identification System of Rheumatic Diseases

Agent Type: Medical Identification System of Rheumatic Diseases

- **Percepts:**
  - General Doctor's perceptions.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- Symptoms, clinical signs, physical examination, heredity, personal data.

- **Actions:**
  - Identifies: joint state, anatomic location, number of joints, joints size.
  - How the patient evolves.
  - Differentiate diseases.
  - Onset type.
  - Inflammatory.
- **Goals:**
  - Affection identification, time reduction, canalization of patient in case of another disease diagnosis.
- **Environmental:**
  - Patient.
  - General Doctor.

Unlike Russel and Norving (1996), in his model *an agent with explicit goals* applied in robotics, software agents type medical diagnosis system, instead of using sensors, software agents are given as percepts (agent receives as perceptions) of input data provided by the general practitioner (symptoms, clinical signs, physical examination, heredity, personal data, evidences and the patient answers) to report the current state of the environment (patient), this information is established a current state on the agent, the agent in this time meets the age, sex, type of disease, whether it is inflammatory or not inflammatory, Identifies: joint state, anatomical region, number of joints, joints size. Which will later be modified to know how will evolve the patient. Immediately questioned actions to take based on a table of terms, when he has taken an action check out the new State in which the patient has evolved to differentiate diseases and turns to question what happens if you take a different action, with this action obtains a goal (identifies the disease). The effectors are the responses of the agent to suggest a group of diseases in hierarchical order which shows the interaction of the user.

Requirements for Software Agents Development

Requirements

It's required, (Pressman, 2002) provide a rheumatic identification by means of an software agents which contains a knowledge base, and the needed algorithms, to make a search process that allows reaching a defined goal.

It's required, create a data bank which contains the information of general data of the patient, age, sex, anatomic location, startup type, inflammatory and so on), clinical signs, symptoms, laboratory analysis.

The human agent (rheumatologist) requirements are: provide the information to elaborate the structured questions, diagnostic approach in rheumatology, the initial evidences, rules and goals to that it is desired to arrive.

The requirements for the computer program are: a user interaction design that allows it to capture the needed information on a structured form, and that shows the same way an answer by means of a text file and images.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Rheumatic diseases that cause inflammation in the joints

1	Gout
2	Infectious arthritis
3	Juvenile arthritis
4	Spondiloarthritis
5	Systemic Lupus Erythematosus
6	Rheumatoid arthritis
7	Degenerative joint disease

Data Analysis

It is necessary to identify and structure the information to facilitate the development of the agent, according to the heuristic of the physician rheumatologist, is described as follows:

The agent identifies a total of 41 diseases, same that are separated into two categories: inflammatory and non inflammatory. Referred to joint.

Diseases that Identifies Software

Inflammatory. See in Table 1, rheumatic diseases that can identify the software.

Description of the Knowledge Base

To search based on goals, we need the Knowledge Base, with well-defined information, established the same, with characteristics of diseases as well as by their categories.

For inflammatory joint disease characteristics are as follows:

1. **Sex:** Sex may refer to:
   a. Male.
   b. Female.
2. **Age:** Inflammatory joints diseases that will be taken into account for the agent are distinguished by a range of age that is presented; these ranges include the following way:
   a. 0 - 12 months or newly born.
   b. 1 - 16.
   c. 17 - 30.
   d. 31-50 years.
   e. 51 years old, or later.

Non inflammatory. See in Table 2, rheumatic diseases that can identify the software.

3. **Onset Type:** The onset type or in other words the way in which the disease occurs in the person concerned may be:
   a. **Acute:** The disease occurs suddenly, in a way not very recurrent.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

   b. **Insidious:** The disease occurs so insistent.
4. **Number of Joints:** The number of joints refers to the total number of joints that are being affected by the disease in the patient taking the following classification.
   a. **Monoarticular:** Disease found only in a joint. I.e. only affects a part of the patient's body.
   b. **Oligoarticular:** Disease is two to four joints. I.e. affects two to four parts of the body of the patient.
   c. **Polyarticular:** The disease is found in more than five joints. I.e. affects more than five parts of the body of the patient.
5. **Affected Region:** The region affected by the patient will be taken in three ways:
   a. Upper limbs.
   b. Lower limbs.
   c. Spine.

In the case of inflammatory diseases they are not given the joint, only the involved region, upper limbs may be the hand without specifying (ulna, radio, Carpus and Metacarpus), illness can involve multiple regions such as it can affect upper and lower limbs or spinal column.

6. **Size of the Joints:** The size of the joints in inflammation may refer to:
   a. **Small:** Metacarpophalangeal, interphalangeal, metatarsophalangeal.
   b. **Great:** Femur, knee, elbows, ankles, shoulders, and wrists.

Here we define the basic features that will be taken into account to enable the agent to the search for inflammatory diseases.

For "non" inflammatory disease characteristics are as follows:

1. **Similar to the Inflammatory:** sex, age, type of home, number of affected joints.
2. **Affected Region:** The region in the patient shall be taken in three ways:
   a. Upper limbs.
   b. Lower limbs.
   c. Spine.
3. **Location:** In the case of "non" inflammatory diseases wanted the part of the body that is affected by the disease, categorized in the following way:
   a. **Upper Limbs:**
      i. Shoulder(s).
      ii. Elbow(s).
      iii. Neck.
      iv. Hand(s).
      v. Wrist(s).
      vi. Finger(s).
   b. **Lower Limbs:**
      i. Foot(s).
      ii. Knee(s).
      iii. Leg(s).
      iv. Hip.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

   v. Ankle(s).
   vi. Thigh(s).
c. **Spine:**
   i. Lumbar.
   ii. Dorsal.
   iii. Cervical.

Some medical symptoms and clinical symptoms are similar in different rheumatic diseases. For this reason, software agents need take several actions to reach a goal.

4. **Clinical Signs:**
   a. Conjunctivitis.
   b. Seizures.
   c. Deformity.
   d. Volume increase of joints.
   e. Digital clubbing.
   f. Lymphadenopathy.
   g. Hair loss.
   h. Cyanosis (hand) or blue hand.
   i. Convulsions (seizure).
   j. Boutonnière Deformity of Finger.
   k. Varus or valgus knee deformity.
   l. Varus or valgus ankle deformity.
   m. Ulnar deviation or ulnar drift.
   n. Cubitus varus (gun stock deformity).
5. **Medical Symptoms:**
   a. Cough.
   b. Fatigue physical.
   c. Chest pain.
   d. Morning stiffness.
   e. Joint tenderness.
   f. Xerostomia(dry mounth).
   g. Dry eyes.
   h. Pyrexia(fever).
   i. Blurred vision.

Design Software Agents

Agent Attributes and Rules of Behavior

Listed in Table 3 are a group of conditions that will allow software agents to take action, which is not enough, therefore need a group of conditions that meet to take the appropriate action is what happens if I take action A? What happens if I take action B? What happens if I take action N…?

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 2. Rheumatic diseases non inflammatory in the joints

1	Deltoid Bursitis	18	Cervical root compression syndrome
2	Meniscus injuries	19	Thoracic outlet syndrome
3	Knee ligaments injuries	20	Carpal tunnel syndrome
4	Chondromalacia patellae	21	Lumbar root compression syndrome
5	Osgood-Schlatter disease	22	sciatic neuritis
6	Prepatellar bursitis	23	Epicondylitis or tennis elbow
7	Anserine bursitis	24	Peroneal syndrome
8	Ankle valgus	25	Growing pains
9	Plantar Fasciitis	26	Fibromyalgia Syndrome
10	Flat feet	27	Multiple myeloma
11	Hallux valgus	28	Febrile diseases
12	rotator cuff Tendinitis	29	Epitrocleitis
13	Back sprains	30	D'Quervain tendinitis
14	Scoliosis	31	Digital stenosis tenosynovitis
15	Spondylosis	32	Dupuytren's Contracture
16	Lumbar disc herniation	33	subtrochanteric Bursitis
17	Vertebral fracture	34	Disorders of the spinal alignment

With the action taken and to meet the rules set out in the knowledge base to identify a disease, the effectors are triggered, showing the interaction of user, the name of the disease, its features and images. For the general practitioner according to its criteria can give its opinion.

In this type of disease clinical signs, symptoms, and affected region, may be similar, which must have clearly defined rules, which are software agents that help identify correctly the condition. Then we'll show some rules to identify the rheumatoid arthritis.

R1: Starts with polyarticular pain.
R2: Volume increase of joints.

Table 3. Identifying inflammatory diseases

Condition	Action
If female	To identify sex
If the age is 7 years	Set to range between 1-16
If the disease is insistent	Sets onset type insidious
If you have an affected joint	Monoarticular
If you have from two to four	Oligoarticular
If you have more than five	Polyarticular
If it occurs in the shoulder	Identifies region affected upper limbs
If it occurs in the knee	Identifies region affected lower limbs

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

R3: It presents between 30 and 50 years old.
R4: It affects the wrists, metacarpophalangeal, interphalangeal joints.
R5: 20% of the cases have extrarticular disease.
R6: It affects shoulders and elbows in more than 50% of cases.
R7: The rheumatoid factor is positive in 70% to 80%.
R8: LC-reactive protein is positive.
R9: X-rays show ankylosis joints.
R10: Its elevated erythrocyte sedimentation and C-reactive protein is positive.

For Non Inflammatory Diseases

The main condition is that the disease is present without joins inflammation, the action taken by the agent is to locate the anatomical region to go to a possible goal, which is simpler to identify because they are very specific conditions of such diseases.

Table 4 shows the agent identifying a condition and taking action to reach a possible goal.

Algorithm

To classify diseases in inflammatory and "not" inflammatory, the search for Software Agents is much simpler, because that some diseases are discarded by groups thus are.

1. **Start:** First the user is asked to identify if the disease is inflammatory.
   a. If it is inflammatory
      i. Go to table of inflammatory diseases.
   b. else
      i. Go to table of Non inflammatory diseases.
2. **Input Percepts:** Calls for the data depending on whether the disease is inflammatory or not
   a. If is inflammatory then:
      i. {Age and sex.
      ii. Onset Type.
      iii. Size of joints.
      iv. Number of joints.
      v. Region affected.
      vi. Signs.
      vii. Symptoms}.
   b. Else "non" inflammatory:
      i. {Age and sex.
      ii. Onset type.
      iii. Region affected.
      iv. Part (s) concerned (s)}.
3. **State Update for Inflammatory Diseases:**
   a. To have entered the age and sex, is looked for in the Data Base diseases that match the age and sex that the user entered and stored. Result1 = Percepts1 ∩ **A,** where are the diseases recorded in the table knowledge base, percepts1 data entered above.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4.

Located	Action	Goal
Upper limbs	Shoulder	Deltoid bursitis
		Rotator cuff Tendinitis
	Elbow	Epicondylitis
		Epitrocleitis
	Hand	D'Quervain tendinitis
		Digital stenosis tenosynovitis
		Dupuytren Contracture
Lower limbs	Hip	Subtrochanteric bursitis
	Knee	Injury of meniscus
		Ligaments injury
		Chondromalacia patellar
		Osgood-Schlatter disease
		prepatellar Bursitis
		Anserine bursitis
	Feet	Ankle valgus
		Plantar Fasciitis
		Flat feet
		Hallux valgus
Back	Spine	Sprain
		Scoliosis
		Spondylosis
		Hernia of intervertebral disc
		Vertebral fracture
General	--	Fibromyalgia Syndrome
		Multiple myeloma
		Febrile diseases

b. To have entered the start type is looked for in the database diseases to coincide with the start type that the user entered and stored. Result2 = percept2 ∩ **A**.
c. To have entered the size of joints diseases that matches the size of joints that the user entered is sought in the database and is stored. Result3 = percept3 ∩ **A**.
d. To have entered the number of joints diseases that matches the number of joints that the user entered is sought in the database and is stored. Result4 = percept4 ∩ **A**.
e. To have entered the affected region seeks a disease that matches the affected region that the user entered in the database and saved. Result5 = percept5 ∩ **A**.
f. To have admitted the sign or signs is sought in the database diseases matching the sign or signs that the user entered and stored. Result6 = percept6 ∩ **A**.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

g. To have admitted the symptom or symptoms database is search diseases that match the symptom or the symptoms that the user entered and stored. Result7 = percept7 ∩ **A**.

Thus the sum of partial results is: Partial Result (PR)

PR = Result1 ∩ Result2 ∩ Result3 Result4 ∩ Result5 ∩ Result7 Result6

4. **Formulate Goal:** With the partial result, it obtained a first goal, performs a preliminary identification, at the moment the software agents known to the goal to be achieved, in other words, so far has identified a group of diseases that share common characteristics.
5. **Formulate Problem:** The problem to be solved is to select only one of the Groups of diseases, whereas the rules established for each of the diseases, we find differences that allow us to establish a new state that leads to a new target.
6. **Plan:** We will establish a plan to select one of the diseases of the identified group. What happens if I take action A? Or, if I take action B?

If taken to action leads to a goal to which we do not know if it is correct. This forces us to create a search-space. (Nilson, 2004)

7. **Search:**

```
while (plan not empty) do {
if (agree state == true) then{
If there are similarities with the provided data (state) establishes an action
leading to a recommendation, by selecting an option in the space of search
(search-space), considering the rules that are specific to each disease (these
are those that allow us to do a proper identification). Emits a response
through the effectors (user interaction) to the environment (general doctor,
patient) environment
If one accepts the answer ends.
}
Else {
If does not return to do a new search
It establishes a new plan calling for new data (creates a new state) to have
options to generate new actions that would lead us to a new target from the
selected groups of diseases (remainder), in every new search, it will decrease
the (remainder) until exhausted. }
Output Action
If there are no more actions to take, displays message that does not have suf-
ficient information to identify a disease of this kind.
end
```


EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Description of Medical Identification System of Rheumatic Diseases (MISRheD)

Clinical Method

In order MISRheD can correctly locate a disease, it is necessary that the information is provided by a general doctor, due to the account with the basic knowledge of medicine. Another person without this knowledge can hardly reach the desired goal.

First there is an interaction between the doctor and the patient to collect information to determine the clinical semiology considering the following factors:

- **Background Inherited Family:** Diseases of the connective tissue, urinary lithiasis, hyperuricemia, psoriasis, uveitis.
- **Particular Personal History:** Urinary lithiasis, hyperuricemia, adverse obstetric events, thrombotic, skin diseases, thyroid disease, uveitis events.
- **Evolution:** Episodic, intermittent, continuous, all types.
- **Joints Physical Examination:** Complete, orderly, symmetric, comparative.
- **Physical Examination:** Deformity, increased volume, functional limitation, snapping, painful points.

Below are the basic information needed from MISRheD to identify the following diseases.

Procedure for Identification of: Rheumatoid Arthritis

- **Female sex:**
  - Age 42 years
  - Inflammatory classification
  - Onset type: insidious
  - Number of joints 5
  - Region affected upper limb
  - Size of small joints
  - Radiological study erosion articulate with decrease of space joint
  - Study laboratory Hematology thrombocytosis, rheumatoid factor immunological studies 75%
  - Procedure for identification: spondiloarthritides
- **Male:**
  - 17 Years old
  - Inflammatory classification
  - Onset type: insidious
  - Number of joints 2
  - Region affected lower limb
  - Size of large joints
  - Radiological study of pelvis: sacroiliitis
  - Laboratory study: histocompatibility HLA-B27 antigen
  - Uveitis

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Procedure for identification of: Gout

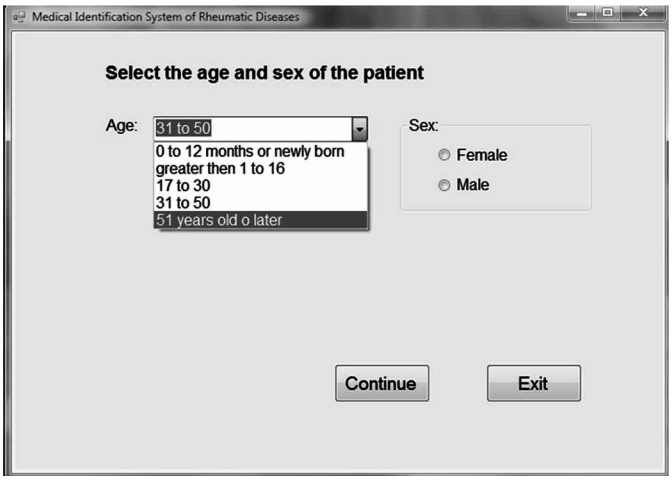
- **Male:**
  - Age 47 years
  - Inflammatory classification
  - onset type: Acute
  - Number of joints 1
  - Region affected lower limb
  - Size of small joints
  - Radiological study: without bone injury
  - Laboratory study: hyperuricemia
  - Previous urinary lithiasis

User Interaction

Description of the Templates for the Capture of Data Description Templates

The template in Figure 1 shows the capture of data to select age and sex of the patient. Figure 2 captures data on the number of damaged joints. The template in Figure 3 captures clinical signs of the patient. Figure 4 captures symptoms of the patient. In the template shown in Figure 5, the general practitioner can select the affected body region.

Figure 1. Templates to select age and sex of the patient



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Templates to select the number of joints

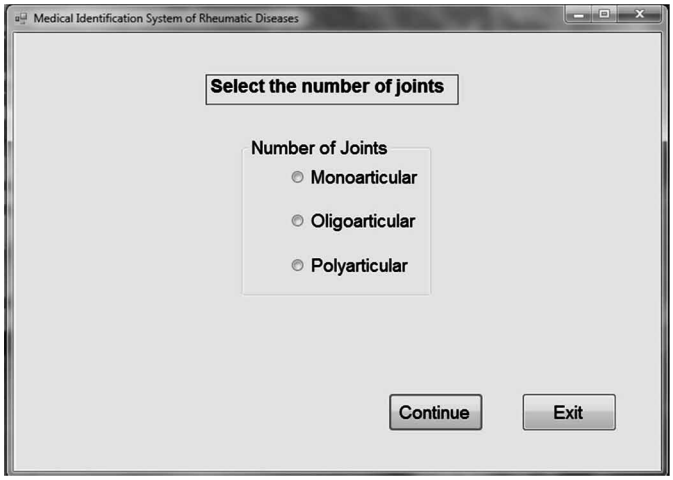
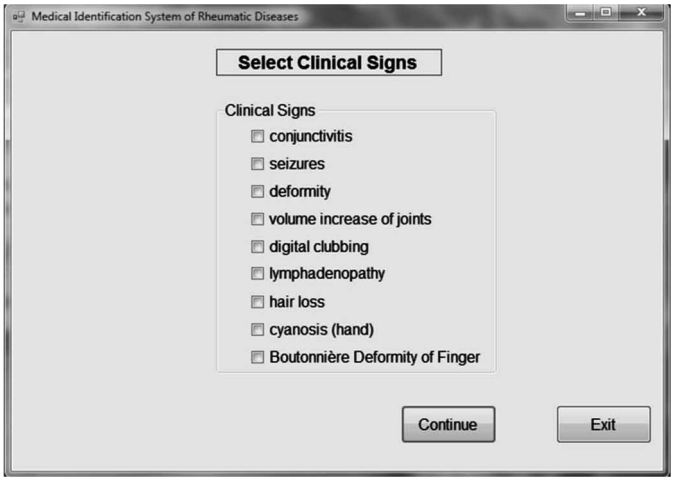


Figure 3. Templates to select clinical signs



SOLUTIONS AND RECOMMENDATIONS

Information to the Population and General Practitioners

Information campaigns and efforts from Government employ an appropriate strategy to the culture of the population. For example in our country, Mexico, the majority of the population is not accustomed to read texts nor to attend talks on cultural themes being these media the principal for disseminating information, in addition there are no records of formal studies of the results and impact obtained by these means of dissemination.

General practitioners who have years on practicing commonly received little information on rheumatic diseases, there is an especial need to offer them courses through the medical institution to which they belong. In this way it can improve its approach to identify a rheumatic condition to ensure that the

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 4. Templates to select symptoms

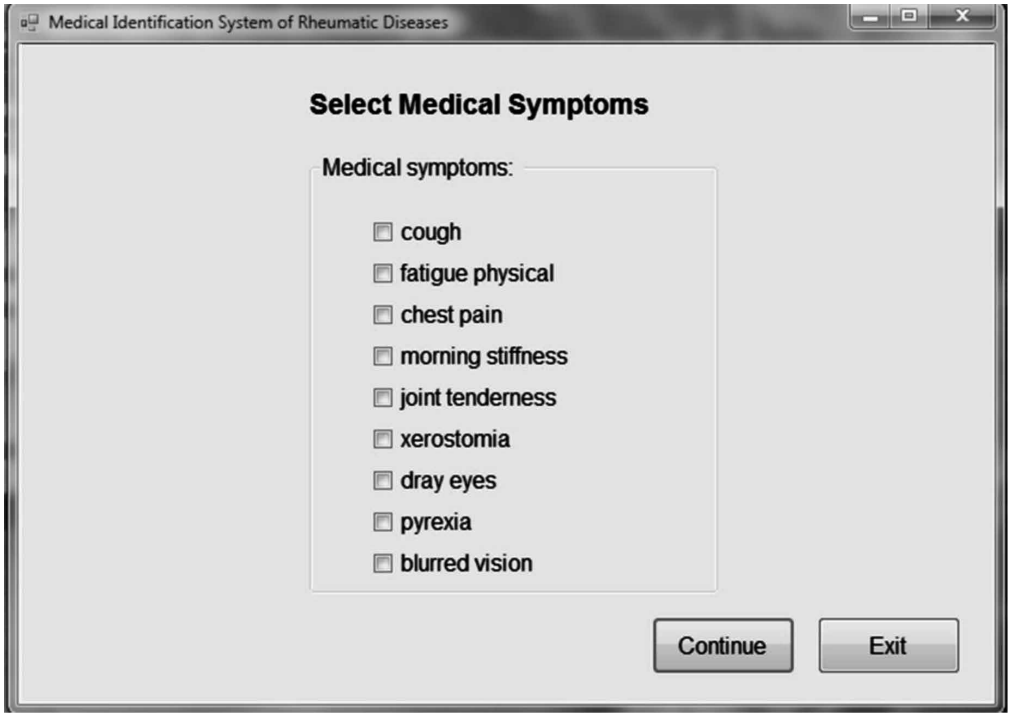


Figure 5. Templates to select region of the body



patient receives timely specialized medical attention. As mentioned above the general medical depends on the detection and prompt attention, which has become a vital part to address the problem in time.

Modifying the culture of the population in terms of popular beliefs of the rheumatic diseases are cured with home remedies, also such diseases are not curable, only controllable. To alleviate ailments people fall into apathy and let the time pass to receive adequate professional attention.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

On the economic side, companies are affected by issued labor disability and pensions granted to its staff, and that is why they should collaborate with the public health institutions requesting that the doctors could detect in time and control these diseases, to avoid large amount of costs that may be unnecessary.

In the social aspect personal disability causing these diseases, becomes a serious public problem that has been increasing in recent years, when a person is affected during his productive age affects also greatly to the family. Give people in productive age alternative effective for the control of the disease, so that in this way they prolong their lives economically active.

In Educational Institutions

Medical schools promote among their students the training of medical specialists in rheumatology to meet the demand of the population and to comply with the ratio of inhabitant/doctor (50/1) that in our country is not enough, in relation to Latin American countries such as Cuba (170/1).

The medical students in their courses receive information and expertise in the specialty of Rheumatology. Encourage the use of new technologies as alternative didactics, in our country this is an educational problem for the lack of resources and training to the medical teaching in the application of teaching tools.

In computer schools that offer courses on information technology, computer and related, promote the development of applications that focus to the medical area to address some of the demanding problems in our country in students and teachers. Projects are being implemented in conjunction with specialists from the sector health.

Physician Rheumatologist

Our country usually the physician rheumatologist provides medical care in his Office and very few are involved in research projects.

It is recommended that they are supported by government agencies or of private initiative for research activities. As mentioned in previous subjects nowadays still much is unknown with what respect to rheumatic diseases. On the other hand medical specialists are encouraged to increase the participation in events that allow the publication of information in books of texts, journals, magazines, proceedings, etc.

FUTURE RESEARCH DIRECTIONS

Perspective

In the future the new trends are to deploy applications using new information technologies in the medical sector to enable decision-making based on the knowledge of specialists and not only information, but complementing information and knowledge to establish a process of reasoning that leads to a certain diagnosis, we know that it is a major challenge for developing countries, because general practitioners do not have the culture of the use of tools of this type and on the other hand the conjunction medical specialist and ICT sector is difficult, because the physician has no researcher training. In the future, the narrow health gap and ICT should be closed.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The social aspect has lagged behind in the development of the technology because it has focused on productivity, development of quality and economic improvements with cost-cutting in enterprises and the labor and productive sector.

From the technological aspect, it is not enough to develop applications to solve economic or business issues, the social aspect is very important, we know that public demand increasingly focuses on a better quality of life and better health services. In this chapter, we propose a way of decreasing the costs involved in attending to an employee with a rheumatic disease, (payments by inability, absenteeism, and lower productivity) that may temporarily or permanently disabled or improve care to an older adult, and as mentioned that these diseases affect a large population of this type. Using MISRA we propose new solutions based on technologies that will improve social and health services to treat a patient in time and take effective control of their disease.

This is neither the only solution nor the only tool, but new applications must be developed where The ICT sector is highly present in this context.

Viability

The viability of this proposal is feasible from the economic point of view, considering that the majority of medical institutions have computer equipment, for the implementation of MISRA requirements is minimal. The technical aspect was discussed in the model and how much the attitude of the physician to accept this proposal was discussed in the paradigm.

Model

On the way in which we are proposing the use of the model goal-based agents related to the medical area, can be very useful for students of computer science with application in artificial intelligence that helps to understand the algorithm and the structure of software agents and allows them later to develop such applications. For the medical students it can show them how the ICT can help them in their professional development. Use of ICT in their professional development, as well as software developers are to be considered an alternative resource for medical researchers.

Paradigm

In our country along the lines of others, to break paradigms and traditions is very difficult, the predominant practice of medicine used to make diagnoses in the presence of the patient, through questioning and exploration complemented by studies of laboratory and Cabinet, using clinical semiology traditional. Not relies on the computing resource, whether an expert system, computerized analysis of images, etc. because there isn't this resource implemented in our public health system. Public health institutions do not have this medium. The technological advancement of today's world requires that all areas of human activity are updated and this includes the medical work. It is desirable that the health institutions have computer hardware in their areas of medical care and that also to have computer tools to facilitate the diagnosis of common and frequent diseases in the various areas of medicine.

Some public institutions of health and in the majority of private institutions electronic formats are used to take control of the record of the patient, consultation to pharmaceutical dictionaries, existence of medicines at the pharmacy of the institution itself, but not on computer auxiliary diagnosis processes.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

This tells us that in the future the general practitioner can accept computer programs to improve their job performance and to enable them to make a reliable diagnosis with high expectations of accepting a new paradigm.

Future Research

A proposal for future research opportunities using software agent in medicine is the application:

Dermatology

The proposed algorithm can be applied wit the next similarity; it's required to create a data bank, which contains the information of general data of the patient (age, gender), visual data (rash, grain, color, anatomic location, size, shape and so on), clinical signs, symptoms, and descriptive images of several sufferings. The human agent (dermatologist) requirements are: provide the information to elaborate the structured questions, the initial evidences, rules and goals to that it is desired to arrive.

The software agent is required of a tables group, as they are enlisted next:

- Table of questions,
- Table of initial evidences,
- Table of conditions,
- Table of goals,
- Table of rules,
- Table of actions,
- Table that shows the pre-diagnosis.

The difference for these software agents is the use of a Bank of images for general practitioner so they can compare the injury of the patient with the disease that is shown in the image and software agents compare these features to perform the search.

Physical Medicine

The goal of software agents is to suggest to the general practitioner when he should send the patient for attention by the specialist in physical medicine and to receive physical rehabilitation therapy. Physical rehabilitation is considered as alternative medicine, for this reason not all doctors resort to it. The areas of medicine that frequently require management by physical medicine as support are Traumatology and orthopedics, Neurology and Rheumatology, however other areas can benefit from its use. An algorithm with the characteristics of the disease that afflicts the patient can achieve a timely delivery to a rehabilitative treatment to enable the recovery of a patient.

Similar to the previous proposal, apply the same algorithm with appropriate amendments, in this case the treatment received by the patient (cold, ultrasound, heat, infrared, hydrotherapy, physical exercise) that the patient needs is going according to the conditions of the joints or the affected part of the body.

The development of a medical diagnosis requires the mental processing of data provided by the patient, the exploration the doctor will take and support studies. For this reason any area of medicine could benefit from the creation of tools similar to which is intending in this chapter.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

The use of CIT as a tool to solve social and health problems, leads to the following:

1. Create multidisciplinary Workgroups. General physician, Rheumatologist, CIT's engineering, students that still are in the formative stages of engineering.
2. The project development and its completion through MISRheD, did not require many financial resources, most valuable resource was the human resource.
3. Takes time to develop such tools, for approximately two years we collected information from the experience of the rheumatologist, with approximately thirty years of professional experience.
4. The shortage of rheumatologists at our country indicates that few people have the knowledge and experience in this specialty, moreover, is a medical area where still much is unknown, MISRheD is a way of preserving part of this experience and transmitting it to the public health community.
5. There is a large backlog in the use of CIT in the institutions of public health in our country; there is the possibility that MISRheD will support telemedicine to apply it in places where there are no rheumatologists.
   The algorithm may be useful for general practitioners, professors, and computer science students, as an assistance that would help them to solve similar problems
6. This tool is an effort between researchers, doctors and students to improve the patients in health care, which are invited to participate in a future in collective projects to solve society's problems.
7. Antirheumatic drugs are needed to control them, such as: Nonsteroidal Anti-Inflammatory Drugs, Corticosteroids, Disease Modifying Anti-Rheumatic Drugs, Slow-Acting Antirheumatic Drugs and Biological Therapy.

REFERENCES

Bertalanffy, L. (2011). Teoría General de los Sistemas. Ed. Fondo de Cultura y Económica.

Braun, J., & Sieper, J. (2004). Biological therapies in the spondyloarthritidesthe current state. *Oxford Journals Medicine & Health Rheumatology, 43*(9). Retrieved august 31, 2014, from http://rheumatology.oxfordjournals.org/content/43/9/1072.full.pdf+html

Cañete, J. (2008). *Manual S.E.R. de las Enfermedades Reumáticas. 3a edición. Sociedad Española de Reumatología*. Editorial Médica Panamericana.

Carbonel, J. (2006). *Semiología de las enfermedades reumáticas*. Madrid: Editorial Médica Panamericana.

Clauw Daniel, J. (2005). *Fibromyalgia and Other Central Pain Syndromes* (1st ed.). Lippincott Williams and Wilkins.

De León, M. (2012). El aprendizaje de las enfermedades reumáticas desde una perspectiva tecnológica. *Revista cubana de Reumatología, 15*(19).

Horn, W. (1999). *Artificial intelligence in medicine.* Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making, AIMDM'99, Aalborg, Denmark.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kats, W. A. (1977). *Rheumatic Diseases: diagnosis and management*. Lippincott.

Kolkata, A. (2012). *Biologics in Rheumatology 2012. Medicine Update 2012 Vol. 22*. Retrieved august 30, 2014, from http://www.apiindia.org/pdf/medicine_update_2012/rheumatology_05.pdf

Lee, C., & Wang, M. (2006). *An Ontology Based Intelligent Agent for Respiratory Waveform Classification.* Advances in applied artificial intelligence: 19th international conference on industrial and engineering applications of artificial intelligence and expert systems, IEA/AIE, Annecy, France. 10.1007/11779568_131

Martins, V. (2010). *Validation of an educational game for teaching cardiovascular assessment*. Investigación y Educación en Enfermería Colombia.

Medical Dictionary. (n.d.). Retrieved august 12, 2014, from http://medicaldictionary.thefreedictionary.com/Antirheumatic+Drugs

Monteagudo, P. (2003). *Software educativo para la enseñanza de la semiología clínica del sistema respiratorio*. (Master's Thesis). Ciudad de La Habana: ISCM de La HabanaCentro de Cibernética Aplicada a la Medicina (CECAM).

Nilson Nils, J. (2004). *Inteligencia artificial una nueva síntesis*. Mc GrawHill.

Noachtar, S., & Peters, A. (2009). *Semiology of epileptic seizures: A critical review*. University on Munich Germany.

Parapar. (2007). Manifestaciones oculares de algunas enfermedades reumáticas en el niño. *Hosp. Pedro Borras*. Retrieved April 12, 2014, from http://www.portalesmedicos.com/publicaciones/articles/634/1/Manifestacionesocularesdealgunasenfermedadesreumaticasenelnino.html

Pressman, R. S. (2002). *Ingeniería del Software. Un enfoque práctico*. Madrid: McGrawHill.

Rejón, C. (2013). *Logic structure of clinical judgment and its relation to medical and psychiatric semiology.* U.S.A. Pub med / U.S. National Library of Medicine.

Resconi, G., & Jain, L. (2004). *Intelligent agents Theory and applications*. Springer Verlag.

Research Unit of the Spanish Society of Rheumatology. (2013). Spanish registry of adverse events of biological therapies in rheumatic diseases. *Biobadaser*. Retrieved August 12, 2014, from http://biobadaser.ser.es/biobadaser/eng/docs/SER_informe_web2012_ENG.pdf

Rich, E., & Knight, K. (1991). *Artificial Intelligence*. Mc GrawHill.

Russell, S., & Norving, P. (1996). *Inteligencia artificial un enfoque moderno*. México: Prentice Hall Hispanoamericana, S.A.

Salter, B. (n.d.). *Nonarticular Rheumatism in Textbook of Disorders and Injuries of the Musculoskeletal System* (3rd ed.). Lippincott Williams.

Saravolats, L. (2007). *The role of Medical Semiology in clinical infectious diseases: Back to the basic*. Lippincott Williams & Wilkins, Inc.

Schumacher Ralph, H., Klippel John, H., & Robinson Dwight, R. (1988). *Principios de las Enfermedades Reumáticas (12th ed.)*. Arthritis Foundation.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Šenolta, L., Vencovskýa, J., Pavelkaa, K., Ospeltb, C., & Gayb, S. (2009). Prospective new biological therapies for rheumatoid arthritis. *Autoimmunity Reviews, 9*(2).

Stone, J. (2009). *Pearls and Myths in Rheumatology*. London: Springer Dordrecht Heidelberg.

Taranta, A., & Markowitz, M. (1981). *Rheumatic fever: a guide to its recognition, prevention, and cure*. MTP Press.

Wright, V., & Harvey, A. (1988). *Ilustraciones diagnósticas en reumatología: Pruebas de autoevaluación*. Interamericana McGraw Hill.

ADDITIONAL READING

Bichindaritz, I. (2010). *Computational Intelligence in Healthcare*. Springer Verlag. doi:10.1007/978-3-642-14464-6

Brahnam, S., & Jain, L. (2001). *Advanced computational intelligence paradigms in healthcare. Virtual reality in psychotherapy, rehabilitation, and assessment*. Berlin, Heidelberg: Springer.

Brooks, F. P. (1995). *The Mythical manmonth, Essays on Software Engineering*. AddisonWesley Pub Co.

Combi, C., Shahar, Y., & AbuHanna, A. (2009). *Artificial intelligence in medicine*: 12th Conference on Artificial Intelligence in Medicine, AIME 2009, Verona, Italy, July 1822, 2009: proceedings. Berlin; New York: Springer.

Dar E. and Ghani M. (2010). Utility Based Agent for Test Paper Generation. International Journal of Control, Automation, and Systems Vol. 1, No. 2, June 2003 ACM Press.

Englebardt, S., & Nelson, R. (2002). *Health care informatics: an interdisciplinary approach*. USA: Mosby.

Håkansson, A. (2009). *Agent and multiagent systems: technologies and applications*: third KES international symposium, KESAMSTA 2009, Uppsala, Sweden, June 35, 2009: proceedings. Berlin: Springer.

Hakim, L. (2007). *Web mobilebased applications for healthcare management*. Hershey, Pa., USA: IGI Global. doi:10.4018/978-1-59140-658-7

Kim D. (2003). *Modeling and Design of Intelligent Agent System.* International Journal of Control, Automation, and Systems Vol. 1, No. 2, June 2003.

Lim, C. T., & Goh, J. C. (2009). 13th International Conference on Biomedical Engineering: ICBME 2008 36 December 2008, Singapore. Berlin: Springer.

McCarty D. & Koopman W. (1993). Arthritis and allied conditions: a textbook of rheumatology Lea & Febiger.

Moreno, A., & Nealon, J. L. (2003). *Applications of software agent technology in the health care domain*. Basel; Boston USA: Birkhauser Verlag. doi:10.1007/978-3-0348-7976-7

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Müller, J., Singh, M., & Wooldrige, M. (2000). *Intelligent agents: agent theories, architectures, and languages*. Berlin: Springer.

Padgham, L., & Winikoff, M. (2004). *Developing intelligent agent systems: a practical guide*. Chichester: Wiley. doi:10.1002/0470861223

Resconi, G., & Jain, L. (2004). *Intelligent agents: theory and applications Berlin*. Springer. doi:10.1007/978-3-540-44401-5

Riaño, D. (2008). *Knowledge management for health care procedures: from knowledge to global care, AIME 2007 Workshop K4CARE 2007, Amsterdam, The Netherlands, July 7, 2007*. Berlin, New York: Springer. doi:10.1007/978-3-540-78624-5

Salter, R. B. (1999). *Textbook of disorders and injuries of the musculoskeletal system: an introduction to orthopaedics, fractures and joint injuries, rheumatology, metabolic bone disease, and rehabilitation*. Philadelphia: Lippincott Williams & Wilkins.

Sarker, R., & Ray, T. (2010). *Agentbased evolutionary search*. Berlin: SpringerVerlag. doi:10.1007/978-3-642-13425-8

Silverman, B. G. (2005). *Intelligent paradigms for healthcare enterprises: systems thinking*. Berlin, New York: Springer. doi:10.1007/b99809

Sordo, M. (2010). *Advanced computational intelligence paradigms in healthcare*. Berlin, Heidelberg: Springer.

Tan, J. (2005). *Ehealth care information systems: an introduction for students and professionals*. San Francisco: JosseyBass.

Tsihrintzis, G. A. (2010). *Intelligent interactive multimedia systems and services*. Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-14619-0

Wickramasinghe, N., Gupta, J., & Sharma, S. (2005). *Creating knowledgebased healthcare organizations*. Hershey Penns, USA: Idea Group Pub. doi:10.4018/978-1-59140-459-0

Yosida, H. (2007). *Advanced computational intelligence paradigms in healthcare. BerlinNew York*. Springer. doi:10.1007/978-3-540-47527-9

Yuan, S., & Yokoo, M. (2001). *Intelligent agents: specification, modeling, and applications*. Berlin: Springer. doi:10.1007/3-540-44637-0

KEY TERMS AND DEFINITIONS

Agents Based on Goals: The goals are to itemize the situations to that it is desired to arrive. The agent program can combine goals with actions and in this way, be able to choose those actions that will allow reaching the goal.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Antirheumatic Drugs: Antirheumatic drugs are drugs used to treat rheumatic diseases. The major classes of antirheumatic drugs include Nonsteroidal Anti-Inflammatory Drugs (NSAIDs), Corticosteroids, Disease Modifying Anti-Rheumatic Drugs (DMARDs), Slow-Acting Antirheumatic Drugs (SAARDs).

Biological Therapy: Treatment to stimulate or restore the ability of the immune (defense) system to fight infection and disease.

Clinical Semiology: In medicine, the clinical semiology is the body of knowledge that deals with the identification of the various pathological manifestations (symptoms and signs) and compiling them into syndromes and how to interpret.

Disease Modifying Anti-Rheumatic Drugs (DMARDs): DMARDs influence the disease process itself and do not only treat symptoms, hence their name. DMARDs also have anti-inflammatory effects, and most were borrowed from the treatment of other diseases, such as cancer and malaria.

Formulate Goal: To set a model for an intelligent agent due to define to achieved the goal, generating new states and testing them against a goal.

Inflammatory: Term to differentiate the types of rheumatic diseases. This is an inflammatory process that occurs in joints. Regardless of nerves or ligaments as prepatellar bursitis, it is the inflammation of the bursa that is positioned in front of the kneecap (patella).

Medical Signs: Term to identify the types of rheumatic diseases. The physician makes his judgments from the information that he gathers during a physical examination of a patient (conjunctivitis, seizures, deformity, volume increase of joints, digital clubbing, among others).

Medical Symptoms: Term to identify the types of rheumatic diseases. They are abnormal perception of a patient (cough, fatigue physical), chest pain, morning stiffness, joint tenderness, etc.).

Nonsteroidal Anti-Inflammatory Drugs (NSAIDs): Drugs belonging to this class bring symptomatic relief of both inflammation and pain, but have a limited effect on the progressive bone and cartilage loss associated with rheumatoid arthritis. They act by slowing the body's production of prostaglandins.

Recommendation: Filtered information is displayed to the user in order make decisions. Software agents suggest rheumatic disease.

Remainder: Selected information group, which matches the search. Rule out disease that does not match your search.

Rheumatic Diseases: They are systemic diseases of unknown origin that happen when the immune system goes haywire. The immune system attacks the linings of joints, ligaments, tendons, muscles, skin, eyes, cause inflammation, swelling, and pain. There are more than 100 rheumatic diseases.

Search Space: The group of records in a database that were selected according to the search that matches the goal that you want to reach (identify rheumatic diseases).

Slow-Acting Antirheumatic Drugs (SAARDs): A special class of DMARDs and the effect of these drugs is slow acting and not so quickly apparent as that of the NSAIDs.

This research was previously published in Advancing Pharmaceutical Processes and Tools for Improved Health Outcomes edited by Tagelsir Mohamed Gasmelseid; pages 136-166; copyright year 2016 by Medical Information Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 74
Social Implications of Big Data and Fog Computing

Jeremy Horne
Inventors Assistance League, USA

ABSTRACT

In the last half century, we have gone from storing data on 5¼ inch floppy diskettes to the cloud and now use fog computing. But one should ask why so much data is being collected. Part of the answer is simple in light of scientific projects, but why is there so much data on us? Then, we ask about its "interface" through fog computing. Such questions prompt this article on the philosophy of big data and fog computing. After some background on definitions, origins and contemporary applications, the main discussion begins with thinking about modern data collection, management, and applications from a complexity standpoint. Big data is turned into knowledge, but knowledge is extrapolated from the past and used to manage the future. Yet it is questionable whether humans have the capacity to manage contemporary technological and social complexity evidenced by our world in crisis and possibly on the brink of extinction. Such calls for a new way of studying societies from a scientific point of view. We are at the center of the observation from which big data emerge and are manipulated, the overall human project being not only to create an artificial brain with an attendant mind, but a society that might be able to survive what "natural" humans cannot.

INTRODUCTION

Fog computing is another link in the chain of information storage extending back from storing data in a handwritten manner to adding machines, vacuum tubes, and floppy diskettes. One creates the information, stores it, and then distributes it. It is easy to confabulate the motivation for its instrumentation, the "why" of the foundation (big data) with how we get it (storage and distribution methods). Two aspects characterize information: its growth and centralization. Fog computing has enabled its distribution not only to computers to billions of devices in a network we know of as the "Internet of things" (IoT), ranging from printers and phones to "smart appliances", like refrigerators and personal security systems. In

DOI: 10.4018/978-1-7998-0951-7.ch074

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

essence, anything that can house a computer chip is a candidate for this network. Accordingly, the scope of information has increased. In keeping with maintaining clarity between the storage of information and its rationale with its distribution, we will first look at big data in depth and its implication and then consider its mode of distribution (fog computing) in a similar light. A prefatory note about the origin of the present article is in order.

You are reading "Social Implications of Big Data and Fog Computing", which is a re-write and updating of "Visualizing Big Data from a Philosophical Perspective" that appeared in the 2018, *Handbook of Research on Big Data Storage and Visualization Techniques* (Segall and Cook, 2018, pp. 809-852). Most of the discussion was not simply of techniques but about the thinking underscoring why big data should be generated at all. Inasmuch as the current paper was invited, it can be assumed that someone "up there" thought that asking the "why" is important. The present article goes further and fits the "why" into a larger framework of how we regard societies and the challenges they are facing in an ever-increasing complex environment. As to the visualization, there are just as many ways of displaying data as there are projects and methods for creating it, as will be seen below.

Most of the technical aspects of the original work have not changed, and I will leave it to others to banter back and forth on all that updated "techie stuff". One can only imagine the further technological development that can and often has occurred during the last one or two years in which many of the sources used for researching the first article appeared. Too, the visualization aspect will not be emphasized here, because one only need read the original article for that information. Typing in "big data visualization" and related phrases will produce hundreds, if not thousands of images, each with its own purpose and technique. This is a trivial exercise without much intellectual content. What has been lacking in all these discussions about big data is similar to what was lacking and as expressed by J. Robert Oppenheimer in the development of the atomic bomb – the social implications of the technology.

There has been another addition to the previous article in the technological aspect, however, that does strengthen the sociological observations: the further development – fog computing - that is closely analogous to a neurophysiology and the argument made in the previous article. As in societies being at least analogously organic, there is a corresponding aspect of organicity in the technology, and that is beginning to assume literally a life of its own. But, we are getting ahead of ourselves a bit.

It is likely that each one of us has encountered some aspect of big data, it often being a buzz word. With fog computing as soon as one picks up a cell phone, works on a computer, or even opens the refrigerator door s/he in effect is one of the tentacles of a gigantic organism with emergent properties that we do not seem able to predict or control. Whenever we interact with a government agency, go to school, shop or use a computer, data is collected with or without our consent. As recent as 20 May 2018 controversy is swirling about Facebook's allowing its guard to be down so that the British company Cambridge Analytica could acquire through the popular social media outlet Facebook and mine its data to promote Donald Trump's U.S. presidential campaign. Insofar as privacy is concerned, if the information is in cyberspace, then logically it is accessible, the only secret in existence being the deodorant. The trail from the origins of data as representations of information ages ago to its present uses and abuses is long. While we marvel at the technology, it is critical to understand why the technology was created in the first place.

There is data (big or otherwise) and then there is its use and fog computing distribution. One should fathom the implications of the first to appreciate what happens with the latter. Otherwise stated and amplified, what applies to the first applies also to the second.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

This article will touch on some of the technical aspects of big data and fog computing, what is collected (as in National Security Agency's bulk data collection programme), and the limits of collecting large quantities of data. Overall, if something can be measured, data can be collected be it the length of every spider's leg, width of every tree or weight of every grain of sand. Yet, there is no dearth of facts but without organization and context there is little or no meaning. An overarching consideration is whether it should be collected at all.

Reactions to technology often are like this where people learn of the cleverness being developed, the "wow", "awesome", and glitz of the gadgets. Big data is no exception as in the wide-eyed reaction to being told of huge data storage capacities and now the ability of persons to access that data directly or through their Internet of things (IOT) devices with fog computing. Older generations accustomed to paper-based card catalog systems in libraries especially are taken aback in learning the big data of the whole of the Library of Congress in the very near future (and even now with 20 ten terabyte (TB) hard drives) can be sitting on their desk at home or on their laptop.

Facts are meaningless in themselves and they need to be connected to each other. Even if a person relates the facts to each other still there is no final coherence unless there is a framework for that coherence, i.e., a theme. There are two basic ways of remembering a situation: rote and as a series of connected events or motif. Describing the beautiful setting of the Grand Canyon including the ecology, geological history and so forth is further impressed upon us by introducing something closer and to which we can translate as a relevant experience the way someone like ourselves interacted with that environment as a story. So, one tells of a story of someone in the Grand Canyon invoking the description of all the things the narrator would like for the audience to remember - the environment, history and so forth. This is the way knowledge was passed down throughout prehistoric times by tales, proverbs, myths and other oral communications (Eliade, 1954). Here, a story is told about the development of big data, that story being analogous to a parable conveying meaning of the assemblage of descriptions about that big data.

Contemporary events may seem to be enveloping us, the proverbial tail wagging the dog. One needs to point only to global warming, exploding populations, mounting political and religious violence, and economic failures. There are luminaries who doubt the capacity of this species even to survive the next 100 years (Firth, 2015; Rees, 2008; Sheldrick, 2014; Carey, 2014). We read of events daily often noting the facts but failing to apprehend the context.

To provide meaning, one has to know what gives rise to the data in the first place including the rationale for it, how it is collected and how it is processed. Science gives the rationale for technology, theory does it for practice and overall the "why" is necessary for the "what" of anything. Besides the ways we think, the "why" of thinking is the subject of philosophy. We are also looking for an explanatory model of events that is consistent. There are immediate purposes for collecting data such as for doing scientific experiments and financial transactions, but one should start wondering about the reasons large volumes of data are collected on persons. Science exists because one wants to control the future, but this is done by understanding the past and such involves measurement and attendant data collection followed by analysis. Big data on us often is collected systematically and a case can be made that it is done for the same reason. Fog computing details how penetrating that data can be.

In many cases it seems that science and the technology emanating from it have been sufficient to solve problems where antibiotics in medicine, phones and computers in communications and automobiles in transportation are obvious examples. In each of these areas we can identify big data sets. There are huge research projects like the Dark Energy Survey which uses a 570-megapixel camera and the 4-meter Blanco telescope at the Cerro Tololo Inter-American Observatory in Chile (Wikipedia, 2018h). Genome

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

sequencing, modeling chemical interactions, analysis of social systems and ecological studies are others. Meteorological agencies are attempting to capture meaning through thermodynamic models of weather systems. Then, there is big data captured in sociological analysis. Searching "sources of big data" and similar phrases on the Internet will produce hundreds of websites and images displaying various projects collecting, analyzing and displaying big data.

The capacity to collect data for all intents and purposes is limitless but its organization, analysis and application are the barriers to its utility. Humans have limitations to perform acts as in memorizing and seeing great distances and have devised tools like computers and telescopes to assist. So, it is with intellectual limitations, creating devices like computers to compensate. Big data may be thought of as a tool, the organization as the way of using it. Whether humans have the native abilities (read "intelligence") to manage the complexity is debatable, a factor contributing to the explanation of why so much research is being done in Artificial Intelligence, particularly with the announcement of the Decade of the Brain (Brain Project, 2015). With its 10 billion synapses (Synapses, 2015; Herculano-Houzel, 2012), the brain and associated consciousness arguably are greatest challenges to those collecting and managing big data. Yet, in considering that we live in a society, one asks about society being an entity of its own a human unto itself. Not only must the human brain encounter challenges of an increasingly complex world but we as a species are in the same situation. Think about this while reading the following philosophical account of big data. Consider that the ultimate visualization of big data is looking at ourselves in the mirror. We will first look at "big data" first and its implications, and then we will examine its distribution and use with fog computing.

BACKGROUND OF BIG DATA AND FOG COMPUTING

What Is/Are Data?

A systematic way of understanding the concept being conveyed by a word is to analyze the meaning of the word itself. Origins and context are a good starting point. The etymology of data is:

…1640s, plural of datum from Latin datum "(thing) given," neuter past participle of dare "to give" (see date (n.1)). Meaning "transmittable and storable computer information" first recorded 1946. Data processing is from 1954. (Data Etymology, 2018)

As a definition, there are these examples as the public first encounters the term:

Noun:

1. A plural of datum.
2. (Used with a plural verb) individual facts, statistics or items of information: These data represent the results of our analyses. Data are entered by terminal for immediate processing by the computer.
3. (Used with a singular verb) a body of facts; information: Additional data is available from the president of the firm. (Data Dictionary dot com, 2018)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

1. Factual information (as measurements or statistics) used as a basis for reasoning, discussion or calculation
2. Information output by a sensing device or organ that includes both useful and irrelevant or redundant information and must be processed to be meaningful
3. Information in numerical form that can be digitally transmitted or processed (Data Dictionary Webster, 2018)

Facts and statistics collected together for reference or analysis. See also datum. (Data Dictionary Oxford, 2018.)

From the basic definition of data and its etymology, one can make some immediate observations.

1. A single piece of information, as a fact, statistic or code; an item of data. (Datum Dictionary dot com, 2015)

Webster's says, "noun plural but singular or plural in construction, data often attributive" (Datum Dictionary Webster, 2015). There is "A piece of information. See also data." (Datum Dictionary Oxford, 2015)

One should note that the technical term is "datum", the singular of data if one is to adhere to the Latin correctly. We will refer to "data" but what is said here applies to the singular "datum". Words common to these are fact, information and statistics. However, as we will see in a moment, the three are not synonyms for what initially is collected but results. One could take a photograph of anything and call that "data". However, we are the ones who give it meaning such as in a fact, information or statistic. The meow of a cat, the ripples on a pond, the smell of a rose, the feel of a porcupine's quill, and the taste of a chocolate bar all are data. The five senses feed to us the reflections of the phenomena of our environment in what we may call "big data."

The term "big data", itself, seems to assume a childlike perspective, a first grade word, "big" being juxtaposed to a much more complex term "data" (as opposed, let's say, to complex data or large information bases). It's coming into use only recently and in a specialized context might suggest that the term is a proper noun, i.e., the first letter of the two words is capitalized. However, common use – lower case "b" and "d" - says otherwise thus giving guidance here (SAS, n.d.). Hence, we will refer to "big data".

It helps to place a datum in its context using a common complexity model, i.e., data, information, knowledge and wisdom (DIKW). This DIKW sets forth:

- Data – mere phenomena
- Information – Phenomena that are processed as being intelligible or not
- Knowledge – Information that is evaluated as to quality or integrity
- Wisdom – Knowledge that is drawn upon to create ideas not found in the knowledge alone (DIKW Pyramid, 2018) (see Figure 1)

Data by itself is phenomena, bits being expressed as zeros and ones in the computer world. Whether it has meaning or not is the job of the information scientist winnowing out whether there is randomness (unpredictability) or patterns. That information is run through an accountability filter, the result being from epistemology (justified belief or knowledge). Synthesizing from that knowledge gives us wisdom.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Data-Information-Knowledge-Wisdom Pyramid (Ibid.)



The subject of this paper is multiples of datum but is it information? Data is to be processed because often what is encountered is unintelligible. Consider:

```
0001, 0010, 0011, 0100, 0101

0111010101010010.
```

Both are data. There are two ways of looking at whether it makes sense. A mathematician or competent computer scientist will recognize the first sequence as binary or base two counting. The second sequence has no discernible pattern. That the first is intelligible as a pattern means that there is information. The second not having any pattern means there is no recognizable information. However, the other way of "making sense out of nonsense" is reversed, in attempting to translate the binary into words. The first does not translate into text, as it is not divisible by eight. The second does translate into "uR". In other words, whether data has meaning or not can be perspectival; it depends upon one's intent, point of view, and so forth. (As an aside, determining if something has content given a perspective – as in seeing the "uR" instead of nonsense – is meta as opposed to object level. One is talking about something rather than directly using it. For example, one may talk in English about the Spanish "mesa" as meaning "table", rather than simply saying in Spanish "poner esto sobre la mesa" – put this on the table.) While there may be information because of meaning does not say there is knowledge.

The collection of data may be by recording information electronically or even by hand. (For example, a person knowing about the census realizes that at least some of the data is collected the old-fashioned way by physically writing down answers.) How to interpret the data comes under "information" - making sense of it. Whether it has any value refers to knowledge and this is largely dependent upon organization and subsequent validation.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The Origin of Data Collection

It is thought that Protagoras (c. 490 c. 420 BCE) said "Man [humans] is the measure of all things" (Wikipedia, 2018n), be it a qualitative description of phenomena or a quantitative measurement. Data collection can be as simple as making a tally mark on a stone, as was done in the Upper Paleolithic period of human development (Wikipedia, 2016) between 10,000 and 50,000 years ago. The 30,000-year-old paintings in the southern French Chauvet-Pont-d'Arc Cave (Wikipedia, 2018d) are data. Any recorded description involves data.

Measurement involves a language that conveys a concept of comparison. Comparison rests upon a fundamental law: something exists because of what it is not, a law of dialectics. One may picture her/himself in a room where everything is of the same shade and texture. There would be no way of discerning one object from another. If the lighting is changed, let us say as a single light bulb being the source of illumination, shadows will appear, each object now contrasted with the other by the shadowing. This simple discernment is comparison. Measurement can yield both the discernment and its magnitude of difference.

Data is about description, description is about comparison and comparison is about measurement. Otherwise stated, measurement, in its most basic form is about empiricism or understanding through the senses or in modern parlance, "sense data". These data are collected through sensors, our bodies having the most immediate sensors: tongue (taste), skin (tactile), eyes (sight), nose (olfactory) and ears (sound). Humans construct devices to extend the abilities of the natural sensors. Fog computing, itself, is a manifestation of how ubiquitous these sensors are.

Perhaps the simplest form of an artificial sensor is the sun dial or even a stick in the ground reflecting the position of the sun in its periodic traversing of the Earth's sky. The specific length of a shadow indicates time of day. The shadow is a piece of data, the information being that the shadow is so many centimeters long, the knowledge being that it is five o'clock in the afternoon. All of these sensors, natural and artificial, collect data, that data being processed to determine if it makes sense or contains information. Whether that information has any value requires higher cognitive faculties to evaluate it, memory, ability to compare and so forth.

To understand the world, we have come to rely on empiricism that which is gained through the senses. Science applications have yielded sensor technology where data about every motion and energy emittance can be sent to us for our devices to process. The explicit way of generating detail was verbalized by Rene Descartes (1596-1650), the modern originator of the analytical method. Descartes said, "...we cannot conceive body unless as divisible … body is by nature always divisible... (Descartes, 1641)". He set the stage for modern deduction theory and analysis thus highlighting the philosophy of the Age of Reason or Rationalism. It is subdivision, after all, that is the origin of big data.

Current Sources of Big Data and Rationale for Generating It

Big data results from measuring complex phenomena. Broad categories are exemplified by:

- Fluid dynamic processes (weather, bodies of water, behavior of gases)
- Nuclear processes (Large Hadron Collider, simulation of nuclear explosions)
- Social interactions (economic, political, educational, sociological)
- Biological descriptions (Human Genome Project, Artificial Intelligence)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In essence, if one wants to produce data, simply subdivide the phenomena or synthesize by relating two or more pieces. Observing phenomena and collecting information on it is guided by basic parameters as in:

- The nature of pattern and randomness
- What to look for
- How to look for it
- What to do with it

The "Techie Stuff" Concerning Big Data

Characteristics of big data often are listed as:

- Volume - How much of it there is
- Variety - How much difference there is in what is collected
- Velocity - How fast the data is coming in and processed
- Variability - How consistent is the characteristic of the data coming in
- Veracity - Refers to the quality of the data received
- Complexity - The number of sources as well as the type of data received (Big Data Characteristics, 2015; Kalbandi, 2015)

"Volume" is discussed here in terms of storage. "Velocity" refers to processing speed, a topic covered here after storage. "Variety", "variability", and "complexity" correspond to organizational issues. "Veracity" is discussed in terms of the data-information-knowledge-wisdom (DIKW) hierarchy discussed above.

Four parameters may be considered as limiting big data: how much potential data there is to be collected, the storage of it, the processing speed and the management (organization, interpretation and application) of it. We now treat each of these in turn.

Collecting data is no problem, as the plethora of sensors (Tae-il, 2013) indicates, as well as our capabilities of Cartesian analysis (Descartes, 1641). Any object can be subdivided to Planck scale, 1.6 x 10^{-35} meters. At this quantum level particles "flick in and out of existence" (Hawking, 2018). To illustrate the possibility of representing every aspect of our universe, one may consider Bostrom's (2003, pp. 243-255) idea that we may be inside of a matrix, part of a computer program. Such would require modeling and managing 1080 bits corresponding to the number of particles in the universe according to one physicist (Heile, 2015) and the same as the Eddington number of protons (Wikipedia, 2018i). This quantity of data far exceeds any current ability to store it, let alone process it, although quantum computing (Wikipedia, 2018o; Khodjasteh, 2013) may provide solutions in the future. Also, it may be possible to approximate a description of the universe by representing groups of these particles. Here, it is the quality of representation and subsequent organization of data that would be the determinant in successful management. Further,

A simulation that is undetectable by science (for instance, both astronomers and quantum physicists) would have to be very large and detailed, it would be very hard to justify its vast costs (it's easy to compute what is the physical volume of computer memory required to try to simulate a unit volume of substance-Avogadro constant is still very large in comparison to the volume of computer memory: The

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

most powerful supercomputer in the world (in June 2015) has "only" circa 1015 bytes of memory, compared with circa 6 x 1023 - particles contained only in 12 grams of unit matter and such supercomputers nowadays still have the weight of hundreds of tons.). (Webb, 2014)

To place data storage in perspective, consider that the whole document set of the U.S. Library of Congress can be stored in 200 terabytes (TB) (Johnston, 2012). Shantanu Gupta, director of Connected Intelligent Solutions at Intel is reported as saying that:

- On YouTube, 72 hours of video are uploaded per minute, translating to a terabyte every four minutes
- 500 terabytes of new data per day are ingested in Facebook databases
- The CERN Large Hadron Collider generates 1 petabyte per second
- The proposed Square Kilometer Array telescope will generate an Exabyte of data per day
- Sensors from a Boeing jet engine create 20 terabytes of data every hour (Higgenbottom, 2012)

To relate data storage quantities to each other, consider Table 1:

In 2010, it was estimated that storing a yottabyte on terabyte size disk drives would need one million city block size data-centers as large as the states of Delaware and Rhode Island combined (Diaz, 2010). However, the writer does not specify how tall the buildings would have to be and as will be seen shortly, the capacity of disk drives keeps growing. Too, Quantum Computing may drastically enlarge the ability to store data currently kept in separate spaces all in one place (Khodjasteh, 2013).

Table 1. Multiples of bytes

MULTIPLES OF BYTES					
Decimal			Binary		
Value		Metric	Value	IEC*	JDEC+
1000	kb	kilobyte	1024 KiB	kiloibyte	KB kilobyte
1000^2	MB	megabyte	1024^2 MiB	megaibyte	MB megabyte
1000^3	GB	gigabyte	1024^3 GiB	gigaibyte	GB gigabyte
1000^4	TB	terabyte	1024^4 TiB	teraibyte	ENIAC – *circa* 1948 (U.S. Army photo)
1000^5	PB	petabyte	1024^5 PiB	petaibyte	
1000^6	EB	exabyte	1024^6 EiB	exaibyte	
1000^7	ZB	zetabyte	1024^7 ZiB	zetaibyte	
1000^8	YB	yottabyte	1024^8 YiB	yottaibyte	

* International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) - https://en.wikipedia.org/wiki/ISO/IEC_80000
+ The Joint Electron Device Engineering Council - https://en.wikipedia.org/wiki/JEDEC_memory_standards#Unit_prefixes_for_semiconductor_storage_capacity

A better representation of this image can be seen in the electronic version of the publication.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

A further perspective on yottabyte storage capacity is:

With recently demonstrated technology using DNA computing for storage, one yottabyte of capacity would require a volume between 0.003 and 1 cubic meter depending on number of redundant backup copies desired and the storage density: "Our genetic code packs billions of gigabytes into a single gram". DNA is much less mature technology than microSDXC cards (for this application) and accompanied by uncertain costs but this suggests potential information density. (Yottabyte, 2018)

Here and elsewhere in this article are descriptions of technological capacities, such as the size of storage devices, figures which have been increasing as advances are made. Accordingly, little attempt has been made to obtain new figures since the publication of the previous article, as it is those figures not only are sufficient to make the point of the arguments but are now even more reinforced by those advancements.

As of this writing (March 2018), for a desktop personal computer (PC) one can purchase across the counter in retail consumer outlets a four Terabyte (TB) hard disk drive (HDD) for under $100. In September 2014, a desktop PC 10 TB HDD appeared (Hard Drive Capacity, 2015). In 2018 the capacity had increased to 12 TB and for under $800. One TB thumb, or flash drives were available at the end of 2015 (Thumb Drive Capacity, 2015). While the prices started out high, as with the current lower storage capacity ones, price drops could be expected, given the downward trend in prices in other technologies. This all says nothing about solid state drives with vastly reduced access times. These are desktops and laptops that the ordinary consumer can purchase. Data storage does not seem to be much of a problem. What, then, about processing speed?

A four-time winner, Tianhe-2, a supercomputer developed by China's National University of Defense Technology, once more took home the title as the world's fastest computer with a performance of 33.86 petaflops (quadrillions of calculations per second) on the Linpack benchmark. To reach that remarkable number, Tianhe-2 uses over 3 million Intel Xeon E5-2692v2 12C cores running at 2.2GHz. It's not the speed of the processor you see in modern-day supercomputers that's important; it's how you get all those processors to work in concert with each other. In addition, its main processors are backed up by Intel Xeon Phi 31S1P coprocessors. (Tianhe, 2015)

Then, "President Obama has signed an executive order calling for the US to build the world's fastest computer by 2025 (Baraniuk, 2015). The supercomputer would be 20 times quicker than the current leading machine which is in China (Supercomputers, 2015).

In terms of cloud storage, the sky is the limit. One only needs the hardware and connectivity. With many free email services, 15 gigabytes of storage are routine. Then, there are dozens of data storage services such as Dropbox, Trend Micro, and Google (Wikipedia, 2018e; Wikipedia, 2018f). Of course, there are pros and cons of such methods, but data storage capacity itself is not an issue, at least for home use and most scientific projects.

Organization of the data is the bottleneck. However, organization is somewhat simplified if one knows the purpose for which the data is sought. Once an organizational scheme is identified, then the sophisticated and often abstract aspects can be presented visually. That a "picture is worth a thousand words" works quite well if that picture is that of the purpose.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Visualizing Big Data and a Research Excursion in Pattern Recognition

Pictorializing big data is for our benefit. A picture is worth a thousand words. The display of the data itself or a picture representing it may give us some insight into the data's meaning that verbal descriptions may not be able to do. For examples, one does an internet search for "big data visualization" and similar phrases, selecting "images for" to see the myriads of ways of presenting big data displays. There are just as many images as there are algorithms to generate them, given an Internet search under "Visualization of big data" and similar phrases represented by the examples above. Many times, visualization techniques are just as much of an art as they are a technical methodology. One visual presentation of what otherwise would be incomprehensible in formulas and words may not make any sense to one person but may be enlightening to another.

How might the data be visualized? It has been said that, as in statistics, there are two basic forms of presentation: to the specialist in the area and to the public. There are processed and unprocessed data where less refinement is required for the specialist. In each area of inquiry, different project teams may collect and process data with different methods.

Particularly relevant to this discussion is open data where:

Thanks to open data we now have services such as App&Town: an application that acts as a guide from your point of departure to your destination using Barcelona's metropolitan transport services. The app offers the best routes and indicates where the nearest stop or station is as well as advising when you have to get on or off; or how about CityBikes, which provides information about the city's Biking system. As a useful aid for the health sector mention should be made of SOS info: an application that shows information about the user (allergies, illnesses, emergency contact details, etc.) and the location of the nearest emergency services. (Spelling in original) (Barcelona Smart City, 2015)

An additional method of visualizing data is with context diagrams showing the interrelation of components in a system (Wikipedia, 2018g). These often convey a crisp and clear meaning about what the data means as opposed to tables of data about voters wanting something from their representatives or figures showing what happens inside the system (as in traffic patterns, revenue, numbers of students per school, etc.).

Now, we get to the deep structure of data and begin our trip of pattern recognition.

It needs to be stated at the outset that bivalency is used to designate the on and off switch conditions in computer architectures. No inherent meaning exists in the way computer scientists design these architectures or the machine language software they manage. Nor is there any particular significance in the zeros and ones in themselves as any symbol could be used to represent the on and off conditions. The bivalent logic that computers use to generate data of any size is a language of a basic order in the universe (Horne, 1997). Here, the logic is specifically applied to how the concept of magnitude is regarded, in this case a base two counting, most commonly expressed by the zeros and ones. Sets of zeros and ones are arranged according to rules or instruction sets, each computer chip architecture having a different way of configuring these. Machine code instruction sets are unique to each processor or processor family, the instructions themselves being physical designs of bit arrangements that represent machine commands. This means that the instruction set belongs to a particular processor class but using generally the same type of architecture (Machine Code, 2018) rather than from any natural ordering, as in binary counting.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

With any digitized phenomenon, including large bodies of data, recursion might be used as a method for detecting patterns not otherwise revealed by mathematical analysis or visualization. The Wuensche and Wolfram designs above result from repeatedly and randomly concatenating binary operators, a type of recursion. Recursion here means that the outputs of a binary formula are fed back in as inputs until the function reappears as an output (Horne, 2006).

There is a natural binary space that emerges from fundamental dyadic computations of base two (Horne, 2006). Contained in a display of zeros and ones, as above, are sets of four representing binary functions where p and q are placeholders for the permutation of zeros and ones, shown in Table 2.

A logician, mathematician, and computer science student will recognize that this table shows a binary counting from left to right, reading the columns downwards: 0000, 0001, 0010, etc., functions f0, f1, f2, f3, onward to the right. Any display of zeros and ones will contain complete or partial sets of functions from the TOLS as in Box 1 (using the arbitrarily selected color code):

Not only is each of these functions recursive, but any function over any other two functions, f*(fn,fp) is recursive as well because there can be only a maximum of 16 iterations (only 16 unique functions) on any branch before there is a repetition of one of the previous (Horne, 2006). Otherwise stated, even if there are 16 unique outputs on a branch, there is repetition on the 17 iteration, i.e., no function can have more than sixteen iterations. This is reflected by a three-dimensional hypercube displaying all 4096 dyadic relations (Horne, 2012). Of 16 panels, there is, for example, f_{13} - ⊃, p contains q, p ⊃ q, – defines deduction shown by Table 3.

The recursion method is shown in Table 4 with the diagram for Function 9 in Figure 2.

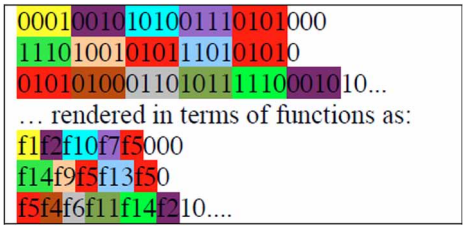
We said that just because there does not appear artistically to be a pattern, it doesn't mean that there is not one as the example above illustrates. In terms of visualization, reducing binary strings to functions may assist in detecting patterns. For example, we have the traditional truth table rendering of (p & q) ' [(r = s) v (p ' r)] in Table 5.

Table 2. Table of Functional Completeness (Horne, 2006)

p	q	f_0	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

A better representation of this image can be seen in the electronic version of the publication.

Box 1



A better representation of this image can be seen in the electronic version of the publication.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Panel 13, f_{13}, p⊃ q (Horne, 2012)

f_{13}	f_0	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}
f_0	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}
f_1	f_{14}	f_{15}	f_{14}	f_{15}	f_{14}	f_{15}	f_{14}	f_{15}	f_{14}	f_{15}	f_{14}	f_{15}	f_{14}	f_{15}	f_{14}	f_{15}
f_2	f_{13}	f_{13}	f_{15}	f_{15}	f_{13}	f_{13}	f_{15}	f_{15}	f_{13}	f_{13}	f_{15}	f_{15}	f_{13}	f_{13}	f_{15}	f_{15}
f_3	f_{12}	f_{13}	f_{14}	f_{15}	f_{12}	f_{13}	f_{14}	f_{15}	f_{12}	f_{13}	f_{14}	f_{15}	f_{12}	f_{13}	f_{14}	f_{15}
f_4	f_{11}	f_{11}	f_{11}	f_{11}	f_{15}	f_{15}	f_{15}	f_{15}	f_{11}	f_{11}	f_{11}	f_{11}	f_{15}	f_{15}	f_{15}	f_{15}
f_5	f_{10}	f_{11}	f_{10}	f_{11}	f_{14}	f_{15}	f_{14}	f_{15}	f_{10}	f_{11}	f_{10}	f_{11}	f_{14}	f_{15}	f_{14}	f_{15}
f_6	f_9	f_9	f_{11}	f_{11}	f_{13}	f_{13}	f_{15}	f_{15}	f_9	f_9	f_{11}	f_{11}	f_{13}	f_{13}	f_{15}	f_{15}
f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}
f_8	f_7	f_7	f_7	f_7	f_7	f_7	f_7	f_7	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}	f_{15}
f_9	f_6	f_7	f_6	f_7	f_6	f_7	f_6	f_7	f_{14}	f_{15}	f_{14}	f_{15}	f_{14}	f_{15}	f_{14}	f_{15}
f_{10}	f_5	f_5	f_7	f_7	f_5	f_5	f_7	f_7	f_{13}	f_{13}	f_{15}	f_{15}	f_{13}	f_{13}	f_{15}	f_{15}
f_{11}	f_4	f_5	f_6	f_7	f_4	f_5	f_6	f_7	f_{12}	f_{13}	f_{14}	f_{15}	f_{12}	f_{13}	f_{14}	f_{15}
f_{12}	f_3	f_3	f_3	f_3	f_7	f_7	f_7	f_7	f_{11}	f_{11}	f_{11}	f_{11}	f_{15}	f_{15}	f_{15}	f_{15}
f_{13}	f_2	f_3	f_2	f_3	f_6	f_7	f_6	f_7	f_{10}	f_{11}	f_{10}	f_{11}	f_{14}	f_{15}	f_{14}	f_{15}
f_{14}	f_1	f_1	f_3	f_3	f_5	f_5	f_7	f_7	f_9	f_9	f_{11}	f_{11}	f_{13}	f_{13}	f_{15}	f_{15}
f_{15}	f_0	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}	f_{15}

A better representation of this image can be seen in the electronic version of the publication.

Table 4. Recursion method (Horne, 2006)

For the p side: F*(fp, fq) = fa F*(fa, fq) = fb BaraniukF*(fb, fq) = fc ….	For the q side: F*(fp, fq) = fa F*(fp, fa) = fb F*(fp, fb) = fc …. Max F*(fp, fq≤16) before there is repetition

This is reduced to a table of functions (Table 6), making it easier to discern regularities.

In Table 6, without the distraction of the detail of the 0s and 1s, one may discern repetitions of functions. Again, the configurations do not necessarily convey any deeper meaning than that there is a regularity. One, however, should ask what "generates it" and moreover, "why", beyond the simple fact a formula exists. That is, what generates the formula?

Latent or unexpected arrangements and processes may exist inside data displays. What results is "emergence", something that cannot be explained by the properties of the individuals used to create the results (Emergence, 2018). All digitized data is in its most basic form bits (zeros and ones). Usually, data storage is according to predefined categories used to collect it. The more vague and coarsely grained the categories, however, the less context and often the less value. Displays of large data sets as zeros and ones may reveal patterns whereas other forms of analysis may not. This works in at least two ways, the first being the existence of patterns in a binary display but no known ordering (or categorization of data) principle and the second where the ordering principle is not known but there results a pattern which then tells the person that principle.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Recursion diagram for f_9

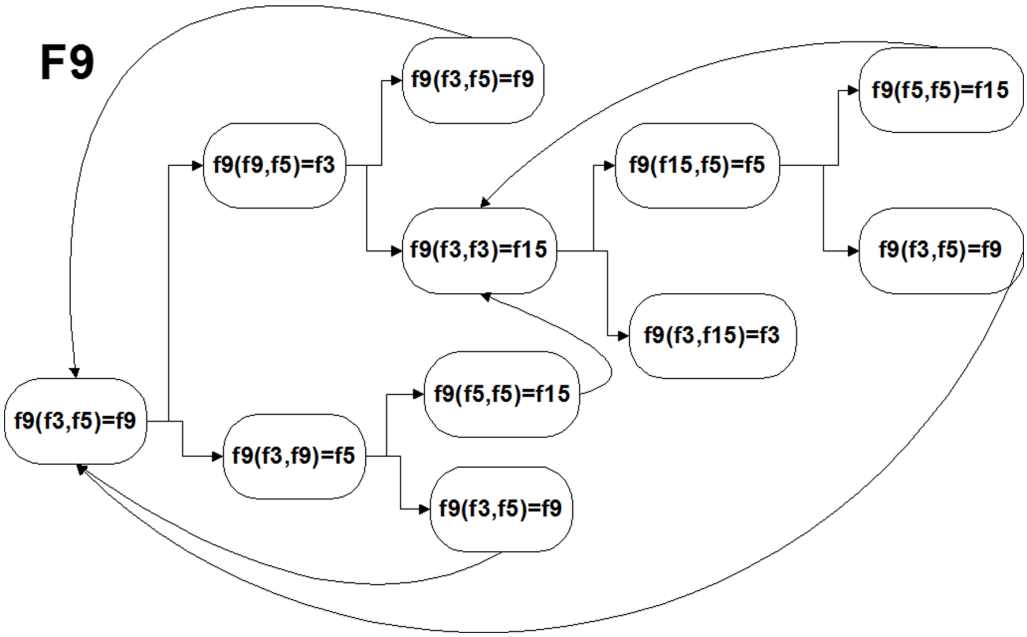


Table 5. Traditional truth table (partial) (Horne, 2012)

p	q	r	s	(p	&	q)	→	[(r	≡	s)	v	(p	→	r)]
0	0	0	0	0	0	0	1	0	1	0	1	0	1	0
0	0	0	1	0	0	0	1	0	0	1	1	0	1	0
0	0	1	0	0	0	0	1	1	0	0	1	0	1	1
0	0	1	1	0	0	0	1	1	1	1	1	0	1	1
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1	1	0	0	1	1	1	1	0	1	0	1	1	0	0
1	1	0	1	1	1	1	0	0	0	1	0	1	0	0
1	1	1	0	1	1	1	1	1	0	0	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

A better representation of this image can be seen in the electronic version of the publication.

Table 6. Truth table in terms of functions (Horne, 2012)

					1		3		1		2		1	
p	q	r	s	(p	&	q)	→	[(r	≡	s)	v	(p	→	r)]
					f_1				f_9		f_7		f_{13}	
f_0	f_0	f_3	f_5	f_0	f_0	f_0	f_{15}	f_3	f_9	f_5	f_{15}	f_0	f_{15}	f_3
f_0	f_{15}	f_3	f_5	f_0	f_0	f_{15}	f_{15}	f_3	f_9	f_5	f_{15}	f_0	f_{15}	f_3
f_{15}	f_0	f_3	f_5	f_{15}	f_0	f_0	f_{15}	f_3	f_9	f_5	f_{11}	f_{15}	f_3	f_3
f_{15}	f_{15}	f_3	f_5	f_{15}	f_{15}	f_{15}	f_{11}	f_3	f_9	f_5	f_{11}	f_{15}	f_3	f_3

A better representation of this image can be seen in the electronic version of the publication.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

For example, there is a pattern, a regularity, in the following images in Figure 3 and Figure 4, generated by randomly concatenating dyadic binary formula but no meaning is obvious beyond that, the issue being that apparent randomness can spawn order.

These patterns of Wuensche (1993) and Wolfram (2002) may not represent or convey any meaning by themselves as in a display of machine code (Wikipedia, 2018k) of zeros and ones. Rather, the art results from a naturalness. East Asian philosophies have mandalas and koans, objects and expressions upon which persons can meditate to gain meaning and perhaps these images may be included. Overcoming randomness is through pattern recognition and such allows us (or so we think) to manage the future. As

Figure 3. Regularity in cellular automata (1) (Wolfram, 2002)

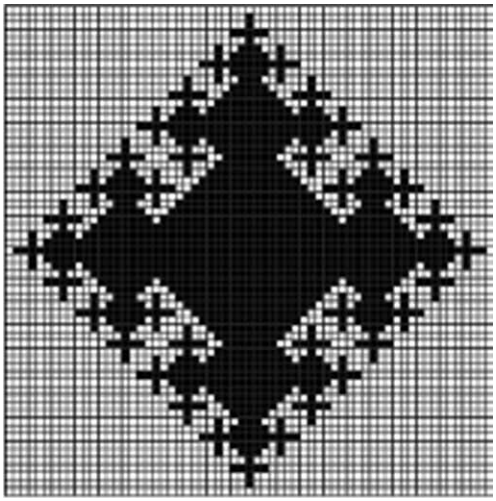
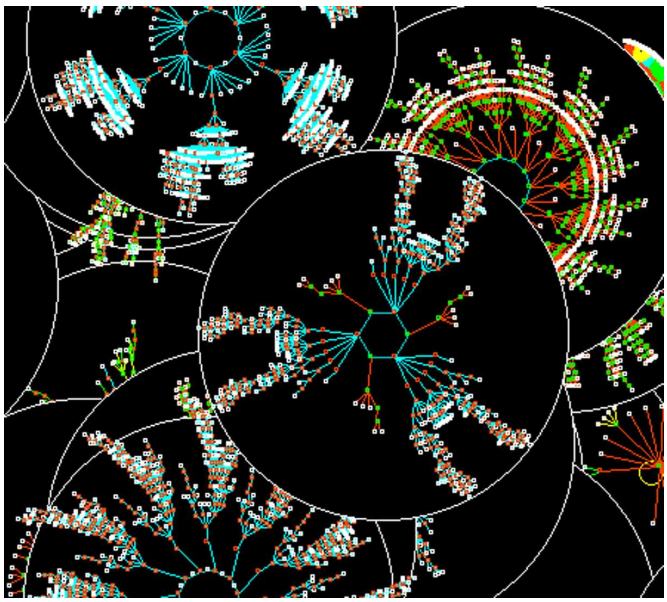


Figure 4. Regularity in cellular automata (2) (Wolfram, 2002)



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

has just been seen, though, regularity does not compel us to repeat it, only to know that an order exists. Above, the regularity in the visualization seems to challenge the supposed randomness used to generate the images. Tautologically, knowledge of order is foundational to answering why it exists.

The second example of discerning a pattern in a binary display and extrapolating the algorithm is exploratory research representative of visualization that may explain data ordering where initial analysis may not. In a more purposeful fashion, most modern computers generating and managing big data use machine language, an example of which is:

```
01101100 01100101 01100011 01110100 01110010 01101111 01101110 01101001
01100011 00100000
01101101 01100101 01100100 01101001 01100001 00100000 01110100 01101000
01100101 01110010
01100101 00100000 01100011 01100001 01101110 00100000 01100010 01100101
00100000 01110110
01100101 01110010 01111001 00100000 01101100 01101001 01110100 01110100
01101100 01100101
00100000 01110000 01110010 01101001 01110110 01100001 01100011 01111001
00100000 01100001
01101110 01111001 01110111 01100001 01111001 00101110
```

To a person unfamiliar with binary code, this display would be meaningless, but its meaning is translated as, "With an electronic media there can be very little privacy anyway" (Binary Translator, 2018).

On the heels of this statement two aspects need to be considered about what is found inside data, other than the ostensible information they may convey. First is a pattern inside the zeros and ones that may emerge despite attempts to mask their ostensible messages. The second is to illustrate that wherever there are zeros and ones there is the opportunity to introduce other information, as in malware.

What is Fog Computing?

This author loves standardization, and the U.S. National Institute of Standards and technology (NIST) is working on a definition, as expressed in its August 2017 "The NIST Definition of Fog Computing" (NIST, 2017, p. 2):

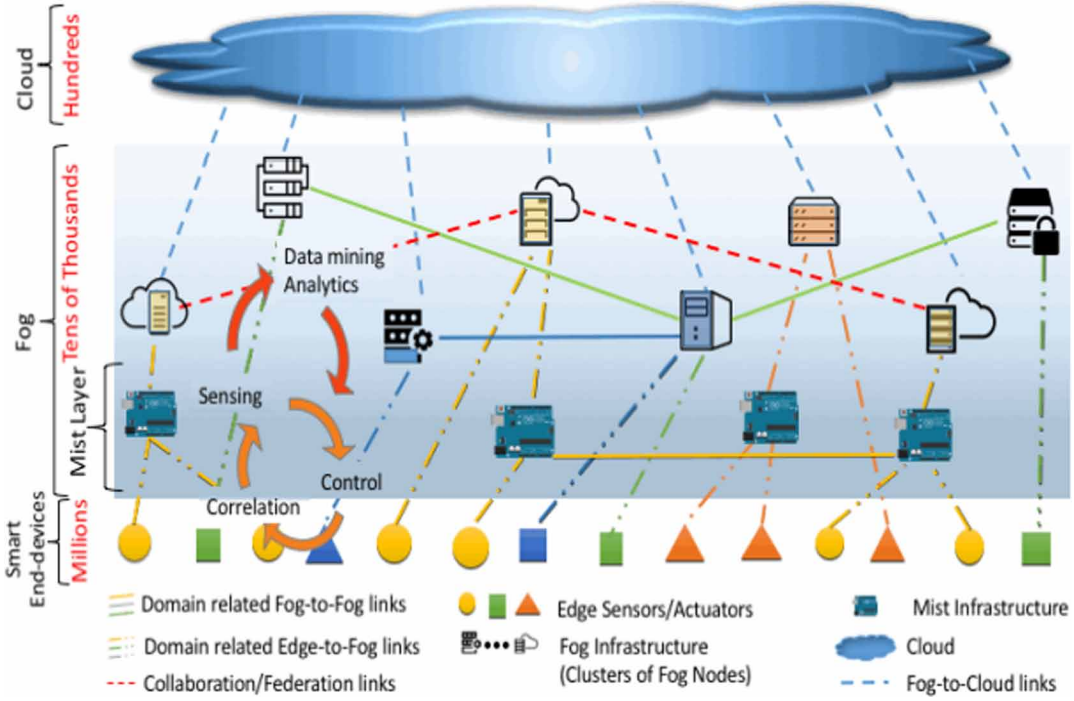
Fog computing is a horizontal, physical or virtual resource paradigm that resides between smart end-devices and traditional cloud or data centers. This paradigm supports vertically-isolated, latency-sensitive applications by providing ubiquitous, scalable, layered, federated, and distributed computing, storage, and network connectivity.

Let us compare NIST's diagram of fog computing to human anatomy.

The cloud might be considered the human body, with the "fog" underneath it the various organs, and the "smart-end-devices" the neurons. Here, the NIST diagram might be a misleading, as big data can and usually is stored in computers, with the devices feeding in an receiving big data. There seems to be a transformation, where big data is being shifted to the cloud, with the computers being its manipulators.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Cloud computing configuration (NIST, 2007, p. 8)



Of course, the "elephant in the room" question is where the big data is being stored in the BIST model, other than the "cloud". Of course, the answer is in computers, albeit especially configured.

Still the hierarchy of number applies, there being many more neurons than nerves, and many more nerves than organs, and so forth. The overall discussion does not change, in that there are what analogously are data reception, transmission, and manipulation devices (cell phones, IoT, etc.) and a passage from these to a central area (cloud, huge data banks, etc.), the whole configuration effectively acting not unlike a huge organism that potentially could develop a life of its own.

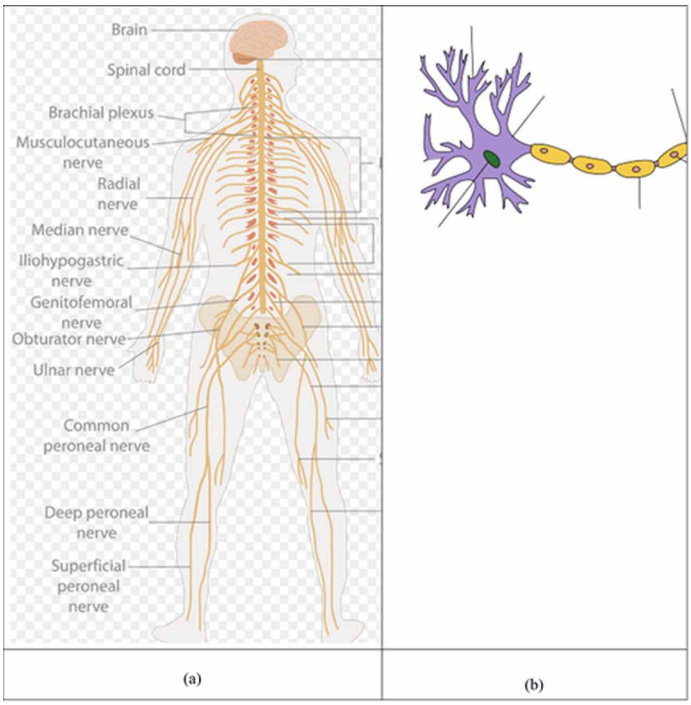
We will see all of this in a social context later, but for now the comparison should be obvious, with all the attendant implications of self-organization, as the discussion about Wolfram, Wuensche, and emergence above indicated. Cloud computing might be regarded as a brain in the Internet overall, or at least that part of the Internet brain responsible for memory (see Figure 6).

THE IMPORTANCE OF BIG DATA AND FOG COMPUTING

Big data concerns us for several reasons. The quantity of phenomena is so large that it makes the job of organizing, interpreting, and controlling it enormous, our frustrations often being reported in terms of "too much information" (TMI) or "information glut". One's ability to process it is complicated further if there is more potentially extraneous data. Then the "fingers" of control with fog computing ultimately manipulated by those generating and applying big data are more ubiquitous than ever, the end-users of data weighing convenience against privacy.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 6. (a) TE-Nervous system diagram (Own Work, 2010); (b) Neuron (Quasar Jarosz, 2009)



There are many relatively neutral projects where huge amounts of data result but to which there is little direct relation to us as in data from surveying the universe. Closer to us, there is information more of a personal concern. Each time one makes a purchase, data is collected, mainly the details about the transaction amount, date, store and item purchased. Beyond that there are other data about the person making the transaction, web data collection forms being a primary example. We are asked about who is making the purchase, why, income, age, sex and so forth. Marketers and advertisers particularly are interested in a buyer's profile for repeat business. Universities collect data on students and faculty and government do on the citizens/subjects in order to manage them.

The importance of what is to be identified comes first. Then, is how to collect it. The quality of collection is next. Organization is next. Then, we need to interpret whether collected data is information or not, as in the DUIKW pyramid discussed previously. These areas may overlap as in identifying the pieces before organizing them. We now come to the eye of the storm, humanity, itself. The philosophy underpinning big data can be discussed as concerning that apart from us as in scientific studies or what persons collect when we, ourselves, are observed. There are individuals and there is the society in which they live; big data exist for both. Philosophy is about asking "why".

Two aspects emerge. The first is about how we relate not only to big data, itself, but how we relate to those who are using it. Besides the privacy issues are social policy analysis and application in achieving ends which may not be agreeable to us, increased complexity of cataloguing use and access, and security. The second aspect is about the organic character of big data and how its development and evolution may track the development evolution of societies. Society may have a collective brain with its memory, so

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

the collection of big data can be regarded as a memory to be accessed and used by artificial intelligence or consciousness. First, though, let us look at what is being collected now as big data and then we can discuss the future in terms of societal development.

The Social Source of Data

Data results from a record keeping, the purpose of which is to understand how to manage the future based on the past. This method is an outgrowth of the millennial-old playing of games but in a more formalized manner. For example, assessing a person's throwing capability by itself may not have much value but there is more meaning when viewed as the person having to compete with warriors with different capabilities. The data is a record of that person's capability. The individual is viewed in a systemic context, the society benefiting from a positive battle outcome. Further on down, the nature of social systems will be discussed, but for now a system is viewed as a collection of elements related by having a common purpose and interacting with an environment so as to be able to survive or adapt.

From various projects emerge big data but one should ask what produces the projects. Big data often is associated either with just itself as in an often-inchoate mass of bits to be processed into information and organized or big data is associated with large-scale undertakings such as designing medicines or analyzing thermodynamics, but these do not simply happen. Organizations initiate them; more precisely, persons individually or collectively with motivations having enough influence in the organizations initiate them. In attempting to describe the future behavior of a society researchers will describe it with a model, much in the same manner a toy airplane "describe" a real one.

Researchers instantiate a model with various hypothetical factors and events to observe possible outcomes. As an example, one speaks of ecological systems. A particular one, as that in any country, is a model. The climate, vegetation, fauna and so forth are identified as elements. One then gives these variables a special condition, as there being a specific type of animal, tree or crop and studies their interactions in a simulation. Another example is that of a city where aspects such as traffic flow, electrical power management and communications management in themselves require big data sets. However, each of these is a subsystem within the framework of a larger one, the city. Here, all the data is collected, processed, interpreted, and disseminated within the larger context of a large dynamic system (the city), analogous to an organism composed of numerous subsystems all working together, a key word here being "organism". The simulation occurs by inserting specific examples of vehicles in a traffic flow, number of megawatts of power being supplied and how many phone calls are being made at a specific time.

One often cannot predict from big data sets precisely what will happen. This means that an outcome cannot be explained by what the pieces are doing or their properties individually. This is called "emergence" (Vintiadis, 2018). In essence a collection of entities with a purpose (a system) seems to assume an activity or life of its own. Such is what appears to be "organic behavior" in big data in the form of the Internet.

Networked computers are subject to compromise when there are one or more deleterious factors involved, malware (Wikipedia, 2018L) being the prime example. Over the years a medical or organic model has been used to address the problem as the "virus" and "anti-virus" terminology illustrates. So too the Internet has assumed a character that can be considered "organic" as a search for "Internet as organism" will show. In a way similar to how big data being generated in drug trials (simulated and real), a simulation may show what effect a specific piece of malware would have on the Internet at a specific time.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

A simple Gedanken, or thought, exercise will place data collection about us in perspective. If there is a phenomenon, it can be detected either naturally or with some aid as in a sensor. This is no less true for ourselves in every organ, cell, movement or behavior as we above with the cellular-scale sensors that can be integrated with body tissue (Tae-il, 2013). One can subdivide to the quantum level and organisms including *Homo sapiens sapiens* (the correct and technical name for human beings) are not exempt. There are numerous projects that collect and manage big data and these in themselves are feats that impress technophiles, but a larger picture emerges when they are seen as articulating as an organic whole. Who would care, were it not for the controversies over data collection and privacy, most of which focus on one's privilege to protect identity and prevent control by others? We are our own commanders, and the more someone else knows about us the greater potential they have to rob our identities, extort us and even replicate us making the individual superfluous.

Again, we need to be asking throughout, what drives big data? It is not simply about describing an individual in her/his entirety even the brain and consciousness. There are large data sets generated by easily identifiable projects, but, overall, we have that background question about the coalescence of large data sets. It may be argued successfully that the full power of the State rests upon the ability to collect, manage, and analyze data and the ability to exert its power (ultimately in a physical manner). We will discuss more formally later the nature and role of the State.

One should notice this discussion has progressed from talking about big data emanating from systems without human involvement, such as thermodynamic systems, quantum-cosmological phenomena (CERN [originally Conseil Européen pour la Recherche Nucléaire]), the Sloan Digital Survey) and ecological research to systems in which humanity plays an increasingly involved role or ones it has created outright as in the uninhabited autonomous systems (UAS), brain research, and human genome projects. It is humans - the creator and observer - to whom data and its management are important, and surely big data is created in these arenas, arguably the largest data sets of all.

Outside the development by governments are non-governmental big data projects that are just as sophisticated in replicating human behavior. Aside from the above references to artificial intelligence and non-hydrocarbon replication of the human brain are modeling and simulation projects of society, itself and for which research data can be obtained. Two prominent examples are *Simulated Society* (SimSoc), started in 1966 and the recent *Second Life* created in 2003. A visitation to these sites (easily accessible by an Internet search with these terms) will illustrate how far we have gone in creating an electronic world in which persons can interact to such a realistic extent that they will go so far as real world trading in currency (as in Linden Dollars for Second Life) and artifacts used in the simulation. We are on the edge of technological advancement not only in the realm of virtual reality but venturing into three-dimensional virtual worlds, where human consciousness could not only be altered but permanently affected and possibly merged into a larger world. In these social simulations it is instructive to observe not only the extent of big data generated but how the virtual and real world might cross over into each other by data manipulation.

Let us look at several examples of how big data is collected on our mentation (intelligence, emotions, psyche, behavior, mind, consciousness, etc.), our daily activities, and various social interactions with respect to the military.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Big Data About Ourselves and Psyche

To understand the conceptual framework for assessing the scope, purpose and ultimate meaning of big data as applied to our personal lives, we need to discuss its context, both historical and current. A central goal of modern scientific endeavor is to extrapolate from the past so as to be able to project to the future so as to be able to manage it. The more we know about objects and processes in our environment, the greater potential we have for managing the future. Immediate concerns over data collection on ourselves focus on privacy but a larger context shapes long-term outcomes of other having data on us. Here, each piece of datum is mapped on to an aspect of our lives. Precisely why would anyone want to do this?

Aside from the everyday data collection at supermarkets by various governments, educational institutions and corporations, we have seen data collection on ourselves in a detailed way but another concerns the way we interact with society.

Perhaps it is natural for a person to react in some way - often by surprise and wonderment - when particular attention is paid, as when someone is taking a picture unexpectedly. Answering a few questions in an initial social exchange is to be expected. "Where are you from? What do you do?", etc. However, in this initial and casual exchange, anxiety heightens the more interrogation proceeds. In exchanges with organizations, a certain amount of data is anticipated to be collected, but again, the more that is collected the greater concern there is about motivation.

In the government area, official business is recorded from the military, driver's licenses, taxes and so forth. For schools, there is student data that can be in detail ranging from name and address to psychological profiles and detailed financial information.

With each encounter with a person or organization arises an opportunity for measurements to be taken of every aspect. Stand back, observe these and other sources of data, and consider collecting all of this data under one roof, thus allowing storage of a complete description of a person's past, present and probable future. So far, we have seen examples of projects that each unto themselves can generate big data, but there is one project beyond that encompassing all of the above and more. Each of us has both a physical and mental health collected by doctors and hospitals and big data reflects this. Going beyond a doctor's visit, physiological data now can be collected at the cellular level, where micron-sized sensors can be integrated with tissue (Tae-il, 2013), these sensors also being a part of fog computing.

In the not so distant future, the health of each any every cell hypothetically might be monitored in a manner analogous to how numerous intelligent communications devices report on their own health (how data is collected, managed and applied) such as smart phones. Fog computing is designed to manage this level of data collection and management granularity. We are on a curious road of complexity that may lead us in unanticipated directions. Thermodynamic and other physical systems present degrees of uncertainties and ranges of possible configurations and behavior that tax even the most advanced of contemporary supercomputers (Tianhe, 2015). Yet, the most complex of all is our brain and associated consciousness.

What follows is not an argument that the research described can be accomplished successfully but that we need to pay attention to what is driving it. The former may be fiction, but the latter is fact. In the past decade, a number of projects have arisen to replicate the human brain, some of them being Blue Brain, Riken Brain Science Institute work and SyNAPSE (Blue Brain, 2015; Riken, 2015; IBM SYNAPSE, 2015; Human Brain Project, 2013). Ray Kurzweil (2015), leader of the Singularity Movement (Singularity, 2015) and Artificial Intelligence advocate categorically states that success along these lines is a distinct possibility. An appreciation for the scope and depth of research on replicating human menta-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

tion ("consciousness", intelligence, memory, etc.) may be gained by reflecting on a sample of available literature on the subject in "Steps toward Memory Preservation v1.6" (2018). Reading the bibliography is an eye-opening research exercise.

A description of physiology is easy compared to that of consciousness associated with the human brain (Toward a Science of Consciousness, 2018). The psychological domain has not been precise until more recently when diagnostic and treatment models became more medically oriented. Today, getting diagnosed and treated more often than not starts with a psychiatrist, a medical doctor specializing in the managing of mental problems. For many years the *Diagnostic and Statistical Manual of Mental Disorders* (DSM) has been used to identify persons with abnormal deleterious mental conditions or disorders (DSM, 2015). However, it is clear that a person emerging from a psychiatrist's office with an evaluation based on the DSM of being mentally healthy can and does engage in violent behavior. That is, nothing in the application of DSM criteria guarantees these people will or will not act violently. One reason is that there is no clearly defined validation mechanism for the DSM. Human behavior until recently has been described primarily by a subjective evaluation of behaviorists relying on observing how a person acts and her/his reports of interactions with the environment. There is no set of clearly defined and uniformly applied metrics to observable conditions; diagnosis is behaviorally based. The DSM is merely a checklist of phenomena to be recorded and judged from professional training and experience. Even though the American Psychiatric Association is transitioning the DSM to the United Nations World Health Organization's International Classification of Diseases (ICD) (ICD, 2015) and even though it may be a more detailed description of a disorder, the same remarks about the DSM apply to it as well.

In 2008, the U.S. National Institutes of Mental Health (NIMH) initiated its Research Domain Criteria (RDoC) project to arrive at diagnosing mental conditions based on an integrated approach including neuroimaging and genetics (RODC, 2018). In NIMH's word:

RDoC is a research framework for new ways of studying mental disorders. It integrates many levels of information (from genomics to self-report) to better understand basic dimensions of functioning underlying the full range of human behavior from normal to abnormal. The RDoC framework consists of a matrix where the rows represent specified functional constructs (concepts representing a specified functional dimension of behavior) characterized in aggregate by the genes, molecules, circuits, etc. used to measure it. Constructs are in turn grouped into higher-level domains of functioning, reflecting contemporary knowledge about major systems of emotion, cognition, motivation and social behavior. (RODC, 2018)

Cognitive neuroscience is a rapidly developing field in which various parts of the brain are found to be responsible for specific behaviors (Brain Behavior, 2015; Wimmer, 2012; Gazzaniga et al, 2009; Kanai, 2011; Dawes and Fowler, 2009; Miller, 2012; Cope, 2014). To replicate the human mind requires at least an accurate and complete description of it as in the Research Domain Criteria project and the BRAIN Initiative (Brain Initiative, 2015). Neuroimaging is here and data collection underpins it. What to do with it shapes what data will be collected and how it is organized. Data collected on us now and of which we have concern about privacy is that reflecting our behaviors. Information on how one acts now can be significant in predicting how the person will act in the future.

There has been advanced a more formidable rationale for wanting to know how a person may act especially in light of increased worldwide terrorism and crime. The Broad Agency Announcement (BAA) BAA 11-65 calls for "New approaches for understanding and predicting the behavior of individuals and

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

groups especially those that elucidate the neurobiological basis of behavior and decision making and the neurobiology of moral judgment, development, and action." Factors motivating the DARPA search include (as stated in BAA 11-65) the need to understand:

- Attitude and habit formation particularly when uniquely influenced by or highly plastic in response to virtual interactions;
- Measuring human propensity to engage in violence against out-group members;
- Mechanisms important to mobilization into violent social movements and groups;
- Pathologies resulting in warfighter accidents, misjudgments and maladaptive behavior. (Broad Agency Announcement 11-65, 2015)

The BAA 11-65 was a first step in researching how that behavior might be preempted. Ultimately, the contract was awarded to Fast Mathematical Algorithms & Hardware.

FMA&H will enable scalability to massive data of their basic analytic database processing system therefore addressing all components of their system. The scalability of their system will be tested and measured on several databases. The first tests will be on hyperspectral imagery where each pixel is a point in 256 dimensions and we have several megapixels. An outcome of the organization is an image segmentation coupled with a related spectral feature sets, integrated with the data morphology, to be achieved automatically with the data viewed as a generic database of size 10 (power of 6) X 256. Other tests will be performed and evaluated for efficiency on massive genetic databases as well as document elections (BAA 11-65 Award, 2015).

Ethical systems are predicated upon what their creators consider what should and should not be done in society. These "shoulds" and "oughts" are translated into the day-by-day standards of behavior, the moral code. Usually, ethical systems are predicated upon value systems or an ethos. These systems may be regarded from either the creator's perspective or the one expected to abide by the system. This consideration is even starker in discussing the ethics of social control. Attempts have been made to produce automated human behavior evaluation software such as Effects Based Assessment Support System (EBASS, 2015) and SENTURION (2015).

That the U.S. Taxpayers are funding research on how they may be controlled, evidenced by BAA W911NF-17-S-0002 (2017), where the Army Contracting Command of the Army Research Laboratory (ARL) is calling for research to

…to support Strategic Land Power Dominance for the Army of 2030 and beyond.… Proposals are expected to be for cutting-edge innovative research that could produce discoveries that would have a significant impact on enabling new and improve d Army operational capabilities and related technologies. (Ibid., p. 1)… The Social and Cognitive Networks programs …supports projects that contribute substantive knowledge to theories about human behavior and interaction and make methodological advancements in modeling and analyzing social network structures (Ibid., p. 54) … Integrated Intelligence, Theory of Mind, and Collective Intelligence (ibid., p55) … identifying and modeling the coevolutionary multiscale dynamics of human neural, cognitive, physical and social systems. (Ibid., p. 60) …develop new theories to understand the dynamic interrelationships between individual/group cognition, decision-making and the role that these influences play on interactions with large and small social systems" (Ibid., p.61).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The U.S. intelligence apparatus has been busy, as well in focusing on the biological aspects of human behavior. Under the Officer of the Director of National Intelligence is The Intelligence Advanced Research Projects Activity (IARPA), inviting researchers to participate in programs, such as:

- MICrONS seeks to revolutionize machine learning by reverse-engineering the algorithms of the brain
- The focus of the ICArUS Program is to understand and model how humans engage in the sense-making process, both during optimal and suboptimal (biased) performance. Of particular interest are cognitive biases related to attention, memory, and decision making.
- The goal of the KRNS Program is to develop and rigorously evaluate theories that explain how the human brain represents conceptual knowledge. In part the evaluation will rest on how well concepts can be interpreted from neural activity patterns using algorithms derived from the theories. In addition to new theories and algorithms, KRNS seeks the development of innovative protocols for evoking and measuring concept-related neural activity using neural imaging methods such as (but not limited to) functional magnetic resonance imaging (fMRI) and magnetoencephalography (MEG).
- The MOSAIC program seeks innovative approaches to unobtrusive, passive, and persistent measurement to predict an individual's job performance.
- IARPA's Janus program aims to dramatically improve the current performance of face recognition tools by fusing the rich spatial, temporal, and contextual information available from the multiple views captured by today's "media in the wild".
- The HFC program seeks proposals for research to develop and test hybrid geopolitical forecasting systems. These systems will integrate human and machine forecasting components to create maximally accurate, flexible, and scalable forecasting capabilities. (IARPA Research Programs, 2018)

The RDOC by the NIMH, BAAs, IARPA, and ethical behavior assessment programmes might be considered a primary investigatory phase of describing how the brain functions. It is said that describing a problem is half the solution. This would be the ultimate of control but control requiring enormous ability to collect data. The technological aspects pale in comparison to the ethical issues. While major scientists encourage research into artificial consciousness, they have warnings.

Each of the above projects – both non-governmental and governmental – is an example of reaction where the investigators respond to how persons and their minds behave. It remains how one can forestall behavior by creating the mind itself.

In January 2015, Nick Bostrom joined the late Stephen Hawking, Max Tegmark, Elon Musk, Lord Martin Rees, Jaan Tallinn among others in signing the Future of Life Institute's open letter warning of the potential dangers associated with artificial intelligence. The signatories "...believe that research on how to make AI systems robust and beneficial is both important and timely and that there are concrete research directions that can be pursued today" (AI Warning, 2015).

There are two directions in collecting data on neuroimaging: individual brains and collective ones. Sociologists like Gustav Le Bon (1896) continued the tradition in arguing that groups of individuals have a collective psyche, one not being the same as an individual. In a true Cartesian approach (Descartes, 1641), one attempts to understand the whole by breaking it down into its constituent parts. If society is an organism, then, we begin by identifying its components: members within it. As the method goes, by understanding how the individuals work, conclusions may be drawn about the whole in an inductive or probabilistic fashion.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Big Data Collected by the NSA on US

Two main US surveillance institutions are the National Security Agency (NSA) and the Department of Defense. Other spy agencies exist, such as the Central Intelligence Agency and Federal Bureau of Investigation, but it can be assumed that if information sharing is not seamless across the full spectrum of the US surveillance network, efforts are ongoing to make this happen (National Infrastructure Advisory Council, 2012). Indeed, a special website dubbed "Intel.gov" refers to "…a multi-agency community" (U.S. Intelligence Community, 2018). The controversial Cybersecurity Information Sharing Act of 2015, S.754 (S.754, 2015, pp. 31-34), was signed into law by U.S. President Obama on 18 December 2015. Companies no longer have to make anonymous data they give to the U.S. Government concerning "cyber threat indicators", which the critics argue are not described in the legislation. If one wants to get somewhat facetious, they could argue that since every major computer operating system has vulnerabilities (National Vulnerability Database, 2015), then there is a "cyber threat indicator" every time someone gets on a computer. Whatever we communicate over the Internet may be deemed a "threat indicator" with all information being monitored by a federal agency. The U.S. Department of Homeland Security says, "Protecting the country from ever-evolving, transnational threats requires a strengthened homeland security enterprise that shares information across traditional organizational boundaries" (U.S. Department of Homeland Security - If, 2018). It is intriguing to note that the DHS phrase "If you see something, say something" is formally registered trademark with the U.S. Patent and Trademark Office (U.S. Department of Homeland Security – Information, 2018). How can anyone miss the parallel between dictatorial spy agencies like the Gestapo and Stalin's NKVD encouraging citizens spying on each other?

Representative of the potential if not the actuality of governments collecting data on their populations is the U.S. National Security Agency (NSA) Utah Data Center (formally called "Intelligence Community Comprehensive National Cybersecurity Initiative Data Center") at Camp Williams near Bluffdale, Utah, between Utah Lake and Great Salt Lake. When one searches for "Utah Data Center" a web page entitled "Domestic Surveillance Directorate" appears with a mock-up but realistic insignia to the left (Utah Data Center, 2018), the astute observer of history being reminded of the vast repositories on information kept by the Gestapo, KGB and Stazi.

A number of verifiable facts are presented, peppered with links to some official U.S. Government website, as well as a major link entitled, "IC off the Record" (2018) with its banner headline, "Truth is coming and cannot be stopped". To the left of "Utah Data Center" page are numerous links to legitimate sources, such as news releases, descriptions, articles, and so forth. At the very bottom of this very official-looking page is:

This is a parody of nsa.gov and has not been approved, endorsed, or authorized by the National Security Agency or by any other U.S. Government agency. Much of this content was derived from news media, privacy groups, and government websites. Links to these sites are posted on the left-sidebars of each page.

However, the imitation is more real than one might think ostensibly. What is not a joke is The U.S. Army Corps of Engineers "project specifics":

- 65 Mega Watts Technical Load to the Raised Floor
- Average load density: 650 watts/square feet (sf)
- Maximum load density: 1600 watts/sf

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- 100,000 sf of Raised Floor in 4 Data Halls
- Tier III Data Center Standard, Concurrently Maintainable
- Flexible Cooling to Data Hall IT Rooms: Water and Air
- Air: 250 watts/sf
- Water: 1600 watts/sf
- Clean Agent Fire Suppression for Data Halls
- Redundant Utilities, 100% Generator & UPS Backup
- Mechanical & Electrical Support Structures
- 2 Substations
- Goal LEED Silver, Other Green Technologies Stressed
- Water usage at full load is approx. 1.7 million gal/day
- Thermal energy storage tanks
- 3 days of fuel storage
- Piping distribution will be primary and secondary
- Secondary pumps will be on UPS
- Cooling towers shall be Cooling Technology Institute
- (CTI) certified or built (Ferguson, Lange, and Capenos, 2009, pp. 10-11)

"Tier III" is a classification created by The Uptime Institute, and recognized as a computer industry standard. It means "Concurrently Maintainable data center has redundant capacity components and multiple independent distribution paths serving the critical environment. For the electrical power backbone and mechanical distribution path, only one distribution path is required to serve the critical environment at any time" (Uptime Institute, 2018, p. 6, Section 2.3) (see Figure 7).

Some other actual facts about the very real Utah Data Center are its:

- Million and a half square feet overall "about a third larger than the U.S. Capitol and the House and Senate office buildings"
- Yottabyte capacity storage (one thousand trillion gigabytes of data)
- $1.2 billion cost
- 200 acres
- 900,000 square feet of technical support and administrative space.**"** (Building size, of course, is not the only factor in considering data storage capabilities; it is what can be placed inside. What is important concerns the ability to store massive amounts of data in such a small space. To illustrate, an average 15 × 15 foot room eight feet high is 259,200 cubic inches, each inch representing the size of a thumb drive. Above was described some of the latest advances in technology, where there now are two terabyte thumb drives. At this point, one can allow the imagination to do the rest in wondering how much data can fit in to a few hundred rooms of this size given what was said above about data storage capabilities. To conceptualize proportions, the NSA Utah Data Center mission critical area may contain 444 of these 15 × 15 foot rooms (100,000 square feet). (UTC data, 2018)

One can perseverate on the Utah Data Center indeed existing and be amazed over the physical construction, but the question is "why", outside the propagandistic reasons of "national security", etc. To uncritically believe NSA's 2013 press release:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 7. Tier requirements summary (Adapted from Ibid., p. 9)

	Tier 1	Tier II	Tier III	Tier IV
Minimum Capacity Components to Support the IT Load	N	N+1	N+1	N After any Failure
Distribution Paths-Electrical Power Backbone	1	1	1 Active and 1 Alternate	2 Simultaneously Active
Critical Power Distribution	1	1	2 Simultaneously Active	2 Simultaneously Active
Concurrently Maintainable	No	No	Yes	Yes
Fault Tolerance	No	No	No	Yes
Compartmentalization	No	No	No	Yes
Continuous Cooling	No	No	No	Yes

Many unfounded allegations have been made about the planned activities of the Utah Data Center. One of the biggest misconceptions about NSA is that we are unlawfully listening in on, or reading emails of, U.S. citizens. This is simply not the case." (Reuters Staff, 2013)

would be the height of naiveté, in light of the ongoing provisions of the Patriot Act, DHS, and what the Electronic Frontier Foundation has to say, all of which can be learned through the Internet and by what has been said to this point.

One also should reflect on the scope of the "Officer of the Director of National Intelligence (emphasis in original) (DNI, 2018). This is left as a research exercise.

As somewhat of a sidebar, to read the original Fed Biz Opps (Federal Business Opportunities) solicitation W912DR-10-R-0015 one accessing the listings even to read the contents of the solicitations and awardees has to be a registered contractor (Fed Biz Opps, 2018).

Big Data Collected by the Military

Aside from the vast empire of the Department of Homeland Security lies the U.S. Department of Defense (DoD). We have discussed its Broad Agency Announcements calling for research on human behavior, but it also has been focusing on modeling and simulation (MSCO, 2015), or automated war, an outgrowth of which is Network-Central Systems, referred to commonly as "Net-Centric Warfare" (Wikipedia, 2018m). As an affirmation of what already has been said about systems and modeling and simulation as the generator of big data sets, the following is presented from the U.S. DoD Modeling and Simulation Coordination Office, *Modeling and Simulation Body of Knowledge* (MSBOK, 2015):

A system is a collection of components organized to accomplish a specific function or set of functions. System types are: (1) units; (2) weapons; (3) platforms; (4) sensors; (5) life support and (6) Command, Control, Communications, Computers and Intelligence (C4I). Systems are further categorized by operating environment which is divided into five areas: (1) individual; (2) air; (3) sea; (4) space and (5) ground. (Ibid., p. 6)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

A model is a physical, mathematical or otherwise logical representation of a system, entity, phenomenon or process. (Ibid., 2015, p. 6)

Human representation refers to the use of a computer-based model within a simulation that mimics either the action of a single human or the collective action of a team of humans. Human behavior representation can model any of the complicated facets of human behavior including ability to reason, ability to change the environment, reaction to comfort and discomfort, susceptibility to injury and illness, emotional responses, communication with others, ability to sense the environment and physical capabilities and limitations. (Ibid., p. 6)

A substantial rationale for identifying systems, creating models of them and testing them through simulations is to obviate the large costs of creating and seeing how they work in real life. The DoD's reasoning also is that, "The need for taking people out of dull, dirty and dangerous missions is a primary motivator for UAST [Uninhabited Autonomous Systems Test]" (UAST, 2015, p. 5). With UAST in mind, we now are ready to consider the larger context of big data, also focusing on the word "intelligence" in all of its forms – natural, artificial, and the very words like "intelligence gathering", "Director of National Intelligence", and out of its office, "Intelligence Advanced Research Projects Activity" (IARPA, 2018).

BIG DATA AND FOG COMPUTING FOR SOCIAL INTELLIGENCE

The Relationship Between Big Data and Societies

To summarize our discussion to this point, one wants ultimately to get knowledge about their environment. Beforehand, data needs to be collected and made sense of by converting it to information. Such information needs to be evaluated as to quality to produce knowledge. Then, the information is vetted and accessed *vis a vis* the repositories to the recipients with fog computing. In order to do any of this, one needs an organizational framework that comes through analysis or subdivision of the environment into constituent parts that are interactive, processing the environment and producing outputs which in turn feed back into the environment and its participants all in the form of what formally is regarded as a system. A system operates in an environment such as medical, social or educational. A particular system is a model. To study how a model acts, one identifies hypothetical situations or examples and observes how the model behaves, i.e., simulates a situation.

All of this is a formal way of generating big data. Yet, we have been discussing big data as an object – its size, who collects it, and even why it might be collected. NSA collects data to "...monitor, strengthen and protect the nation." IARPA collects it to understand human behavior better and possibly control it. Modeling and simulation projects (SimSoc and Second Life) allow people to live out their fantasies in a virtual world, and government projects want to see what happens when variables are manipulated so as to at least think about prescriptions for societies.

So, back to the fact that prompted this discussion about systems in the first place, one should be reminded about the nature of the entity that would want to do this. We can put the gamers aside and focus on those thinking about real-world uses and applications, i.e., governments. Governments are focal points for social decision-making, and it can be argued that societies, themselves are organisms. This next section traces the development and nature of modern societies, those same societies that are generating

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

big data, that data reflecting more about the nature of the society generating it. Rather than having been more passive up to this point – social institutions collecting and using big data – that big data is part of a larger scope involving the very nature of these societies and what they do. More explicitly and in finding a common denominator of big data collection, that collection centers about the nature of thinking, mind, intelligence, consciousness, and so forth and finding ways of directing, shaping, or even re-creating it.

Following is a model for discussing the "why" of data in context. In essence big data is analogous to a memory and in the computer world that data is stored electronically. However, so goes the organization of society, so follows technological structure. One may think of a snake skin or animal armor; as the creature grows, so does its outer surface. Hence, we should examine the societal analog of memory and data storage, but to do so we need to identify the locus. In the organism, it is in the neurological structure, the more complicated loci being the brain.

Parallel to the emerging structure of data generation, storage, and distribution is the organization of society. Such organization is not only mechanical but is dynamic, as the parts of a body have their mechanical aspects, but each being an integral part of the whole corpus that embodies them. However to study these entities requires an approach, a discipline, better situated than the current political or social sciences.

What is this to be a science about? It is not enough to say "a study of power", "politics", or "societies". What is about these things that would warrant the word "science"? Extrapolating from the past to project to the future as the foundation of science is not done very well by those purporting to be social or political scientists. This section sets forth a framework for making sense of how big data relates to a science of societies. Treating societies as organisms – dynamic entities – will help us better answer questions like "How could an institution such as the NSA, come about, despite the long history and tradition of mistrusting 'the government'?". Data is collected not as a recreation but for a specific and usually scientific purpose: discerning patterns from the past and present so as to be able to predict and manage the future. The NSA data is about us and the society and unless we see its collection in the context of the political, the society, it won't make much sense. It is time for those studying social and political science to realize this in a modern way. Let's start by looking at the foundation of societies, focusing upon their being special dynamic systems, in particular, organic ones.

How Societies are Organized

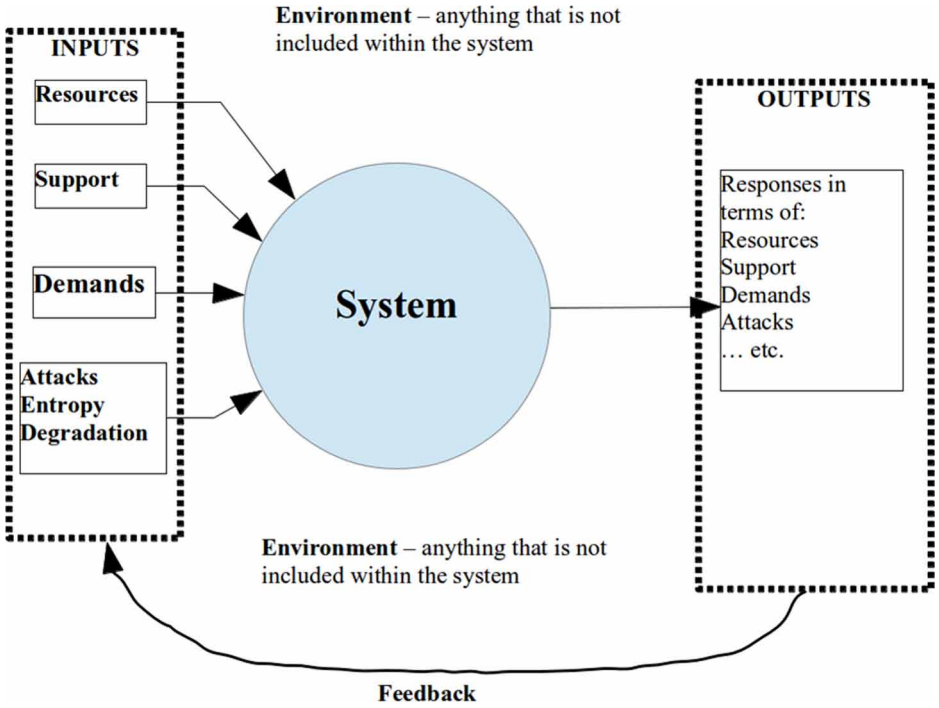
Late in the 1960s and mid-1970s, cybernetics was applied to social phenomena as systems analysis. Organizational behavior is predicated upon the dynamics of structures as systems. A system is a collection of elements in which each element bears a relationship to the other and the collection of elements has a purpose (Buckley, 1968, p. 430; Callaos, 1994, p. 3). In perspective, on a larger scale, examining the behavior of a collection of systems and systems of systems is in the realm of complexity studies and at this level big data becomes explosive. The explicit rationale for the systems analysis method as opposed to any others is that it follows directly from basic logic which is inherently a bootstrapping method. Logical proof allows one to posit assumptions within a set of definitions, rules and so forth and draw conclusions. Otherwise put, various elements are arranged according to a scheme-a system-where there are defined inputs, outputs and various processes. The theoretical underpinnings of a system applied to a particular set of circumstances produces a model, that model being tested with instantiations with simulation.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

As Social Contract philosophers (e.g.: Hobbes and Rousseau) discussed societies as organisms starting in the 17th century, modern systems analysis identifies social components collected into functioning entities living in environments, intaking, outputting, and processing factors, either maintaining themselves or adapting (Buckley, 1968; Wikipedia, 2018r). For example, we may consider a general systems as a framework into which data can be related, i.e., a form of data visualization, a context diagram, Figure 8.

The system takes demands and supports as inputs and processes the outputs as decisions and actions, which are processed by the environment and fed back to the system. Outputs, decisions and actions of authorities (Buckley, 1968, p. 434) are consequences of the behavior of system members. How the system processes inputs is a function of its ability to adapt. A political system must be able to allocate values and be able to have members accept those allocations of values. Stress and tolerance for it determines system viability. A cybernetic approach to systems theory in political science refers to an organism (system) having a behavior, a condition requiring consciousness, mind and psychology. Indeed, the expression "body politic" alludes to this. The word "corporation", a basic social unit, is about the "corpus", or body. It may be argued that, while there is no homunculus (as in an entity coordinating or being the center of consciousness), there is a center of organization, i.e., decision-making bodies, as in the three branches of government – legislative – executive – judicial (analogous to the brain of a body), where each of the other system's constituents (as in the electorate) have no sense of the organism's overall motion, direction or organizational-level consciousness. Such would explain why that long tradition of individualism and mistrust of government does not translate into coherent widespread opposition to institutions like the NSA. Somewhat analogous is a muscle exhibiting pain upon exercise directed by the brain, but the muscle, itself, is powerless to countermand the directive, except by breaking down.

Figure 8. A context diagram of a generalized system (Courtesy of the author, Jeremy Horne – drawn especially for this paper)



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

A key feature of systems as observed by Ross Ashby (1954) is that most are dynamic because they have to be adaptive to survive. How does an ecosystem fare with the arctic as a model simulated by polar bears affected by climate change? Note that one may change the granularity of analysis by studying the population of polar bears as the system modeled by their being in a specific location and this model being simulated with a particular month. There is minimally a primitive organicity or organic character to a system, each part of the system articulating in concert with another to achieve the overall aim of survivability, adaptability or improvement.

The Nature of Organicity

All organic things are systems, but not all systems are organic. It is the "organic" aspect of societies that helps us place big data in better perspective. Yet, the question is begged: "what is organic?" An organism is seen as a dynamic entity that depends for its existence on parts that interact with each other and cannot exist independently of that entity. Note that "organicity" (having the characteristics of an organism) has to do nothing with the material out of which an organism may be made, be it a hydrocarbon or not. "Organic" rests upon the concept of what "life" means, and this question by no means has been finally resolved. However, and for our purposes and pulled from myriads of sources the following commonly are listed as criteria:

- Life has complex patterns.
- A living thing has metabolism (processing inputs as food to provide sustenance to sustain itself)
- Something living can interact dynamically with the environment
- Killable
- Able to maintain itself (or even adapt) despite environmental changes or even modify the environment.
- All the parts depend upon each other for their existence.
- The entity is able to reproduce itself.

Searching for Erwin Schrödinger, Carl Sagan, John Casti, and Alan Turing, among others, will enrichen the discussion, this being left as a research exercise. By no means is there any universal agreement about clear-cut lines between what life is and what it is not (Trovonov, 2012; Weber, 2018). Now, let's couple systems and organicity to discuss the highest form of social evolution: the State.

The State and its Models

Plato thought in his Republic that the society is an organism carefully describing how each part should be qualified and function.

However, in founding the city we are not looking to the exceptional happiness of any one group among us but as far as possible, that of the city as a whole. (Plato, 420b, 1968) Have we any greater evil for a city than what splits it and makes it many instead of one? Or a greater good than what binds it together and makes it one? (Republic 462A, 1968)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Thomas Hobbes' Leviathan (a Biblical sea monster) writes in the Introduction of "Nature (the art whereby God hath made and governs the world)," and from this of societies not only being organic but artificial:

Nature (the art whereby God hath made and governs the world) is by the art of man, as in many other things, so in this also imitated, that it can make an Artificial Animal. ... Art goes yet further, imitating that Rationall and most excellent works of Nature, Man. For by Art is created that great LEVIATHAN called a COMMON-WEALTH, or STATE, (in latine CIVITAS) which is but an Artificiall Man; though of greater stature and strength than the Naturall, for whose protection and defence it was intended; and in which the Sovereignty is an Artificiall Soul as giving life and motion to the whole body; The Magistrates and other Officers of Judicature and Execution, artificiall Joynts; & …. (Hobbes, 1651, Introduction; spelling and emphasis in the original)

Rousseau thought of human societies being organic as well:

At once, in place of the individual personality of each contracting party, this act of association creates a moral and collective body composed of as many members as the assembly contains votes and receiving from this act its unity, its common identity, its life and its will. This public person, so formed by the union of all other persons, formerly took the name of city,1 and now takes that of Republic or body politic; it is called by its members State when passive, Sovereign when active and Power when compared with others like itself. Those who are associated in it take collectively the name of people and severally are called citizens, as sharing in the sovereign power and subjects as being under the laws of the State. (Rousseau, 1761, p. 44)

Perhaps by far the greatest development of a coherent idea of the State was Georg Wilhelm Friedrich Hegel (Hegel, G.W.F., 1833/1896/2001) in his *Grundlinien der Philosophie des Rechts, (Philosophy of Right)* ("§" references are section numbers locatable in any edition).

The state is the realized ethical idea or ethical spirit. It is the will which manifests itself, makes itself clear and visible, substantiates itself. It is the will which thinks and knows itself, and carries out what it knows, and in so far as it knows. (§257, p. 128)

[The State has] a higher authority, in regard to which the laws and interests of the family and community are subject and dependent. On the other side, however, the state is the indwelling end of these things, and is strong in its union of the universal end with the particular interests of individuals. Thus, just so far as people have duties to fulfil towards it, they have also rights" (§261)

...Political disposition is given definite content by the different phases of the organism of the state. This organism is the development of the idea into its differences, which are objectively actualized. These differences are the different functions, affairs, and activities of state. By means of them the universal uninterruptedly produces itself, by a process which is a necessary one, since these various offices proceed from the nature of the conception, …(§ 269, p. 136)

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

After Hegel and then Darwin, the Victorian period, represented by Herbert Spencer descrying the social state (Spencer, 1896, p. 608, et. seq.) and Émile Durkheim (Durkheim, 18982, p. 123), as in his reference to social organism, and "For what else essentially constitutes the body social if it is not the social space) (Ibid., p. 203)", brought forward more the view of society as organic and evolving with a collective consciousness, just like one of Darwin's organisms (Organic Theory of the State, 2015).

Most relevant is Durkheim's (1858 – 1917) (1893/1984, p. 77) saying:

All the functions of society are social, just as all the functions of an organism are organic (Ibid., p. 81).

[C]o-operative relationships do not carry with them any other form of sanctions. Indeed, special tasks, by their very nature, are exempt from the effects of the collective consciousness. This is because if something is to be the object of shared sentiments, the first condition is that it should be shared, that is, present in every consciousness, and that each individual may be able to conceive of it from a single, identical viewpoint (Ibid., p. 82).

In the end this law plays a part analogous in society to that of the nervous system in the organism. That system, in effect, has the task of regulating the various bodily functions in such a way that they work harmoniously together. Thus it expresses in a very natural way the degree of concentration that the organism has reached as a result of the physiological division of labour.

*Yet if from this comparison it is legitimate to induce, with some degree of probability, that economic functions are not of a kind to be placed under the immediate influence of the s***ocial 'brain',** *it does not follow that they can be isolated from all regulatory influence. (Ibid., p. 164) (emphasis added).*

We will return to the "social brain" later, but suffice it to say that these smatterings of writing about society being more coherent than a loose assemblage of entities and at least analogously a living being illustrate more profound thinking than most politicos realize. Suffice it to say, we rarely see any philosophy of society like this in mainstream public debate. It is a worthy research exercise and learning experience to read these (and many more) thinkers who proposed that societies are organic. Phrases like "society as organism", "superorganism", "organic societies" and related phrases will open up a new world to those accustomed to the prevalent thinking that social arrangement are just mechanical ones. For example, regarding society as an organism implies that human being are integrated and interdependent, the more harmonious the greater chance of survival and growth. Indeed, the organizational model of governments and economic units is the corporation. A recent U.S. Supreme Court ruling (Citizens United vs. Federal Election Commission, 2015) concluded what had been assumed under this model, corporations are persons, complete organic units. Sociological data collection, like that in other systems areas, concerns describing organic entities and their components including us.

Characteristic of the Social Organism

Whether this direction towards a central authority collecting so much detailed data on us has a nefarious purpose is not immediately material. The effort is coherent and the effect of the coalescence of these events/factors shaping that collection may result in the same as if it were intentional. For argumentation's sake and to make this controversial (a nature of philosophical inquiry) discussion more dynamic, one

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

may consider what the most elite of policy makers potentially may think about, an argument in behalf of the Devil. Such a discussion technique follows an accepted scientific methodology of attempting to disprove the null hypothesis (Popper, 1959) or that in logic of the indirect proof, the common denominator of each being the attempt to develop a contradiction. What might policy makers contemplating social control from an elitist perspective – hypothetical or intentional-consider in their collection of big data and its management? Let us look at what they are facing and why the organic model makes the most sense in identifying and solving social problems.

Challenges to the State as Organism

Set before societies is unreservedly an existential challenge to which traditional responses most likely will not be adequate. Much has been written about increasing adverse complexity brought about by exploding populations, global warming, environmental degradation, and greater technology (Homer-Dixon, 2007; Tainter, 2003). Because of these as major factors, coupled with others, scientists are now discussing the "Sixth Great Extinction", the question of whether *homo sapiens sapiens* will even be around as a subspecies in a hundred or so years. A cursory search under "sixth great extinction", or "Holocene extinction" will display hundreds of entries, many of them peer-reviewed scientific articles on the subject; these are recommended reading (e.g. Ceballos, 2015). Organizations like the World Wildlife Fund, Post-Carbon Institute, The Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES), and National Geographic Society have numerous pieces on the subject, those being retrievable by a simple search in their websites. So goes the species, so goes its organization(s).

This species cannot wait for the wheels of participatory decision making with all its infighting to meet the challenge (Kintisch, 2015). It is incumbent on anyone valuing the continuation of this human species and its environment to research the Sixth Great Extinction and think how a social consciousness resting upon the bedrock of big data might address. That there does not exist the State is reflected by governments meaningfully and totally assuming ownership of problems that are existential in nature. Let us focus on some. This discussion is needed to help provide a larger context within which big data is collected and used, each problem having its own data set. The nature of the society in which we live is a major determinant of that collection of data, its applications, and how we relate to that use.

Even with populations exhibiting passion for their fellow members of this species, indifference, tolerance of events, desensitization and gradualism (Lifton, 1986) allow mass suffering and death. There is a growing burnout of social caregivers that contributes to social paralysis in front of danger (Pace, 2014).

Given that so many people are getting on line with very affordable computers and intelligent communications devices, security has become such a highly complex issue that there may not be any real protection but only management of an attack that is seen as inevitable. Protective measures are chimeras not unlike whack a mole. As soon as one barrier is erected, the monster is inside the compound again. This problem is exacerbated because people not only have access to the computer but they are increasingly dependent upon its integrity in their daily lives. Millions of persons use antiquated software that was deficient to begin with and they do not do basic maintenance mainly because of being ignorant (Microsoft, 2015). Society, itself - education, commerce, government, etc. - also depend upon that integrity for basic functioning.

That computer security is a growing threat to communications stability (Computer Security, 2018; U.S. Cert, 2018; Sans, 2018; CERT, 2015; Wikipedia, 2018L; SC Magazine, 2018), may be symptomatic of a larger problem of how a growing population responds to material (read "economic") challenges. A

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

prison guard knows all too well what an inmate can do with limited resources in fabricating contraband (Contraband, 2018). We now are witnessing what a population can do unfettered and with the same sociopathic mentality. While the one percent of the U.S. Population being sociopathic (NIMH-sociopathy, 2018) seems to be a low figure, it translates into 3.2 million persons (U.S. Census, 2015). This says nothing of at least 11 million others under severe mental duress (NIMH Results, 2018).

Given this mentality and a question of whether a large part of the population lacking ethics (Cattermole, 2010; Jordan, 2015; Ceballos, 2015; National Geographic, 2006) or intelligence (as in the IQ bell curve), we see more extremist movements and gangs that seem to be capable of meeting the consequences of alienation (Weber, 1992). It may be argued persuasively that the lack of employment, recreational facilities, social organizations or other institutions that traditionally have socialized individuals has been a main contributing factor to the social entropy (DeAngelis, 2009). Increasing millions are under psychological duress (NIMH Results, 2018) as the 353 incidents of multiple shootings (alone) would seem to indicate (Multiple Shootings, 2018).

More controversially is the question whether this subspecies has the intelligence and will enough to manage the onrushing complexity so as to save itself from extinction. Allowing the average to make critical decisions could be fatal. Considering this means critically regarding perhaps the most cherished myths of modern times, that of "democracy".

Despite all the talk about "democracy" and its being the rationale for so many modern wars, the backdrop of Plato and Aristotle has persisted throughout time. Plato in his *Republic* and Aristotle in *Politics* argued throughout that a participatory society (referring then explicitly to democracy) cannot last without a substantial educated middle class. One might add that the electorate needs to be ethical and be willing to participate in decision making. The classic argument is that the uneducated poor will easily be swayed by demagogues, be more prone to settle problems through violence (not having the capacity to think critically), and be susceptible to financial rewards in return for support. Persons not having a basic knowledge (Alsop, 2014), of even the geography where the contemporary affairs occur (let alone the affairs, themselves) will not be able to select their representatives intelligently. A knowledge of basic world events aside, more egregious is that one in four of Americans thinks the Sun goes around the Earth and only half realize that antibiotics are ineffective against viruses (National Science Foundation, 2014). It may be asked, "How can leaders and policy makers be selected intelligently those who will be faced with complexities far exceeding knowledge equal to that of an eighth grade student?" Recent world events seem to illustrate that social upheavals against dictatorships rarely culminate in participatory governments, the Middle East after the so-called Arab Spring being a case in point. There is enough "big data" to substantiate the thesis that violent approaches to problem solving are positively correlated with ignorance and low critical thinking skills. However, one must be intelligent to be educated and use those skills.

While substantial literature exists on multiple intelligences (Wikipedia, 2018t; Gardner, 1993; Gardner, 2004), one may ask whether the average person (average IQ) can understand today's complexities let alone manage them successfully (IQ, 2018). A persistent argument exists (Wikipedia, 2018b) that if society is composed largely of persons of average or lower than average intelligence, as reflected in a popular book, *The Bell Curve*, policy should be made only by intellectual elites (Hernstein, 1994). It is beyond the scope of this paper to delve into the IQ debate but suffice it to say, reams of peer-reviewed literature have appeared on the subject as an Internet search under "IQ" and similar terms will indicate. There is ongoing argumentation about whether there has been a peak in the percentage of persons attaining quality education. In 1983, The Carnegie-Mellon Institute issued *A Nation at Risk* (National Com-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

mission, 1983), describing a National security risk existing because of the generally poor educational attainment of the US population. In the middle of the second decade of the 21st century, matters have not changed significantly. According to the U.S. Department of Education in 2013 – 30 years later-only 39% of high school students were academically prepared for college in mathematics and 38% in reading (DOE Statistics, 2013). If data from organizations like the Literacy Project Foundation are credible, concern is raised about 50% of U.S. Citizens being functionally illiterate, i.e., reading below the eighth grade. (Literacy Project Foundation, 2018). Even being able to read at a minimal level does not even begin to address serious issues like evaluating information for its quality (including identifying good sources of information), let alone critically thinking about it. The National Science Foundation surveys referenced above result partially from illiteracy but a question remains why persistently so many persons are illiterate let alone knowledgeable in basic facts.

Such deficiencies have major policy implications on an international scale, where ignorant voters select policy makers with the power ultimately to press buttons on nuclear weapons systems. It may be argued that the pattern of world events has not changed since ancient times with the exception of the capacity to destroy being greater and radically more complex technology being able to magnify human error especially beginning with World War One and now with vast nuclear weapons stockpiles. Societies have become enormously complex with even the best and brightest seemingly making questionable judgments (atomic bomb, bioweapons, and fossil fuels).

These may all be factors besetting the average and below average person, but as systems become even more complex, one may ask about the increasing portion of the more intelligent being drawn into this group, Dr. Theodore Kaczynski, the Unibomber, being case in point. At some point, even those with an above 150 IQ (or whatever standards are used to gauge intelligence, multiple and qualitatively different IQs included) may be incapacitated in managing social complexity. In an extreme case, one only need look at the *Einsatzgruppen* (Nazi death squads), where "Of the 25 *Einsatzgruppen* and *Einsatzkommando* leaders, 15 of them bore the title of PhD, most of them doctors of jurisprudence or philosophy. (*Einsatzgruppen*, 2006)." One asks the same of the Nazi doctors (Lifton, 1986).

Again, the above argumentation may not be necessarily sound, let alone valid, but one needs to consider it as a possible motivation for why data is being collected as BAA-1165 and the various intelligence websites illustrate. If the population cannot control itself even to the point of protecting its own environment, the question is raised, who will? For the present one may refer to the technocratic movement as a response to these problems (Technocracy, 2018), especially in considering such big data laden projects as those of NSA, NIMH, the BAAs and brain modeling.

What are the Implications of Fog Computing?

It was mentioned that data can and often is acquired from sensors, perhaps the most personal example being those placed inside of a pill-sized object and the whole assembly being swallowed in order to report on what is happening inside the digestive tract. Typing in "microsensors", "miniature sensors", "radio frequency sensors", and similar phrases should prompt the imagination in envisioning how detailed data collection might be (Minin, I., 2011). On the outside, radio frequency (RF) tags can be affixed to just about any object, with that object being tracked. RF tags now are being embedded in the skin for various purposes. It is not a question, then, of whether objects can be tracked and data collected on them. Where does that data go, who manipulates it, who manages it, and why? A number of situations were alluded to above, which can be the core of lively debate, but what of the future?

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

What applies to cloud also applies to fog, as the latter merely is the interface between humans and big data. One may think of peripheral nerves and appendages in the form of drones, IoT devices, phones, etc. We saw how the The National Institute of Standards and Tchnology (NIST, 2017) fog computing diagram resembles neuroarchitecture. Hence, we can see in a society as organism model big data as memory, the "social brain" using it, and fog computing as a way of individual cells receiving and sending via the social brain big data.

If societies can be regarded either literally or analogously as organisms and computer technology in the form of data collection and management and artificial mentation parallel the organism in themselves assuming an organic character, we now see ourselves as at least technologically in the completion stage, where fog computing providing that last segment of detail to enable complete manipulation of the technological organism reflecting the society. As the organism takes on its emergent characteristics, so does the technological ensemble.

Towards a Solution

If perceived deficiencies of the average individual and the societies in which they live are some of the motivations driving social control, what, then, is to be modeled? The above discussion on simulated societies may offer a starting point in that modeling and simulating human interactions in what amounts to a social laboratory can reveal roads to possible solutions. It has been said that principles of organization underpin the utility of large data sets.

Collecting data and identifying it as information about physiography, thermodynamic interactions in the atmosphere and precipitation levels are in the context of a larger framework of meteorology, each area of investigation articulating with another to describe a system. Data sets reflecting a description of DNA, diseases and the make-up of organisms are within the context or framework of promoting the health of those organisms as systems. For humans living together on this planet, it is the society that has been regarded as an organic system having a health. A reading of the literature of organic thinkers, Hegel and Durkheim in particular, will present a viable way of diagnosing and treating social and political problems. Durkheim's "social brain" is foundational.

For thousands of years humans have created tools to do what they could not alone. Simple machines (levers, wheels, inclined planes and so forth) are combined into complex ones. From the science that provides the understanding of our world comes the technology to manage it. Computers are tools to memorize, compute and analyze. Now, we have a society confronted with problems threatening species survival that even the best and brightest individually or collectively may not be able to solve, as Homer-Dixon and Tainter (discussed above) have alluded. To this point we have described a repository of information (data storage), methods of collecting that data, and some of the uses to which it might be put. The next step is connecting the myriads of parts of society – its members and technology to that repository and its manipulators. A clue is with cybernetics and scientists becoming involved their "scientific study of control and communication in the animal and the machine." as stated by Norbert Weiner (1948). But, first, we need a science of society and its subdomain, politics.

The New Social and Political Science

Social and political use of big data generally re issues normally under the purview of sociology and political science, but they remain only as issues, rather than being viewed in a wider context of the nature

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

of the society, itself, that society assuming a particular character that can help explain the larger role that big data is playing in our lives.

Political science has been thought to be a misnomer since time the term was presented formally in 1752 by David Hume, who wrote that:

… politics may be reduced to a science So great is the force of laws and of particular forms of government and so little dependence have they on the humors and tempers of men, that consequences almost as general and certain may sometimes be deduced from them as any which the mathematical sciences afford us. (Hume, 1742)

Aristotle has been credited as the founder of political science (Aristotle and Political Science, 2015), and Aristotle's teacher, Plato, wrote extensively about the thinking behind and outcome of social behavior in the form of constitutions. However, the words "political science" do not appear in any of Plato's or Aristotle's works. The study of peoples' (demos) institutions polis (city) as a result of human behavior (ethics) was not regarded so much as a science as a philosophy or a way of thinking. More precisely politics was about how power is acquired, distributed and used, power being the ability to make another do one's will (What is Political Science, 2015). One of, if not the modern founder of the systemic studies of politics under the rubric "political science" was Henry Baxter Adams in 1880 in establishing the "Johns Hopkins Studies in Historical and Political Science" (Adams, 2015).

Years of debate ensued about how much of a science was there of politics. After all, science could be listed in their order of exactitude and value in predicting: math, physics, chemistry, biology, psychology and the so-called social sciences including political science (Political Science History, 2015). For the most part any technical rigor to the field has been in collecting and analyzing data from survey research, but because these surveys are behaviorist-oriented, with those surveyed being highly ephemeral in their views, precision might be regarded to that on par with predicting the outcome of a horse race. The recent attempts to prohibit the U.S. National Science Foundation from funding political science as in the failed Coburn Amendment SA 2631 (2015) illustrate this.

What may be giving force to a more legitimate study of societies as a science is a model that has been developing for millennia, one of the State being an organism, where diagnosis and treatment of problems would be analogous to medicine a thematic approach, more coherent that the cacophony of what exists now. Overlain on this is a focus on identifying the "social brain", asking the pertinent question of whether its collective mentation is sufficient to meet future challenges, especially the Sixth Great Extinction. Heightened big data capacities, as well as more sophisticated management are at issue. It may be added, too, that political philosophy is conspicuously absent in political and social science departments, the preference being for mechanical descriptions and piecemeal Band-Aid type remedies for problems.

Political science (and social science) to be a science must follow scientific methods, but in contemporary times, with an increasingly interdependent and diverse world, it needs to be interdisciplinary. This is to say that political and social science students need to address not only what they are studying now but other fields, like biology (focus on organicity), artificial intelligence, neuroanatomy, and perhaps most important philosophy (consciousness studies, logic, and ethics) if this discipline is to have any thematic unity and coherences oriented to organic societies and social brain.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

FUTURE DEVELOPMENTS AND RESEARCH

What might the ultimate meaning of big data be? We have seen "data points" above that indicate future technological development, the primary areas being in data storage capacity and processing speeds both addressed by quantum computing. To gain some perspective on how rapid technology can evolve, look at ENIAC, the vacuum tube-based computer sponsored by the U.S. Army in 1943, where in a large room the 30-ton machine of thirty separate units relied on forced-air cooling used 19,000 vacuum tubes, 1,500 relays and hundreds of thousands of resistors, capacitors, and inductors, and required about 200 kilowatts of electrical power. It was a complex operation for it to do square roots and long division (Weik, 1961). It is almost impossible to comprehend what may happen in the next 70 or so years given our moving ever so closer to replicating the human brain.

Yet, one should track development of big data where there is the greatest complexity in organization and management as well as in the areas of social control. In answer to the problems of democracy, referring back to Plato and Aristotle, policy makers may be thinking about developing tools to accomplish what humans alone cannot do among other things getting humans to live peacefully with each other. Big data is all about complexity both concerning the project for which it is collected and the big data itself in terms of collection methods, organization, interpretation and management. Somewhat of a paradox is emerging in that the level of complexity either natural (before any human intervention) or of humans' own making-far outstrips the ability of humans to deal with it. It was stated that the ability to replicate the universe is questionable at best. At some juncture, there is a crossover point, where after so many Cartesian subdivisions, the emerging complexity simply from a big data standpoint may be impossible to record, let alone be managed. We are out-analyzing ourselves, as it were. What emerges very likely will be something unpredicted.

We have seen how the discussion about big data started from a rather simple description of large individual systems, such as those involved in weather modeling, could be seen in larger contexts, leading into data about modeling the human brain and perhaps consciousness, itself, culminating in modeling all of society, inductively described by what happens with individuals. For this last, consideration is given to developing the ultimate tool, Durkheim's social brain, but that brain being human-made. This is a logical outcome of transhumanism (to be researched) and Kurzweil (2018), where, starting from simple prosthetics humans gradually are transformed into non-hydrocarbon beings.

There can be a science of societies and politics if there is an interdisciplinary program, not only consisting of historical and descriptive accounts of societies and politics but of scientific research methods, biology (as in focusing on what life is), neuroimaging, artificial intelligence, and above all, philosophy to impart the whole realm of study with meaning. At its foundation are social groupings as organisms, each organism having a "brain", each brain as a repository of "big data", with various "programs to manage it, but overall with a "high-road" ethos as the motivation for sustaining the organism.

Given the discussion about whether Earth can support this species in the future and much more long-range, the ultimate demise of the universe, thinking has emerged about reproducing humans digitally, the complete elimination of hydrocarbon-based life, big data far beyond creating an artificial brain. This digital immortality was proposed by Frank Tipler in his 1994 work *The Physics of Immortality* (Tipler, 1994). It may be pseudoscience or not but as with science fiction inspiring real technological development, this Ph.D. Physicist's work might have the same effect (Digital Immortality, 2018). Indeed, fog computing now is giving us those peripheral nerves which allow us to complete that mapping from hydrocarbon-based systems through non-hydrocarbon-based ones and one to full digitization of hu-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

manity. In any case, it is a bellwether of thinking about how to manage seemingly irresolvable complex problems. Ray Kurzweil (Kurzweil, 2015), referred to above as a leader of the singularity movement, sees such a day coming. As defined,

The technological singularity is a hypothetical event related to the advent of artificial general intelligence (also known as "strong AI"). Such a computer, computer network or robot would theoretically be capable of recursive self-improvement (redesigning itself) or of designing and building computers or robots better than itself. Repetitions of this cycle would likely result in a runaway effect an intelligence explosion (Chalmers, 2010; Loosemore, 2011) where smart machines design successive generations of increasingly powerful machines, creating intelligence far exceeding human intellectual capacity and control. Because the capabilities of such a superintelligence may be impossible for a human to comprehend, the technological singularity is the point beyond which events may become unpredictable or even unfathomable to human intelligence. (Wikipedia, 2018s)

By regarding society as a natural organism, one looks at the evolution of computer networks and artificial intelligence with all that big data, processing ability, and fog computing to incorporate everything having a chip and those holding that object that may equal or even exceed present "natural" capabilities.

The concept is not new if one remembers Alan Turing and John Von Neumann. In a more sophisticated form John Glad (Glad, 2008) argues that we can shape our own evolution in a cyborg fashion. We have come full circle to our creating ourselves just as Von Neumann with his Universal Constructor (Universal Constructor, 2015) and Turing (Turing, 1950) suggested. A "minor" problem exists - bias. Turing stated,

...a machine undoubtedly can be its own subject matter. It may be used to help in making up its own programmes, or to predict the effect of alterations in its own structure. By observing the results of its own behaviour it can modify its own programmes so as to achieve some purpose more effectively. These are possibilities of the near future, rather than Utopian dreams. (Turing, 1950)

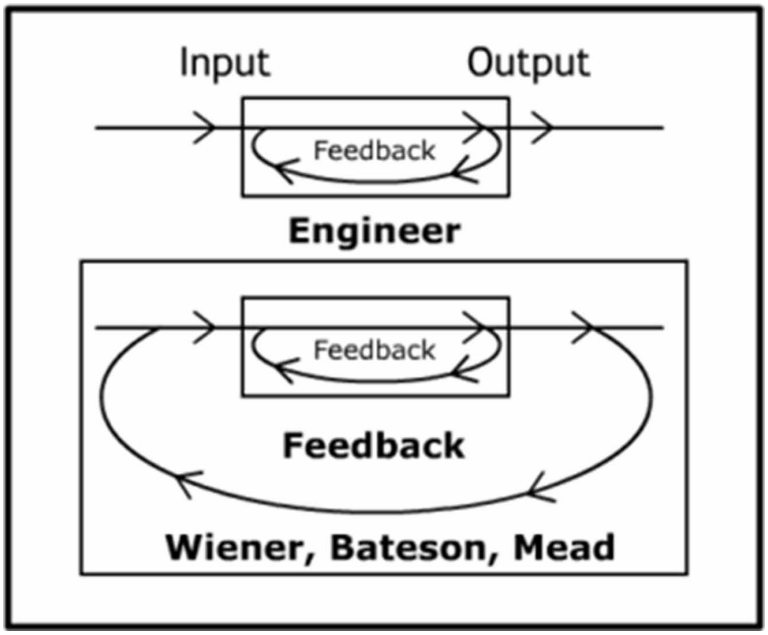
For argument's sake, let us say that the outcome of the ultimate big data project-the creation of a mind equivalent to ourselves-is feasible (and, again, we are not arguing that it ever can be successful, only that persons are working as though it could be). Clearly, the observer has become a participant, i.e., second order cybernetics and recursion (Wikipedia, 2018q). That artificial mind would be seeing us through ourselves, a case of reflexivity, a topic that has generated its own and very large body of literature (Wikipedia, 2018p). Immediately, we are confronted with emergent systems immanent with human bias. A data visualization as a context diagram may appear as in Figure 9.

If one is to posit that the culmination of big data with fog computing will be in the form of reproducing ourselves as in the artificial brain, one should be free to discuss some consequences having to do with human bias and the differences between an organic and a non-hydrocarbon-based way of doing things. Such is even truer about replicating society as a being unto itself. Too, the question is not so much as whether it is possible to do all of this down the road but the thinking in that direction and the implications of these attitudes.

In emergence, one needs to address autopoiesis or self-creation (Wikipedia, 2018a). After all, modeling a social brain is what we are doing. In the same discussion about the seeds of creation being within ourselves, one must also address in a dialectic fashion, autodestruction analogous to a death gene (Prescott, 2012). What if, in emergence, as in systems acting on their own (Singer, 2009), this tendency

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 9. Second order cybernetics feedback loop (Côté, 1976)



becomes predominant? The U.S. Department of Defense surely has these concerns as illustrated in one of its Broad Agency Announcements:

This new and expanding challenge to the test community is being driven by rapidly accelerating technology developments in cognition, control, autonomy, sensors, computing, software and communications resulting in the development of advanced unmanned military-technical capabilities being deployed at an ever expanding pace. This is consequently straining the verification and validation methods as they are practiced today (BAA UAST-0002, p. 13).

Another consideration is that of entropy. It is said that the Universe is tending towards a heat death where all energy will be equally dispersed (Wikipedia, 2018j). Entropy also results when information emanates from chaos (as in the singularity that was the origin of the Universe) produced what we have today and will continue to produce until its heat death end. We have as Boltzmann entropy, the equalization of particles giving an evenness. Information entropy is regarded as the ability to predict (Shannon, 1948; Chaitin, 2001). It is uniformity that allows the experimental scientific method (Whewell, 1847), but is not uniformity randomness? Yet, this is a uniformity of process but it would appear that entropy in this sense already has occurred. In all the scientific development it becomes hypothetically possible to predict the future. Given what we have just said, we will be approaching a maximum entropy, in essence, a condition being, as the Universe is tending, one of maximum information dispersal. It will be a maximum Cartesianism as it were to test our ability to identify every element. Information entropy then may be that autodestruction. Our quest for "knowledge" is that autodestructive process with big data the barometer.

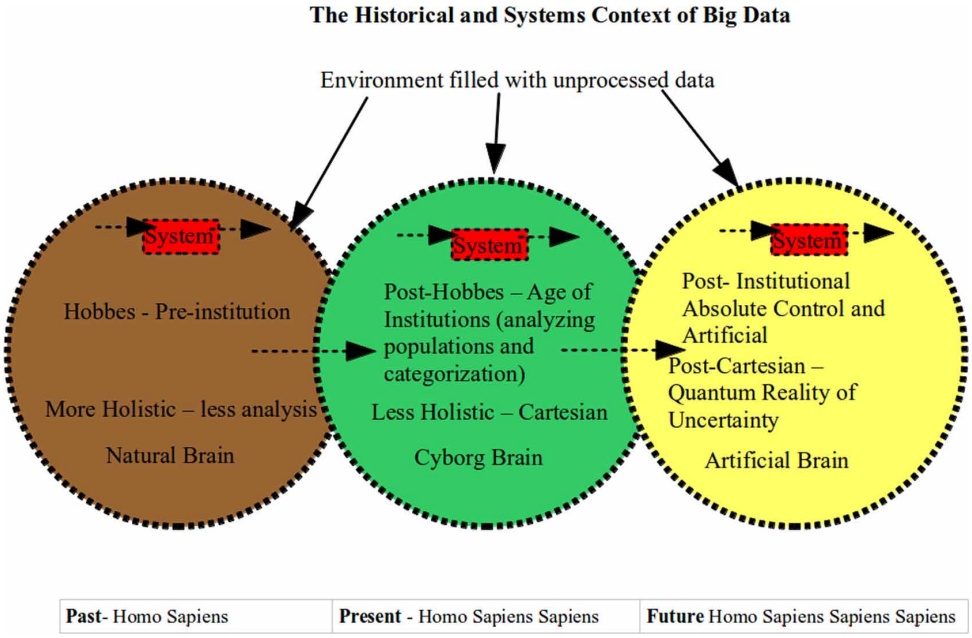
EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

CONCLUSION

Data exists every time we measure something or tell someone about something. There may be too much of it for our own good. Ostensibly, it seems we can manage it by inventing storage and distribution technologies. However, "big" elides to "unmanageable" when there are insufficient organization and analysis methods. It is at this juncture where the conversation about "big data" goes from describing the technological wonderment to the more serious one of not only the organization but the purpose of big data. This brings in more profound considerations of philosophy (metaphysics, ethos, etc.). If we cannot manage the answers to this question, the technological aspects of management become minuscule even comical. Think of society as a child building a house of cards of big data and poking around it with fog computing.

Large projects, exemplified by genome sequencing, weather modeling, or quantum research in particle accelerators are simple in comparison to modeling global combat, the complexities of the human brain, or society as a conscious organism. Such complexity raises the question of humans' ability to manage it. It may be the case that in a situation similar to humans needing jack for the car they can't alone lift. Humans need a tool to help think of solutions for problems they cannot solve with their natural brains. It is in this context we can have our discussion of big data, especially in light of our discussion on the U.S. intelligence network and agencies, the National Institute of Mental Health's Research Across Domains (RDOC, 2015), the various artificial brain projects, and the Broad Agency Announcement calls for identification and detention of mentally disordered persons as potential terrorists (Broad Agency Announcement 11-65, 2015). One asks about the concept of Durkheim's "social brain", extending digital immortality (Tipler, 1994) to society as a whole, especially if this species is asking whether it will survive.

Figure 10. Our context diagram (Courtesy of the author, Jeremy Horne – drawn especially for this paper)



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The more data accumulated the more we have measured our world and presumably the more we have understood it. Knowledge acquisition overall has as its main objective the enhancement of our ability to predict (and control) our future. The question immediately relevant in this paper is who acquires that data and who manages it? Without a science of society and politics, the anarchy of big data will continue, the trend leading where most anarchistic endeavors have in the past, to destruction and iron-fisted oppressive rule.

These broader considerations of "big data" and fog computing make it all the more imperative to conduct research not simply into the technological aspects of the subject but more importantly the reasons for it. So, in the end, the reader is left with a visualization of data within the larger framework of what has been described here linearly Figure 10:

From this, one only hopes that there will be a simulation of a second order cybernetic model that will at least preserve this species and its environment. Let's look into that mirror again and reflect on all of what has been said here.

AUTHOR'S NOTE

A note on Wikipedia: This article is only a seed for further thinking about big data. Numerous citations of Wikipedia appear, not because they necessarily are peer-reviewed sources (though many articles are) but because they provide a general overview of the topic and moreover lists of further readings upon which the content is based. In addition, this information is what the general public often encounters (and only this one). One should refer to the content and references as any competent scholar would and see, for the most part, that most importantly, the Wikipedia presents a useful starting point for research.

In terms of other references, they are not meant to be complete but only what the author of this paper thinks are representations of typical evidence supporting the arguments. A number of works for which there are no downloadable versions can be purchased on line through various booksellers. Many can be read on line via places like Gutenberg Project or https://archive.org/.

REFERENCES

Abdollahian, M., Baranick, M., Efird, B., & Kugler, J. (2006). *Senturion - A Predictive Political Simulation Model*. Retrieved 28 December 2015 from www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA454175

Alsop, H. (2014). Americans surveyed: Misunderstood, misrepresented or ignorant? *The Telegraph*. Retrieved from http://www.telegraph.co.uk/news/worldnews/northamerica/usa/10640690/Americans-surveyed-misunderstood-misrepresented-or-ignorant.html

Army Research Office. (2017). Broad Agency Announcement for basic and applied scientific research (BAA - W911NF-17-S-0002 – 20017). U.S. Army Contracting Command, Aberdeen Proving Ground, NC. Retrieved from https://www.arl.army.mil/www/pages/8/W911NF-17-S-0002.pdf

Ashby, R. (1954). *Design for a brain*. New York: Wiley; https://ia801406.us.archive.org/3/items/designforbrain00ashb/designforbrain00ashb.pdf doi:10.5962/bhl.title.6969

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Azevedo, F. A., Carvalho, L. R., Grinberg, L. T., Farfel, J. M., Ferretti, R. E., Leite, R. E., ... Herculano-Houzel, S. (2009). Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *The Journal of Comparative Neurology*, *513*(5), 532–541. doi:10.1002/cne.21974 PMID:19226510

Bamford, J. (2012, March 15). The NSA is building the country's biggest spy center (watch what you say). *Wired*. Retrieved from https://www.wired.com/2012/03/ff_nsadatacenter/all/

Baraniuk, C. (2015). Supercomputers: Obama orders world's fastest computer. *BBC News*. Retrieved from http://www.bbc.com/news/technology-33718311

Barcelona Smart City. (2015). http://smartcity.bcn.cat/es/

Beering, S. (2014). *Science and technology: Public attitudes and understanding*. Retrieved from http://www.nsf.gov/statistics/seind14/content/chapter-7/chapter-7.pdf

Binary Translator. (2018). Q-Bit. Retrieved from http://www.qbit.it/lab/bintext.php?text=&enc=Encode+text+to+binary+code

Blue Brain. (2018). The Blue Brain Project – A Swiss Brain Initiative. Retrieved from http://bluebrain.epfl.ch/

Brain Project. (2000). Project on the Decade of the Brain. Retrieved from http://www.loc.gov/loc/brain/

Brewin, B. (2018). Army awards $1.2 billion contract for NSA data center in Utah. *Nextgov*. Retrieved from http://www.nextgov.com/cybersecurity/2010/09/army-awards-12-billion-contract-for-nsa-data-center-in-utah/47647/

Buckley, W. (Ed.). (1968). *Modern systems research for the behavioral scientist*. Chicago: Aldine Publishing Company. Retrieved from http://philpapers.org/rec/BUCMSRhttps://en.wikipedia.org/wiki/Walter_F._Buckley

Callaos, N. (1995) Metodologia Sistemica De Sistemas: Concepto y aplicaciones-libre. Caracas, VE: Universidad Simon Bolivar.

Carey, B. (2014). Stanford biologist warns of early stages of Earth's 6th mass extinction event. *Stanford News Service*. Retrieved from http://news.stanford.edu/pr/2014/pr-sixth-mass-extinction-072414.html

Cattermole, T. (2010). Why we aren't as ethical as we think we are. *Gizmag*. Retrieved from http://www.gizmag.com/we-arent-as-ethical-as-we-think-we-are/14572/

Ceballos, G., Ehrlich, P. R., Barnosky, A. D., Garcia, A., Pringle, R. M., & Palmer, T. M. (2015). Accelerated modern human induced species losses: Entering the sixth mass extinction. *Science Advances, 1*(5). Retrieved from http://advances.sciencemag.org/content/1/5/e1400253

Center for conscious studies. (n.d.). The University of Arizona. Retrieved from http://consciousness.arizona.edu

CERT. (2018). *Computer Emergency Response Team – Software Engineering Institute*, Carnegie-Mellon University. Retrieved from http://www.cert.org/

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chaitin, G. (2001). *Exploring Randomness. Theoretical interlude---What is randomness? My definitions*. London: Springer-Verlag. Retrieved from https://d-nb.info/960010319/04

Chalmers, D. J. (2010). The singularity: A philosophical analysis. *Journal of Consciousness Studies*, *17*, 7–65. Retrieved from http://consc.net/papers/singularity.pdf

Chalmers, D. J. (2013). *The Matrix as metaphysics*. Oxford, United Kingdom: Oxford University Press. Retrieved from http://consc.net/papers/matrix.html

Cisco. (n.d.). Big Data. Retrieved from http://www.cisco.com/c/en/us/solutions/data-center-virtualization/big-data/index.html

Citizens United vs. Federal Election Commission (2010). *Appeal from the United States District Court for the District of Columbia No. 08–205.* Argued March 24, 2009—Reargued September 9, 2009—Decided January 21, 2010. Retrieved from http://www.supremecourt.gov/opinions/09pdf/08-205.pdf

Computer Security (2015). *SANS.* Retrieved from www.sans.org

Congress.Gov. (2015). *Coburn Amendment S. Amdt. 2631 to H.R. 2847*. Retrieved from https://www.congress.gov/amendment/111th-congress/senate-amendment/2631

Cope, L. M., Ermer, E., Gaudet, L. M., Steele, V. R., Eckhardt, A. L., Arbabshirani, M. R., ... Kiehl, K. A. (2014). Abnormal brain structure in youth who commit homicide. *NeuroImage. Clinical*, *10*(4), 800–807. Retrieved from http://www.sciencedirect.com/science/article/pii/S2213158214000588. doi:10.1016/j.nicl.2014.05.002 PMID:24936430

Côté, M. (1976). Redrawn from Co Evolution Quarterly and Cybernetics Society Conference brochure (image). *Wikimedia*. Retrieved from https://commons.wikimedia.org/wiki/File:SOCyberntics.png

Daily Mail Reporter. (2013). Is the NSA building a $800million data center to spy on Americans? Utah desert facility code-named Bumblehive 'will monitor emails of U.S. citizens.' *Daily Mail*. Retrieved from http://www.dailymail.co.uk/news/article-2309906/Is-NSA-building-1-2billion-data-center-spy-Americans-Utah-desert-facility-monitor-emails-U-S-citizens.html

Data Dictionary Webster. (n.d.). Data. Retrieved from http://www.merriam-webster.com/dictionary/data

Datum Dictionary dot com (2015). Datum. Retrieved from http://dictionary.reference.com/browse/datum

Datum Dictionary Oxford. (2015). Datum. Retrieved from http://www.oxforddictionaries.com/us/definition/american_english/datum

Datum Dictionary Webster. (2015). Datum. Retrieved from www.merriam-webster.com/dictionary/datum

Dawes, C. T., & Fowler, J. H. (2009). Partisanship, voting, and the dopamine d2 receptor gene. *The Journal of Politics*, *71*(3), 1157–1171. Retrieved from http://jhfowler.ucsd.edu/partisanship_voting_and_drd2.pdf doi:10.1017/S002238160909094X

DeAngelis, T. (2009). Understanding terrorism. Monitor on Psychology, 40(10), 60. *American Psychological Association*. Retrieved from http://www.apa.org/monitor/2009/11/terrorism.aspx

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Deathcamps.org. (2006). Einsatzgruppen. Retrieved from http://www.deathcamps.org/occupation/einsatzgruppen.html

Descartes, R. (1641). Meditations on first philosophy. *The Internet Encyclopedia of Philosophy*. Retrieved from http://infomotions.com/etexts/philosophy/1600-1699/descartes-meditations-746.htm

Diaz, J. (2010). The one hundred trillion dollars hard drive. *Gizmodo Media Group*. Retrieved on 28 December 2015 from http://gizmodo.com/5557676/how-much-money-would-a-yottabyte-hard-drive-cost

Dictionary.com. (2015). Data. Retrieved from http://dictionary.reference.com/browse/data

Digital Immortality. (2018). *Digital Immortality*. Retrieved from https://en.wikipedia.org/wiki/Digital_immortality

DNI. (2018). Office of the Director of National Intelligence. Retrieved from https://www.dni.gov/index.php

DOE Statistics. (2013). *Department of Education Statistics- Nation's Report Card*. Retrieved from http://www.nationsreportcard.gov/reading_math_g12_2013/#/preparedness

DSM. (2015). Diagnostic and Statistical Manual of Mental Disorders (DSM-5). Retrieved from http://www.dsm5.org/Pages/Default.aspx

Durkheim, E. (1982). *Rules of sociological method*. New York: Macmillan. Retrieved from https://monoskop.org/images/1/1e/Durkheim_Emile_The_Rules_of_Sociological_Method_1982.pdf doi:10.1007/978-1-349-16939-9

Eliade, M. (1954). *Cosmos and history (The Myth of the Eternal Return).* New York: Harper Torchbooks. Retrieved from http://users.uoa.gr/~cdokou/MythLitMA/Eliade-EternalReturn.pdf

Etymonline. (n.d.). Data Etymology. Retrieved from http://www.etymonline.com/index.php?term=data

Federal Business Opportunities. (2008). UAST - Unmanned Autonomous Systems Technology BAA UAST-0002. Retrieved from https://www.fbo.gov/index?s=opportunity&mode=form&id=dbb918ff38f72d71f8f6165890a7e86a&tab=core&_cview=1

Federal Business Opportunities. (2010). Y—Utah Data Center. Retrieved from https://www.fbo.gov/index?s=opportunity&mode=form&id=633411ff6e764cdffd3df8732b8a9cd6&tab=core&_cview=1

Federal Business Opportunities. (2012). Defense Sciences Research and Technology. Retrieved from https://www.fbo.gov/index?s=opportunity&mode=form&tab=core&id=31cf5a859cfa2764da7e1b8d54a79d1c&_cview=0

Ferguson, S., Lange, T., & Capenos, C. (2009, November 13). Utah Data Center (UDC) [presentation]. U.S. Army Corps of Engineers. Retrieved from https://info.publicintelligence.net/UtahDataCenter.pdf

Firth, N. (2010). Human race 'will be extinct within 100 years', claims leading scientist. *Daily Mail.* Retrieved from http://www.dailymail.co.uk/sciencetech/article-1287643/Human-race-extinct-100-years-population-explosion.html

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Gardner, H. (1993). *Frames of mind: the theory of multiple intelligences*. New York: Basic Books. Retrieved from http://www.tecweb.org/styles/gardner.htmlhttps://en.wikipedia.org/wiki/Theory_of_multiple_intelligenceshttps://howardgardner.com/papers/

Gardner, H. (2004). A multiplicity of intelligences: In tribute to Professor Luigi Vignolo. Retrieved from https://llk.media.mit.edu/courses/readings/gardner-multiple-intelligences.pdf

Gazzaniga, M. S., Ivry, R. B., & Mangun, G. R. (Eds.). (2009). *Cognitive Neuroscience: The Biology of the Mind* (4th ed.). New York: Norton Publishing. Retrieved from https://www.amazon.com/Cognitive-Neuroscience-Biology-Mind-4th/dp/0393913481

Glad, J. (2008). *The future of human evolution*. Schuylkill Haven, PA: Hermitage Publishers. Retrieved from http://www.whatwemaybe.org/txt/txt0000/Glad.John.2008.FHE.Meisenberg-abridgement.r004.en.pdf

Google Images. (n.d.). Contraband images. Retrieved from https://www.google.com.mx/search?q=inmate+contraband&client=ubuntu&hs=Si7&channel=fs&tbm=isch&tbo=u&source=univ&sa=X&ved=0ahUKEwiw0d-9oajJAhXXm4gKHa6FAq0QsAQIKA&biw=1272&bih=856

Gordon, G. (2013, July 3). Hints surface that NSA building massive, pervasive surveillance capability. *Stars and Stripes*. Retrieved from https://www.stripes.com/news/us/hints-surface-that-nsa-building-massive-pervasive-surveillance-capability-1.228664

Gun Violence Archive. (2018). Multiple shootings. (Retrieved from http://www.gunviolencearchive.org/

Hawking, S. (2015). *Into a Black Hole, Stephen Hawing's Page*. Retrieved from http://www.hawking.org.uk/

Hegel, G. W. F. (2001). *Philosophy of Right* (S.W. Dyde, Trans.) Ontario: Batoche Books, Ltd. (originally printed in 1833). Retrieved from https://socserv2.socsci.mcmaster.ca/econ/ugcm/3ll3/hegel/right.pdf

Heile, F. (2015). Is the total number of particles in the universe stable over long periods of time? *Huffington Post*. Retrieved from http://www.huffingtonpost.com/quora/is-the-total-number-of-pa_b_4987369.html

Herculano-Houzel, S. (2012). The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(Suppl. 1), 10661–10668. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22723358 doi:10.1073/pnas.1201895109 PMID:22723358

Higginbotham, S. (2012). As data gets bigger, what comes after a yottabyte? *Examiner.com*. Retrieved from https://gigaom.com/2012/10/30/as-data-gets-bigger-what-comes-after-a-yottabyte/

Hobbes, T. (1651). *Leviathan*. London: Andrew Crooke. Retrieved from http://www.gutenberg.org/files/3207/3207.txt

Homer-Dixon, T. F. (2007). *The Upside of Down: Catastrophe, Creativity, and the Renewal of Civilization*. New York: Knopf.

Horne, J. (1997). Logic as the language of innate order in consciousness. *Informatica, 21*(4), 675-682. Retrieved from http://www.vxm.com/link.LogicLanguage.html

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Horne, J. (2006). Recursion of binary space as a foundation of repeatable programs. *Journal of Systemics, Cybernetics and Informatics*, *4*(5), 73–77. Retrieved from http://www.iiisci.org/journal/CV$/sci/pdfs/P277173.pdf

Horne, J. (2012). A new three dimensional bivalent hypercube. *NeuroQuantology: An Interdisciplinary Journal of Neuroscience and Quantum Physics*, *10*(1), 20–30. Retrieved from http://www.neuroquantology.com/index.php/journal/article/view/361 doi:10.14704/nq.2012.10.1.361

Human Brain Project. (2013). *Human Brain Project*. Retrieved from https://www.humanbrainproject.eu/

Hume, D. (2015). *Essays moral, political and literary 1742, 1752*. The Founders' Constitution. Chicago: The University of Chicago Press. Retrieved from http://press-pubs.uchicago.edu/founders/documents/v1ch11s3.html

IARPA. (2018). Intelligence Advanced Research Projects Activity. Retrieved from https://www.iarpa.gov/

IARPA Research Programs. (2018). Retrieved from https://www.iarpa.gov/index.php/research-programs

IBM. (n.d.). Big data analytics. Retrieved from http://www.ibm.com/big-data/us/en/

IBM SYNAPSE. (2015). *IBM -Systems of Neuromorphic Adaptive Plastic Scalable Electronics (SyNAPSE)*. Retrieved from https://www.darpa.mil/program/systems-of-neuromorphic-adaptive-plastic-scalable-electronics

ICD-10 Version. 2016 (2016). World Health Organization. Retrieved from http://www.who.int/classifications/icd/en/

Ionescu, D. (2013). Thumb Drive Capacity. Kingston's 1 TB thumb drive packs massive storage capacity in tiny package. *PCWorld*. Retrieved from http://www.pcworld.com/article/2024220/kingstons-1tb-thumb-drive-packs-massive-storage-capacity-in-tiny-package.html

Johnson, K. A., & Becker, J. A. (2015). The whole brain Atlas. *Harvard Medical School*. Retrieved 28 December 2015 from http://www.med.harvard.edu/aanlib/home.html

Jordan, R. (2015, June 19). Stanford researcher declares that the sixth mass extinction is here. *Stanford News*. Retrieved from http://news.stanford.edu/news/2015/june/mass-extinction-ehrlich-061915.html

Kalbandi, I., & Anuradha, J. (2015). A brief introduction on big data 5Vs characteristics and Hadoop technology. *Procedia Computer Science*, *48*, 319–324. Retrieved from http://www.sciencedirect.com/science/article/pii/S1877050915006973 doi:10.1016/j.procs.2015.04.188

Kanai, R., Feilden, T., Firth, C., & Rees, G. (2011). Political Orientations Are Correlated with Brain Structure in Young Adults. *Current Biology*, *21*(8), 677–680. Retrieved from http://www.cell.com/current-biology/fulltext/S0960-9822(11)00289-2 doi:10.1016/j.cub.2011.03.017 PMID:21474316

Khodjasteh, K., Sastravan, J., Hayes, D., Green, T. J., Biercuk, M. J., & Viola, L. (2013). Designing a practical high-fidelity long-time quantum memory. *Nature Communications*, *4*(2045). Retrieved from http://www.nature.com/ncomms/2013/130619/ncomms3045/full/ncomms3045.html PMID:23784079

Kintisch, K. (2015). Climate crossroads. *Science*, *350*(6264), 1016-1017. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/26612930

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Le Bon, G. (1896). *The crowd, A Study of the Popular Mind*. Kitchener, Ontario: Batoche Books. Retrieved from http://socserv.socsci.mcmaster.ca/~econ/ugcm/3ll3/lebon/Crowds.pdf

Lifton, R. J. (1986). *The Nazi doctors*. New York: Basic Books. Retrieved from https://www.scribd.com/doc/200776030/Robert-Jay-Lifton-The-Nazi-Doctors

Literacy Project Foundation. (2018). *Literacy by the Numbers*. http://literacyprojectfoundation.org/community/statistics/

Loosemore, R., & Goertzel, B. (2011). Why an intelligence explosion is probable. *Humanity Media*. Retrieved from http://hplusmagazine.com/2011/03/07/why-an-intelligence-explosion-is-probable /

Mearian, L. (2014). WD leapfrogs Seagate with world's highest capacity 10TB helium drive, new flash drives. *Computer World*. Retrieved from http://www.computerworld.com/article/2604311/computer-hardware/wd-leapfrogs-seagate-with-world-s-highest-capacity-10tb-helium-drive-new-flash-drives.html

Microsoft. (2015). *Are you protected from online threats?* [Infographic]. Retrieved from http://mscorp.blob.core.windows.net/mscorpmedia/2013/04/7536.Microsoft_5F00_SIR_5F00_infographic_5F00_long.png

Miller, G. (2012). Investigating the psychopathic mind. Science, *321*(5894), 1284-1286. Retrieved from http://www.sciencemag.org/content/321/5894/1284

Minin, I. (2011). *Microsensors*. London: InTech. Retrieved from https://www.intechopen.com/books/microsensors

Mondal, S. (2015, February 22). *Organic theory of the state*. Retrieved from http://socialscienc.blogspot.com/2015/02/organic-theory-of-state.html

Morel, O. T., Wong, M. E. Y., Goerger, S. R., Kwinn, M. J., & Dodge, R. C. (2006). *Effects Based Assessment Support System* (EBASS). Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a448132.pdf

Morel, T. O. (2006). Effects Based Assessment Support System (EBASS) [technical report]. United States Military Academy West Point. Retrieved from [REMOVED HYPERLINK FIELD]http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA448132

MSBOK. (2015). *Modeling and simulation body of knowledge*. Retrieved from http://www.msco.mil/documents/_25_MSBOK20101022DistA.pdf

MSCO. (2015). *Modeling and Simulation Coordination Office*. Retrieved from http://www.msco.mil/

National Center for Education Statistics. (n.d.). Reading Performance. Retrieved from http://nces.ed.gov/programs/coe/indicator_cnb.asp

National Commission. (1983). *A nation at risk: The imperative for educational reform.* National Commission on Excellence in Education. Retrieved from https://www.edreform.com/wp-content/uploads/2013/02/A_Nation_At_Risk_1983.pdf

National Geographic Society. (2006). *National Geographic-Roper Public Affairs 2006 geographic literacy study*. Retrieved from https://media.nationalgeographic.org/assets/file/Roper-Poll-2006-Highlights.pdf

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

National Infrastructure Advisory Council. (2012). *Intelligence Information Sharing: Final Report and Recommendations*. Retrieved from https://www.dhs.gov/xlibrary/assets/niac/niac-intelligence-information-sharing-final-report-01102012.pdf

National Institute of Mental Health (NIMH). (2018). *Research Domain Criteria*. Bethesda, MD: National Institutes of Health (NIH). Retrieved from https://www.nimh.nih.gov/research-priorities/rdoc/index.shtml

National Vulnerability Database. (2018). *Computer Security Resource Center*, National Institute of Standards and Technology. Retrieved from https://web.nvd.nist.gov/view/vuln/search

Neuman, S. (2014). 1 in 4 Americans thinks the sun goes around the earth, survey says. The Two-way. *NPR*. Retrieved from http://www.npr.org/sections/thetwo-way/2014/02/14/277058739/1-in-4-americans-think-the-sun-goes-around-the-earth-survey-says

NIMH-sociopathy. (2018). *Antisocial Personality Disorder*. Retrieved from http://www.nimh.nih.gov/health/statistics/prevalence/antisocial-personality-disorder.shtml

NIST. (2017). The NIST Definition of Fog Computing. NIST Special Publication 800-191 (Draft). Gaithersburg, MD: National Institute of Standards and Technology. Retrieved from https://csrc.nist.gov/CSRC/media/Publications/sp/800-191/draft/documents/sp800-191-draft.pdf

NSA. (2018). IC off the Record. Retrieved from https://nsa.gov1.info/dni/2018/index.html

Ollman, B. (2017). *What is Political Science? What should it be*? Retrieved from https://www.nyu.edu/projects/ollman/docs/what_poly_sci.php

Oracle. (n.d.). Streamline your Big Data platform. Retrieved from https://www.oracle.com/big-data/index.html

Oxford Living Dictionaries. (2015). Data. Retrieved from http://www.oxforddictionaries.com/us/definition/american_english/data

Pace, P. R. (2014). Compassion fatigue a growing concern. *NASW News*, *59*(1). Retrieved from http://www.socialworkblog.org/nasw-news-article/2014/01/self-care-in-social-work-compassion-fatigue-a-growing-concern/

Popper, K. R. (1959). *The Logic of Scientific Discovery* (J. Freed & J. Freed, trans.). New York: Basic Books. Retrieved from http://strangebeautiful.com/other-texts/popper-logic-scientific-discovery.pdf

Prescott, B. (2012). Gene helps predict time of death. *Harvard Medical School*. Retrieved from http://hms.harvard.edu/news/gene-helps-predict-time-death-11-26-12

Quasar Jarosz. (2009). Neuron Hand-tuned. https://commons.wikimedia.org/wiki/File:Neuron_Hand-tuned.svg

Reuters Staff. (2013). U.S. agency denies data center to monitor citizens' emails. *Reuters*. Retrieved from https://www.reuters.com/article/us-usa-security-nsa/u-s-agency-denies-data-center-to-monitor-citizens-emails-idUSBRE93E11O20130416

Riken Brain Science Institute (RBSI). (n.d.). Retrieved from http://www.brain.riken.jp/en/

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Rousseau, J. J., & Cole, G. D. H. (Eds.). (1761). *The Social Contract and Discourses*. Toronto, Canada: J.M. Dent and Sons. Retrieved from http://oll.libertyfund.org/titles/638

SANS Institute. (n.d.). Retrieved from www.sans.org

SANS Institute. (n.d.). Newsletters: NewsBites. Retrieved from https://www.sans.org/newsletters/newsbites

SAS. (n.d.). What is Big Data? Retrieved from http://www.sas.com/en_us/insights/big-data/what-is-big-data.html

SC Magazine. (n.d.). Retrieved from https://www.scmagazine.com/

Segall, R. S., & Cook, J. S. (2018). *Handbook of Research on Big Data Storage and Visualization Techniques*. Hershey, PA: IGI-Global. Retrieved from https://www.igi-global.com/book/handbook-research-big-data-storage/179829 doi:10.4018/978-1-5225-3142-5

Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal, 27(*623-656), 379-423. Retrieved from http://math.harvard.edu/~ctm/home/text/others/shannon/entropy/entropy.pdf

Sheldrick, G. (2014). '100 years left' scientists warn of TRIGGER to mass biological event WIPING OUT humans. *Express News*. Retrieved from http://www.express.co.uk/news/nature/493422/End-of-the-World-scientists-warn-biological-extinction-event-in-100-years

Singer, P. W. (2009). *Wired for war: The robotics revolution and conflict in the 21st century*. New York: Penguin. Retrieved from http://wiredforwar.pwsinger.com/

Spencer, H. (Ed.). (1897). *The principles of sociology*. New York: D. Appleton and Company. Retrieved from https://ia902305.us.archive.org/22/items/prinsocial01spenuoft/prinsocial01spenuoft.pdf

Tae-il, K., McCall, J. G., Jung, Y. H., Huang, X., Siuda, E. R., Li, Y., . . . Bruchas, M. R. (2013). Injectable, cellular-scale optoelectronics with applications for wireless optogenetics. *Science, AAAS*, 340(6129), 211-216. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/23580530

Steps toward Memory Preservation v1.6. (2018). Retrieved from https://docs.google.com/document/d/184MAffoiRo70JrYqG9GIcleBZRPoH1iqSAaJP3fJ97I/edit#heading=h.34vw9iiczx8a

Tainter, J. A. (1990). The collapse of complex societies [Abstract]. Retrieved from https://books.google.com.mx/books?id=M4H-02d9oE0C&printsec=frontcover&dq=%22collapse+of+complex+societies%22&hl=en&ei=1fCTTu7XG6TjiALbtMjjBA&sa=X&oi=book_result&ct=result&redir_esc=y#v=onepage&q&f=false

Technocracyinc. (n.d.). Retrieved from http://www.technocracyinc.org/

TED. (2008). *Sir Martin Rees: Earth in its final century?* [YouTube video]. Retrieved from https://www.youtube.com/watch?v=3qF26MbYgOA

The Emirr. (2011, November 27). Nervous System Diagram [image file]. *Wikimedia*. Retrieved from https://commons.wikimedia.org/wiki/File:TE-Nervous_system_diagram.svg

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Theaetetus, P. (1921). *Plato in twelve volumes* (Vol. 12). (H. N. Fowler, Trans.). Cambridge, MA: Harvard University Press. Retrieved from http://www.perseus.tufts.edu/hopper/text?doc=urn:cts:greekLit:tlg0059.tlg006.perseus-eng1:152a

Tice, P. (2015). Serious psychological distress (SPD). Table 1.17a. In *NIMH Results from the 2014 national survey on drug use and health: Mental health detailed tables* (p. 195). Retrieved from http://www.samhsa.gov/data/sites/default/files/NSDUH-MHDetTabs2014/NSDUH-MHDetTabs2014.pdf

Trifonov, E. N. (2012, February). Opinion: What is life? The *Scientist*, *16*. Retrieved from https://www.the-scientist.com/?articles.view/articleNo/31730/title/Opinion--What-Is-Life-/

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*(236), 433–460. doi:10.1093/mind/LIX.236.433

Uptime Institute. (2018). Data Center Site Infrastructure Tier Standard: Topology. Retrieved from https://uptimeinstitute.com/uptime_assets/4f5c1d5a03c7c51036ee73b524e0769fa6790d12bf01f52e2198a4e0cc7b6888-00001D.pdf

U.S.-Cert. U.S. Computer Emergency Readiness Team. (n.d.). Retrieved from https://www.us-cert.gov/

U.S. Census Bureau. (2018). US and World Population Clock. Retrieved from http://www.census.gov/popclock/

US Congress. (2015). *S.754 2015 Division N-Cybersecurity Act of 2015*. Retrieved from https://www.congress.gov/114/bills/s754/BILLS-114s754es.pdf

U.S. Department of Homeland Security - If. (2018). If you see something, say something. Retrieved from https://www.dhs.gov/see-something-say-something

U.S. Department of Homeland Security - Information. (2018). Information Sharing. Retrieved from https://www.dhs.gov/topic/information-sharing

U.S. Intelligence Community. (2018). We are the Intelligence Community. Retrieved from https://www.intelligence.gov/how-the-ic-works#our-organizations

Utah Data Center. (2018). *Domestic Surveillance Directorate*. Retrieved from https://nsa.gov1.info/utah-data-center/

Vaughan-Nichols, S.J. (2015). Six Clicks: The six fastest computers in the world. *ZDnet*. Retrieved from http://www.zdnet.com/pictures/six-clicks-the-six-fastest-computers-in-the-world/

Vintiadis, E. (2018). Emergence. *Internet Encyclopedia of Philosophy*. Retrieved from https://www.iep.utm.edu/emergenc/

Weaver, M. (2017). What Do You Know About The National Security Agency Data Center At Camp Williams? *Facilities Magazine*. Retrieved from http://facilitiesmagazine.com/utah/buildings/nsa-utah-data-center

Webb, J. (2014). Met Office to build £97m supercomputer. Jonathan Webb Science reporter. *BBC News*. Retrieved from http://www.bbc.com/news/science-environment-29789208

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Weber, B. (2018). Life. In The Stanford Encyclopedia of Philosophy (Summer 2018 ed.).

Weber, M. (1992). *The Protestant ethic and the spirit of capitalism*. London, UK: Routledge. Retrieved from https://archive.org/details/protestantethics00webe

Webopedia. (n.d.). Big Data. Retrieved from http://www.webopedia.com/TERM/B/big_data.html

Weik, M. H. (1961). The ENIAC story. Aberdeen Proving Ground, MD: Ordnance Ballistic Research Laboratories. *The Journal of the American Ordinance Association*. Retrieved from http://ftp.arl.mil/mike/comphist/eniac-story.html

Whewell, W. (1847). *The philosophy of the inductive sciences*. London, UK: John W. Parker. https://archive.org/details/philosophyofindu01whewrich

Whitman, R. (2017, January 5). Kingston announces world's largest capacity 2TB flash drive. *Extremetech*. Retrieved from https://www.extremetech.com/electronics/242123-kingston-announces-worlds-largest-capacity-2tb-flash-drive

Wiener, N. (1948). *Cybernetics, or control and communication in the animal and the machine*. Cambridge, UK: MIT Press. Retrieved from http://uberty.org/wp-content/uploads/2015/07/Norbert_Wiener_Cybernetics.pdf

Wikipedia. (n.d.). Big Data. Retrieved from https://en.wikipedia.org/wiki/Big_data

Wikipedia. (n.d.). *Data, Information, Knowledge, Wisdom Pyramid*. Retrieved from https://en.wikipedia.org/wiki/DIKW_Pyramid

Wikipedia. (n.d.). Emergence. Retrieved from https://en.wikipedia.org/wiki/Emergence

Wikipedia. (n.d.). Frank J. Tipler. Retrieved from https://en.wikipedia.org/wiki/Frank_J._Tipler

Wikipedia. (n.d.). Henry Baxter Adams. Retrieved from https://en.wikipedia.org/wiki/Herbert_Baxter_Adams

Wikipedia. (n.d.). Nick Bostrom. Retrieved from https://en.wikipedia.org/wiki/Nick_Bostrom

Wikipedia. (n.d.). *Ray Kurzweil: Future predictions*. Retrieved from http://en.wikipedia.org/wiki/Ray_Kurzweil#Future_predictions

Wikipedia. (n.d.). Research domain criteria. Retrieved from https://en.wikipedia.org/wiki/Research_Domain_Criteria

Wikipedia. (n.d.). Self-replicating machine. Retrieved from https://en.wikipedia.org/wiki/Self-replicating_machine

Wikipedia. (n.d.). Technocracy. Retrieved from https://en.wikipedia.org/wiki/Technocracy

Wikipedia. (n.d. The bell curve. Retrieved from https://en.wikipedia.org/wiki/The_Bell_Curve

Wikipedia. (n.d.). Utah Data Center. Retrieved from https://en.wikipedia.org/wiki/Utah_Data_Center

Wikipedia. (n.d.). Von Neumann universal constructor. Retrieved from https://en.wikipedia.org/wiki/Von_Neumann_universal_constructor

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Wikipedia. (2018). Prehistoric Numerals. Retrieved from https://en.wikipedia.org/wiki/Prehistoric_numerals

Wikipedia. (2018). Yottabyte. Retrieved from https://en.wikipedia.org/wiki/Yottabyte

Wikipedia. (2018a). Autopoiesis. Retrieved from https://en.wikipedia.org/wiki/Autopoiesis

Wikipedia. (2018b). Bell Curve. Retrieved from https://en.wikipedia.org/wiki/The_Bell_Curve

Wikipedia. (2018c). Big Data Characteristics. Retrieved from https://en.wikipedia.org/wiki/Big_data#Characteristics

Wikipedia. (2018d). Chauvet Cave. Retrieved from https://en.wikipedia.org/wiki/Chauvet_Cave

Wikipedia. (2018e). Cloud Storage. Retrieved from https://en.wikipedia.org/wiki/Cloud_storage

Wikipedia. (2018f). Comparison of Backup Services. Retrieved from https://en.wikipedia.org/wiki/Comparison_of_online_backup_services

Wikipedia. (2018g). Context Diagram. Retrieved from https://en.wikipedia.org/wiki/System_context_diagram

Wikipedia. (2018h). Dark Energy. Retrieved from https://en.wikipedia.org/wiki/Dark_energy

Wikipedia. (2018i). Eddington Number. Retrieved from https://en.wikipedia.org/wiki/Eddington_number

Wikipedia. (2018j). Heat Death of the Universe. Retrieved from https://en.wikipedia.org/wiki/Heat_death_of_the_universe#Origins_of_the_idea

Wikipedia. (2018k). Machine Code. Retrieved from https://en.wikipedia.org/wiki/Machine_code

Wikipedia. (2018l). Malware. Retrieved from https://en.wikipedia.org/wiki/Malware

Wikipedia. (2018m). Net-Centric. Retrieved from https://en.wikipedia.org/wiki/Network-centric_warfare

Wikipedia. (2018n). Protagoras. Retrieved from https://en.wikipedia.org/wiki/Protagoras

Wikipedia. (2018o). Quantum Computing. Retrieved from https://en.wikipedia.org/wiki/Quantum_computing

Wikipedia. (2018p). Reflexivity (social theory). Retrieved from https://en.wikipedia.org/wiki/Reflexivity_%28social_theory%29

Wikipedia. (2018q). Second Order Cybernetics. Retrieved from https://en.wikipedia.org/wiki/Second-order_cybernetics

Wikipedia. (2018r). Systems Analysis. Retrieved from https://en.wikipedia.org/wiki/Systems_analysis

Wikipedia. (2018s). Technological Singularity. Retrieved from https://en.wikipedia.org/wiki/Technological_singularity

Wikipedia. (2018t). Theory of Multiple Intelligences. Retrieved from https://en.wikipedia.org/wiki/Theory_of_multiple_intelligences

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Wikivisually. (n.d.). Utah Data Center. Retrieved from https://wikivisually.com/wiki/Utah_Data_Center

Wikivisually. (n.d.). Utah Data Center. Retrieved from https://wikivisually.com/wiki/Utah_Data_Center#cite_note-reuters1-4

Wimmer, G. E., & Shohamy, D. (2012). Preference by association: How memory mechanisms in the hippocampus bias decisions. *Science*, *338*(6104), 270-273. Retrieved from http://www.sciencemag.org/content/338/6104/270.abstract

Wolfram, S. (2002). *A New Kind of Science*. Wolfram, Champaign, IL: Wolfram Media. Retrieved from http://www.wolframscience.com/nksonline/toc.html

Wuensche, A. (1993). Tools for researching Cellular Automata, Random Boolean Networks, multi-value Discrete Dynamical Networks, and beyond. Discrete Dynamics Lab (DDLab). Retrieved from http://www.ddlab.com/

KEY TERMS AND DEFINITIONS

Broad Agency Announcement (BAA): A BAA is a solicitation to the general public for one of its members or organizations to provide under contract, in the words of the U.S. Code of Federal Regulations, 48 CFR 35.016(a), "scientific study and experimentation directed toward advancing the state-of-the-art or increasing knowledge or understanding."

Cloud Storage: A repository of data on a remote storage medium often distributed through multiple servers accessible through the Internet and owned and managed by a provider.

Cellular Automaton: A grid of any number of dimensions composed of regularly-sized spaces or cells each of which may be in an active or inactive state, the whole assemblage of which assumes distinct conditions or goal states and through each unit of time each cell's condition being dependent upon the condition of one or more other cells.

Complexity Theory: Thinking underscoring the identification of simplicity or difficulty required in understanding the construction, analysis or resolution of a problem particularly related to systems.

Deduction: In an argument, if the premises are found to be true, the conclusion follows with certainty; the conclusion is found entirely within the set of premises; a closed system.

Dyadic Binary Formula: A composition of terms in a two-variable or base-2, system with a set of operators according to rules establishing how to compose an expression. Examples: $p \rightarrow q$, Rxy & Pyx, A + B (where "A" and "B" are the only two variables); 00111 * 10010

Emergence: Unexpected phenomena appearing (and often having a regularity or pattern) from a collection of apparently unrelated elements and where the elements themselves do not have the characteristics of the phenomena and that phenomena itself is not contained deductively within the elements (See "deduction").

Induction: In an argument, the conclusions follow only probabilistically from the premises. The scope of the conclusions is larger than the set of premises individually or collectively.

Quantum Computing: A computer architecture relying on the physics of quantum mechanics where the units or qubits of transmitting data may be in a multiple set of states simultaneously and where the transmission of data requires the "collapse" of one or more the qubits to a definite state.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Recursion: In its simplest sense: repeatability. For example, when the outputs or results of a formula are fed back into the formula as inputs, the outputs will reappear. In the computer world, an overall solution requires the solution of its smaller constituents. One statement can generate any number of objects, meaning a specific program can describe limitless computations. A computer program reflects recursion by having a function call itself inside a program.

Reductionism (Cartesianism): Subdivision of a whole into parts normally meaning that the whole can be described by the sum of those parts.

Second Order Cybernetics: The introduction of the observer or creator of a situation as a part of what is being observed.

Visualization Technique: Simply and broadly stated, a method of pictorializing or generating images of a symbolized: such as verbal or numerical - description.

This research was previously published in the International Journal of Fog Computing (IJFC), 1(2); edited by Sam Goundar and Kashif Munir; pages 1-50; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Section 7
Emerging Trends

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 75
State of the Art Recommendation Approaches: Their Issues and Future Research Direction in E-Learning A Survey

Bhupesh Rawat
Babasaheb Bhimrao Ambedkar University (BBAU), India

Sanjay K. Dwivedi
Babasaheb Bhimrao Ambedkar University (BBAU), India

ABSTRACT

Recommender systems have been used successfully in order to deal with information overload problems in a wide variety of domains ranging from e-commerce, e-tourism, to e-learning. They typically predict the ratings of unseen items by a user and recommend the top N items based on user's profile. Moreover, the profile can be enriched further by using additional information such as contextual data, domain knowledge, and tagging information among others for improving the quality of recommendations. Traditional approaches have not been effective in exploiting these additional data sources. Hence, new techniques need to be developed for extracting and integrating them into the recommendation process. In this article, the authors present a survey on state of the art recommendation approaches their algorithms, issues and also provides further research directions for developing smart and intelligent recommender systems.

INTRODUCTION

E-learning is an acronym for electronic learning which makes learning content available online for learners to access from anywhere and at any time. It is broadly inclusive of all the forms of education that make use of technology and multimedia. Online communication with real people may or may not be involved but the focus of E-learning is usually more on the learning content than on the communication between learners and tutors (Fetaji et al., 2011). E-learning systems provide huge benefits in comparison to traditional education systems (Zhang et al., 2004). However, they have several drawbacks

DOI: 10.4018/978-1-7998-0951-7.ch075

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

such as users spending a lot of time in searching the required information and same learning material is provided to all the learners regardless of their interest, skill, and level of knowledge among others. Recommender systems play an important role in this context by suggesting relevant items to learners based on their profile.

Recommender systems have drawn the attention of the research community since the appearance of the first paper on collaborative filtering (Adomavicius et al., 2005). Existing literature (Lu et al., 2015) suggests that among all the recommendation techniques, collaborative filtering, content-based, hybrid and knowledge-based have been used relatively more in the domain of e-government, e-business, e-commerce, e-library, e-learning, and e-resource.

The rest of this article is organized as follows. Firstly, we discuss major recommendation approaches, their critical issues, algorithms and input/output required by them. We also present some popular real systems based on these approaches. Then we provide further research direction which could help in developing more intelligent recommender system. Finally, the paper concludes with the discussion section.

BACKGROUND

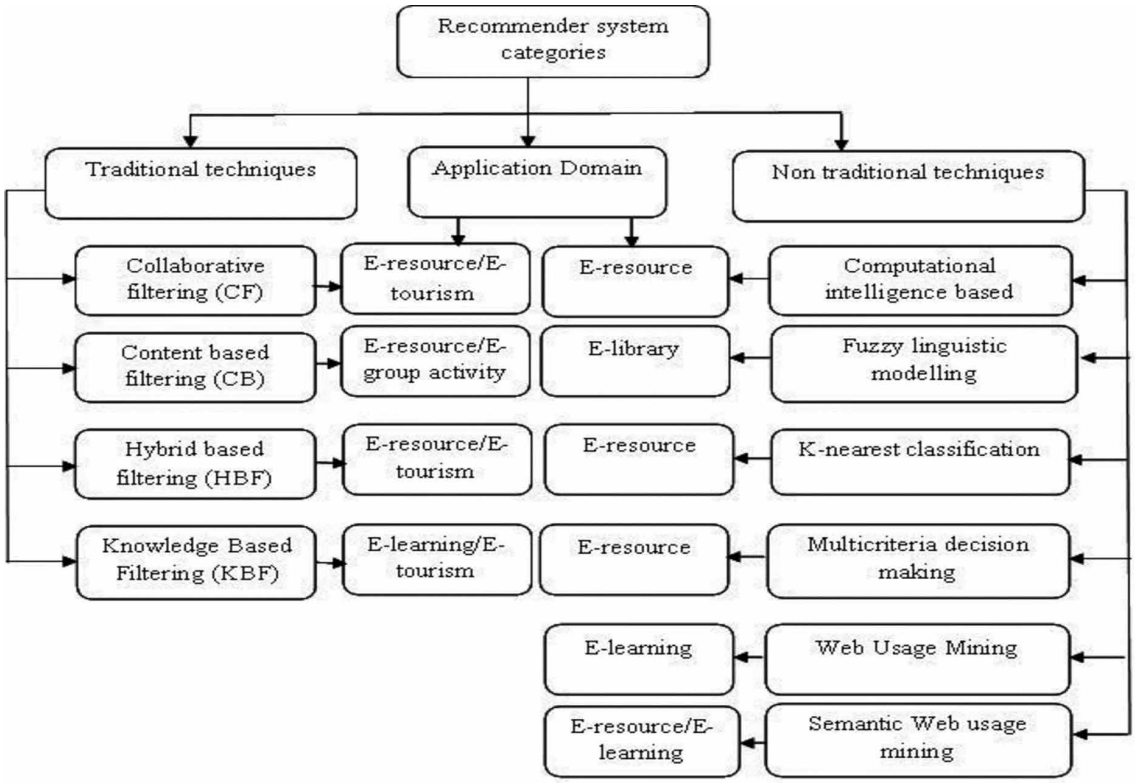
The roots of recommender systems lie in cognitive science (Rich, 1979), information retrieval (Salton, 1989) and approximation theory (Powell, 1981). The recommender system appeared as an independent research area in the mid-1990s (Resnick et al., 1994). Recommender systems are software tools suggesting for items to be of use to a user (Resnick et al., 1997). They assist users in selecting the right item from a vast repository on the web, hence overcoming information overload problem. Conventional recommendation approaches such as content filtering, collaborative filtering, and more recently developed web usage mining have been successfully used in several domains. However, they are unable to exploit additional sources of information such as domain knowledge, tag information, and annotation among others which could significantly improve the quality of recommender systems (Pazzani et al., 1997). Non-traditional approaches such as computational intelligence, semantic web usage mining, multi-criteria decision making, and fuzzy linguistic modeling have recently emerged as an extension of the traditional approaches. In conventional systems, recommendations are made based on the profile of user which comprises of either user's ratings or description of item experience by a user (Gemmis et al., 2009). Moreover, users' unwillingness to provide ratings leads to sparse "user item rating matrix" which causes poor quality of recommendations. However, additional information about user and item can enrich a user's profile and hence improve the quality of recommendations.

TYPE OF RECOMMENDATION APPROACHES

We have broadly categorized recommendation techniques into "traditional techniques" and "non-traditional techniques" as shown in Figure 1. Unlike earlier survey papers on recommendation systems, this paper also includes web usage mining (WUM) and semantic web usage mining (SWUM) which play a major role for developing next generation of recommender systems.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. Types of recommendation approaches with their domains



Collaborative Filtering (CF)

Collaborative filtering (CF) is one of the most widely used approaches in the recommendation system (Herlocker et al., 1999). They recommend an item by matching the preferences of a target user with similar users (also called neighborhood of the target user). CF-based systems are classified into memory based and model-based. The memory-based methods compute the similarity of an active user with similar users without creating a model (Rafter, 2010). They are not fast and scalable which is required in real-time systems, hence model-based approaches were developed. They use "user-item ratings" for building a model without using the entire dataset which provides scalability and reduces runtime. They make use of the Bayesian network, clustering, and rule-based approaches to improve the quality of recommendation systems (Portugal et al., 2015).

The inputs to collaborative filtering systems are a set of all users which is represented by the symbol 'U' and a set of all possible items which is denoted by the symbol 'I'. The goal of CF is to predict the rating of item i by a user u which is denoted by R(u,i). User's ratings are stored in a two-dimensional "user-item rating matrix". The output of collaborative filtering system is a set of items (i_1, i_2, i_3……i_k) where k <= I which are ranked and top N items are suggested to the user.

Collaborative systems have been used in many application areas such as movies, article, and products among others. However, altered Vista (AV) is one of the first collaborative systems used for recommending learning resources (Walker et al., 2004). The AV comprises a database of learners' evaluations

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

of learning resources which can be shared to other learners' as word of mouth recommendations about the quality of resources. "Web-based Peer Grader" is another popular CF based system which aims to improve learners' skill by reviewing and evaluating the solutions of their peer (Gehringer et al., 2001). LARGO (Pinkwart et al., 2006) is another popular collaborative system in the field of legal argumentation to evaluate the verdict (decision rule) of the supreme court of US. It makes the interpretation of legal argument textually. Diagrams are used within LARGO to visualize court arguments as hypothesis testing. The approach is designed to support first-year law students in learning legal argumentation skill by creating a graphical model of oral arguments and presenting them weakness in their models. The approach is able to reveal three types of weaknesses in arguments. LARGO uses graph grammar to handle structural and context weaknesses. Collaborative filtering is employed to handle content weaknesses. RACOFI (rule applying collaborative filtering) (Anderson et al., 2003) approach integrates two recommendation techniques. Ratings are elicited from learners on learning resources and inference engine is used for discovering the association between learning resources to be recommended to learners. Scaffolded writing and rewriting (SWoRD) is another popular collaborative system. Due to long text homework, tutors are unable to evaluate the whole text consequently learners don't get feedback on their performance. The tool deals with this problem by making use of peer review process. Experimental evaluation of the tool shows the improvement in the performance of learners as compared to single expert feedback systems (Cho et al., 2007).

Collaborative filtering systems could be benefitted by using the tagging information associated with social networking sites (Mathes, 2004). It helps to enrich a user's profile which leads to recommend more relevant items to a user. Tagging allows users to label their resources (images, audio etc) and assign them some meaningful words called tags (Golder et al., 2005). These tags can also be used to retrieve user information. Users can assign any tag to their resources; however, it leads to the problem of identifying most appropriate tags.

The collaborative approaches are more reliable than content filtering as they can capture 'realistic relationships' among items based on similar preferences of identical users. However, CF-based systems suffer from sparsity (Sarwar et al., 1998) and cold start problem (Iaquinta et al., 2011). Sparsity means not having adequate users' ratings to compute correct prediction. As the rating matrix gets sparse scalability makes it difficult to generate recommendations with the huge dataset. The cold start problem occurs when the system does not have enough ratings to make a recommendation. However, CF-based systems provide reasoning of items being recommended which is one of the important features of a recommender system.

Majority of collaborative filtering systems use k-nearest neighbor (Cover et al., 1967), association rule mining algorithm (Aggarwal et al., 1994) and matrix factorization algorithm (Takacs et al., 2008) among others for making recommendations. The k-NN is more suitable for building collaborative systems due to their similar functionality (Adomavicius et al., 2005). The matrix factorization is used because it allows us to discover latent features underlying the interaction between user and item(Yeung,2010). Collaborative filtering systems based on association rule mining algorithm are quite effective as compared to other recommendation approaches (Lin, 2000).

Content-Based Filtering (CB)

Content-based recommender systems make a prediction using the features of an item rather than ratings of other users (Basu et al., 1998). They consist of "content analyzer", "Profile learner" and "Filtering component". The inputs to the "content analyzer" are web pages, documents and product description.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The content analyzer aims to perform initial pre-processing on unstructured data (i.e. text) to make it suitable for the next processing step. The "Profile learner" takes input from the "content analyzer" in the form of "structured item representation" and generates "general user profile" using machine learning algorithms. This component also requires user feedback for updating user's profile to recommend more relevant items. The "filtering component" takes "user profile" as input for recommending relevant items to users by matching their profile with the items to be recommended.

CB approaches are unaffected by sparsity and gray sheep problems due to their independence from "user-item ratings". Moreover, they can be utilized in domains such as movies, news article, web page, and television program.

Majority of recommender systems based on content filtering use Bayesian approach (Ericson, 2013), decision tree (Lucas, 2012), neural network (Alvarez, 2011), rule learning (Hariri, 2013) among others. Bayesian approach is used due to its reduced complexity in Bayesian calculations. Content filtering systems based on neural network are gaining popularity due to their ability to differentiate between two different classes of items rapidly and accurately.

CB approaches are categorized into case-based reasoning (CBR) and attribute-based techniques (ABT). CBR recommends those learning objects which are strongly correlated to the objects which a user liked in the past. The quality of items is improved as more users' ratings are collected in the system. However, it suffers from overspecialization and sparsity problems. Overspecialization occurs due to recommending similar items to a user repeatedly. CBR is more useful in teaching and learning field. For example, Pixed (Project Integrating eXperience in Distance Learning) is an implementation of CBR which is a hypermedia ontology-based system. In yet, another CBR based system best navigational path is provided to the learner (Huang et al.,2006). An instructional planner was proposed in (Elorriaga et al., 2000) which adapts the sequences observed in logs to create instructional sequences for a complete course. Attribute-based techniques recommend objects by matching their attributes to the learner profile. They address 'cold start' problem as their recommendations are not based on the learner behavior data.

Existing literature suggests that content-based filtering has been used relatively less in e-government domain and more in e-resource and e-group (Lu et al., 2015). This is primarily due to the privacy policy of data which makes it difficult to get users and items data from the government organization. CB approaches have the following issues:

Cold Start Problem

Cold start problem occurs when a recommender system does not have enough ratings either of a user or item to make recommendations to a user. This is due to either the user has not rated items or a new item was added to the site recently which makes it difficult to provide high-quality recommendations. Cold start problem is of three types: New user problem, new item problem, and new system problem. The new user problem occurs when a user has not used an item from the repository. The new item problem arises when a new item is added to a site without any ratings from a user and the new system problem arises when the system is new and has just started to operate and does not have any ratings either from the user or the items. The cold start problem has its origin in recommendation system (Iaquinta et al., 2011).

One of the most challenging tasks of the recommendation system is to improve the recommendation accuracy for the new user (inactive) or new item (rarely rated item). Cold start problem causes missing ratings in user-item rating matrix. Several solutions (Zhang et al., 2010; Iaquinta et al., 2011) were suggested to address this issue. Furthermore, missing ratings could also be filled with their default values

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

such as the middle value of average user or item (Breese et al., 1998). Moreover, content filtering approach (Degemmis et al., 2007) and autonomous agents (Good et al., 1999) have also been used to deal with missing ratings.

Sparsity

Sparsity refers to insufficient information in the "user-item matrix" (R) for making recommendations controlling the quality of recommendations (Sarwar et al., 1998). With the huge number of items and users in many large-scale applications, the rating matrix can still be extremely sparse even when many events have been recorded which leads to the situation where the similarity between two given users is zero, rendering collaborative filtering useless. Due to sparsity, a rule-based recommender system cannot give any recommendations. Sparsity occurs because some users' rate only a very small fraction of the available items, hence making it difficult to find a sufficient number of common items in multiple user profiles. The problem often occurs in web usage mining and collaborative filtering as compared to other recommendation approaches. In (Chen, 2011), the authors relied on some standard dimensionality reduction techniques to alleviate sparsity. However, this approach suffers from the problem of some of the interesting items may be removed from the final recommendation list. The authors (Owoc et al., 2012) suggested two solutions to this problem. The first solution is to rank all the discovered rules based on the degree of intersection between the left-hand side of each rule and the user's active session and then to generate the top k recommendations. This approach eliminates constraint of having to obtain the complete match with the left-hand side of the rules. The second solution is to employ collaborative filtering in which the system finds "close neighbors with similar interest to the target user and makes recommendations based on the close neighbor's history. In (Popescul et al., 2001) they used probabilistic models for unified collaborative and content-based recommendations in the sparse data environment.

Scalability

With the increasing number of users and items, the existing recommender systems need more resources such as CPU and memory for processing of large information. Most of these resources are used for identifying users with similar taste and descriptions. This problem is solved by considering the combinations of various types of filters and system physical improvement. Moreover, several attempts (Georgiou et al., 2010; Aggarwal et al., 1999) have been made to deal with the problem of scalability. One of the widely used methods is matrix factorization (Koren et al., 2009) which provides several latent models such as SVD (single value decomposition) and SVD++. The latter is an advanced version of SVD and is used to incorporate other sources of feedback from the user (Koren et al., 2011) in order to improve prediction accuracy. The goal of latent models is to convert user and item variables to the same latent factor space.

Over Specialization Problems

In over specialization a user model solely relies on tags in his rated items (Kim et al., 2011). This makes it difficult to recommend new items other than the user has previously rated. For example, if a user is interested in "action movie" but not comedy movies then he/she will never get recommendations for the movie which is a mix of action and comedy. Users are restricted to receive recommendations which strictly match with the information in their profile. Overspecialization prevents the user from discover-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

ing new items and various available options other than suggested by his profile. However, the diversity of recommendation is an important feature of all recommendation system. The problem is solved using genetic algorithms and provides a wide range of alternatives (Silva et al., 2010). In (Alphy et al., 2015) they have proposed WebBluegillRecom-annealing dynamic recommender system which gives dynamic recommendations each time a user visits a site thus overcoming specialization issue. In yet another work (Ruiz-Iniesta et al., 2012) a recommendation framework is proposed which is a prototype of the knowledge-based system and overcome the problem of overspecialization. Knowledge-based filtering is also used to address overspecialization (Trewin, 2000).

Hybrid Based Recommender Systems (HBRS)

Hybrid Based Recommender Systems combine two or more recommendation techniques such as collaborative filtering (CF), content-based (CB) or knowledge-based (KB) in order to utilize the strength of each one to improve the overall quality of recommendations (Burke, 2002). For example, CF relies on user ratings for predicting the ratings of a new item to be recommended. On the other hand, CB approach does not depend on ratings from a user and based its recommendations on the features of items. Numerous combinations can be created by using the basic recommendation techniques (Burke, 2007) in which he has discussed existing and possible HBRS and also proposed a novel system namely Entrée which combines knowledge-based approach and collaborative approach for recommending restaurants.

Existing literature shows that hybrid models perform better than single recommendation techniques (Soboroff et al., 1999; Pazzani, 1999). In (Savic et al., 2011) they have presented a multi-criteria hybrid recommender system based on analytic hierarchy process(AHP), trend analysis and data comparison in order to evaluate the quality of e-learning. The goal of AHP is to view a problem in terms of hierarchy starting from the top through the intermediate level to the lowest level.

Although hybrid recommender systems have been successfully employed in wide variety of applications, they have the following issues:

1. The constraint with switching hybrid (one of the forms of hybrid techniques) is to have prior information of confidence value or external criteria.
2. Another issue with HBRS is of producing the new rank score based on the individual ranks generated by each component. One of the solutions to this problem is simply adding the individual ranks such as KB_rank (2) + CB_rank (4)->Mixed_rank(6).

Hybrid filtering systems require various input such as user profile (which stores users' preferences about item) and contextual parameters such as (time of day, time of year), demographic data (age, location, gender, etc.) or community data, item or product features (title, genre, actors), different knowledge models (such as rule-based, Bayesian network, etc.). The output of HBRS is a list of items (i_1, i_2, $i_3 \ldots i_n$) with their predicted score such as (0.8, 0.3, 0.5).

Existing literature shows that HBRS are mostly used in the domain of e-government, e-resource, and e-tourism (Lu et al., 2015). A comprehensive study on the possible combination of basic recommendation techniques has been reported in (Burke, 2002).

Among several machine learning algorithms, ANN (Artificial neural networks) has shown good results in combining output from various recommendation models (Hsu et al., 2007). Another approach used is an ensemble of a classifier in building hybrid recommender system as reported in (Tiemann et al., 2007).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Knowledge-Based Recommender Systems (KBRS)

These systems predict the utility of an item by using some appropriate utility function even when no knowledge about the preferences of an item is available. Case-based reasoning system (CBRS) is one of the popular KBRS (Bridge et al., 2006). A suitable similarity function is employed for learning how an item fulfills user's need. One of the forms of CBRS is "constraint-based recommender systems" which are used when a solution is needed in order to fulfill a specific requirement of the user (Felfering et al., 2008). Some conditions must be satisfied by such system for making effective recommendations. For example, if a client 'X' wants to purchase customized software from a company let's say 'ABC' then the set of constraints imposed by the client might be the following: a) A client asks for distributed software system b) The software should be run on the open source platform such as Linux/UNIX c) the client also specifies the number of users who will be using the software d) He states hardware constraints such as minimum required memory and type and speed of CPU. In this scenario, case-based recommender system would be more appropriate as collaborative filtering requires frequent ratings of an item over a period of time which is not possible in the above situation where the client would purchase software once or twice a year from the company which leads to sparse rating matrix and results in collaborative technique useless.

The following issues need to be addressed for the success of KBRS:

1. One should be able to explicitly define knowledge base.
2. Good domain knowledge is required.
3. Must have expertise in representing domain knowledge.

The following inputs are required by KBRS for achieving their goal:

1. User properties, such as type of learner (beginner or advanced), learning style (visual learner or pragmatic learner)
2. Item properties, for example, fiction, non-fiction among others
3. Constraints, they are provided by learner such as courses must be recommended depending upon the type of learner
4. A part of an ontology, to be obtained from database schema
5. Contextual knowledge
6. Filter conditions, provided by a learner such as a book's price should not exceed as stated by a learner

The output of KBRS is the set of items to be recommended. KBRS in particular "case-based recommender" employs several technologies (Felfering et al., 2008).

Several techniques have been employed in the construction of knowledge-based recommender systems. Some of the popular ones are critiquing which is a popular interaction style and was first proposed by (Burke,1996), natural language processing (Zanket et al., 2008), finite state automation (Felfernig et al., 2006).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Computational Intelligence Based Recommender Systems (CIBRS)

Computational intelligence based recommender systems can recognize patterns, exhibit computational adaptability, have fault tolerance, and their error rate approximate human performance (Bezdek, 1994). Computational intelligence is a discipline of artificial intelligence which uses various techniques such as fuzzy sets (fuzzy logic, fuzzy numbers), artificial neural networks (Backpropagation neural networks, self-organizing map, feedforward networks), evolutionary computing (genetic algorithms, genetic programming, evolutionary programming), swarm intelligence and artificial immune system among others. These technologies have made recommender system more robust, effective and context-aware (Huang et al., 2006).

For recommender system based on computational intelligence, the input consists of user preferences from an active user and contextual parameters and the output produced is top N recommendations (list of items with their predicted value).

The reviews of existing literature on CIBRS show the following issues:

1. Overfitting is one of the problems of neural networks (NNs).
2. The black box effect is another problem which is due to the difficulty in selecting its parameters.
3. Neural networks take relatively more processing time due to training with a large number of samples data.
4. Difficulty in interpreting learned neural network.

Computational intelligence techniques have been known for not only improving recommendation accuracy but also able to deal with issues found in traditional recommendation approaches such as sparsity, cold start, and scalability. Majority of computational intelligence techniques such as fuzzy sets (Zhang et al., 2013), artificial neural networks (Devi et al., 2010) and evolutionary computing (Hernandez, 2008) among others have been used in building context-aware recommender systems.

Fuzzy Linguistic Modelling (FLM)

Fuzzy linguistic modeling is a tool which uses the concept of a linguistic variable in order to deal with qualitative assessment (Zadeh, 1975). It has been used in varieties of the problem such as decision making (Herrera et al.,1996) and clinical decision making (Degani et al., 1998), political analysis (Arfi, 2005) among others. FLM is of two types namely the 2-tuple FLM (Herrera et al., 2000) and the multi-granular FLM (Herrera et al., 2000).

There have been several works reported based on fuzzy linguistic modeling. In one such work (Yager,2003) fuzzy set is used in order to represent an item in a recommender system. In yet another work (Zenebe et al.,2009) the authors have defined a set of properties for an item and set of values for each property. Existing literature suggests that Fuzzy linguistic modeling is mostly used in the domain of e-government, e-commerce, e-library and e-learning (Lu et al., 2015).

In spite of their widespread applications, fuzzy linguistic modeling has the following issues:

1. Difficulty in the interpretation of TFN (trapezoidal fuzzy numbers)
2. The granularity of uncertainty.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In order to produce the desired output fuzzy linguistic modeling requires the following inputs and output:

1. Crisp input, for example: high (1), low (0)
2. Fuzzy input, obtained after fuzzification between 0 and 1
3. Fuzzy output, obtained after passing through fuzzy inference engine
4. Fuzzy rule database, for example if… then
5. Crisp output, obtained after defuzzification

Recommender systems based on fuzzy linguistic modeling employ Markov chain model (Huang et al., 2009), OWA (ordered weighted averaging functions) which have become very popular among the fuzzy community for associating a weight with a relative value of an input (Yager,1988).

K-Nearest Neighbours (k-NN)

The k-nearest neighbor is a non-parametric method which stores the entire training dataset in memory and compares the test instance with each of the training instances in order to find its class. K-NN is simple and effective in many cases (Hand et al., 2001).

Although *k*-NN has been used as one of the widely used approaches of collaborative filtering, it has the following issues:

1. To determine the value of k(number of neighbors)
2. Formation of the neighborhood is computationally expensive
3. The requirement of large memory in case of a big dataset

K-NN needs various inputs such as training dataset, N input/output pairs. The output consist of $y_i \in \{1, ..., C\}$ for classification; a discrete variable for classification; a discrete variable $y_i \in R$ for regression; a continuous (real-valued) variable.

Among all instance based classifier such as KStar (John et al., 1995), LWL (Locally weighted learning) (Frank et al., 2003) and IBk (Aha et al., 1991), IBk is widely used for building collaborative filtering (CF) system due to its functional similarity with CF.

Multi-Criteria Decision-Making Recommender Systems

Multi-criteria recommender systems in contrast to traditional single criteria recommender systems consider multiple criteria in order to compute the utility of an item for a particular user (Adomavicius et al., 2005). For example, in traditional systems rating function is defined as follows:

$$R: U \times I \rightarrow R_0 \quad (1)$$

where U is the set of all users, I is the set of all possible items to be commended and R_0 is the predicted rating. However, single criteria recommender systems are unable to compute the utility of an item effectively as an item (movie) might be rated from different aspects such as story and visual effects in a

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

movie recommendation system. A user might provide ratings of 3 to story and 5 to visual effects. On the other hand, a multi-criteria recommender system predicts the utility as follows:

$$R: U \times I \rightarrow R_0 x \times R_1 x \times .... \times x R_k \quad (2)$$

The input to multi-criteria systems is same as single criteria recommender systems with the only difference of considering multiple attributes of an item and asking users' preference on these attributes. The output is a set of top N items to be recommended for the user.

MCDM's (multi-criteria decision making) have the issue of intrusiveness which occur due to explicitly asking additional information from a user. Several works have been reported to address this issue (Caglayan et al., 1997; Konstan et al., 1997). Another issue is handling incomplete data which occurs more in the multi-criteria systems as users are usually unwilling to provide ratings and these systems require more data from a user on multiple attributes of an item. Expectation maximization algorithm has been successfully used to deal with this issue (Ghahraman et al., 1994). However, several other machine learning techniques could also be explored in this context (Portugal et al., 2015). The issue of constructing the item evaluation criteria is critical in multi-criteria systems as it directly affects the quality of recommendations. MCDM requires information on several attributes from a user hence they must be carefully chosen in order to compute the utility of an item.

Algorithms employed by multi-criteria recommender systems are broadly classified into heuristic (Adomavicius, 2007; Manouselis et al., 2007) and model-based approaches which include, probabilistic modeling (Adomavicius, 2007) and multilinear singular value decomposition (Li et al., 2008).

Web Usage Mining (WUM)

Web mining is the application of data mining techniques to the content, structure, and usage of Web resources (Liu,2011). It is broadly categorized into web content mining, web structure mining and web usage mining. Web content mining concentrates on the content of individual pages, which is contained in the HTML (Hypertext markup language), script and code that generates a page. In (Sebastiani, 2002) they have presented a detailed review of text mining.

Web usage mining refers to the use of data mining techniques to automatically discover and extract useful information from the web document and services (Facca et al., 2005). It is becoming more and more popular among users due to easy access to the log file as well as its applicability in CRM (Customer relationship management). The goal of CRM is to provide not just the quality products and services but also make, retain and grow customers (Kotler,2004). This is possible once we know the needs and preferences of the customers by applying data mining techniques to customers' usage data in order to get the profile of customers. In Web usage mining, the primary web resource is a record of the requests made by visitors to a Web site, most often collected in a Web server log (Srivastava et al., 2000).

WUM process consists of three major steps. The first step is "data preparation" which takes its input as "site files" and "web server log." The data preparation step comprises several sub-steps such as data cleaning, user identification, transaction identification, usage statistics, and support filtering. The output of this step is "user transaction file" which is fed to the "usage mining phase." The transaction records are inputted to the various mining algorithms such as clustering and association rule mining which outputs "frequent itemsets" and URL cluster. The third step is generating recommendation which takes its input from the "usage mining stage" and outputs set of rules which are recommended to a user.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Majority of recommender systems based on web usage mining uses association rules, sequential patterns (Sungjune et al., 2008), and clustering (Liu, 2011). The major application areas for WUM are personalization, system improvement, site modification, business intelligence and usage characterization (Srivastava et al., 2000). Moreover, web content or site structure can be integrated with usage data in order to improve the accuracy of the personalization system (Liu, 2011). Web usage mining primarily collects its data from the log file of a web server which contains all the users' data visiting the site such as IP address, pages visited and visiting time among others. It collects, analyzes and process this data to discover user access patterns. The result of this analysis provides us with useful insight into interesting usage information and patterns which can be further used to predict the web pages a user is likely to be interested. Table 1 presents recommender systems based on web usage mining using popular machine learning algorithms in e-learning:

Issues of Web Usage Mining

Although web usage mining techniques have been able to overcome the limitations of traditional recommendation approaches, they have the following issues:

Table 1. E-learning recommendation systems based on web usage mining

S.No	Methods and Algorithms	Results
1	The Combination of k-means and apriori algorithm(Aher et al.,2013).	Recommending best combination of courses
2	Naive Bayes algorithm (Kotsiantis et al.,2003).	Predicting students with high probability of drop out
3	Web usage mining and decision tree (Cho et al.,2002).	Improved and effective product recommendation
4	Web usage mining and information retrieval(Khirbi et al.,2008).	Recommend relevant links to an active learner.
5	Multi-attribute evaluation and fuzzy method(Lu,2004).	Providing learning material recommendations.
6	Association rule mining (Zaiane et al.,2002).	Suggesting activity or shortcuts to learners.
7	Web usage mining coupled with content based and collaborative filtering(Claypool et al.,1999).	Recommending relevant links to active user.
8	Association rule mining and collaborative approach for recommendation to courseware Filtering (Garcia et al.,2009).	Courseware Recommendation
9	Bayesian approach (Bekele et al.,2005).	To predict the performance of a student
10	Content-based analysis, collaborative filtering and data mining techniques (Hsu et al.,2006).	To generate recommender score by which help the student to select appropriate lessons?
11	Combining oversampling techniques and cost sensitive learning (Nguyen et al.,2009).	For improving prediction/ classification results.
12	Collaborative algorithm (Michael et al.,2007).	For recommendation of online course Enrollment.
13	Collaborative, Content-Based and Demographic Filtering (Pazzani 1999)	For recommending restaurants.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Privacy

Privacy is one of the major issues of web usage mining. It is drawing a lot of attention of research community due to constantly growing nature of the world wide web. The issue of privacy is based on the fact that majority of users want anonymity while surfing the web. While administrators are interested in discovering users' access patterns which requires applying data mining techniques to user's browsing data, users are generally unwilling to disclose their browsing information. Moreover, this information would help in improving the design of a site and meet the requirements of different users leading to improving user's browsing experience. The challenge here is to make a compromise where user's identity is not disclosed and site administrators can get required data for their analysis. Furthermore, there should be regulations in place which prevent user's data being sold to third party and site administrators can perform analysis on usage data. The users should also be informed about the site's policy so that they can make an informed decision about disclosing their personal information. The world wide web consortium has started an initiative called P3P (*Platform for Privacy Preference*). It allows a site to publish their privacy policies in a machine-readable format. When a user first visits a site, the browser matches the privacy setting of the site with the security settings configured by the user. If a match is found satisfactory then user's request is allowed otherwise a negotiation protocol is used which is acceptable to the user. Another aim of P3P is to provide guidelines for independent organizations which can ensure that sites comply with the policy statements they are publishing (Srivastava et al., 2000). The European Union has taken a lead in setting up a regulatory framework for internet privacy and has issued a directive which sets guidelines for processing and transfer of personal data. Some of the privacy issues have been reported in (Facca et al., 2005).

In order to get the most accurate and valid recommendation, the system must have a large amount of possible information about the user such as demographic data and the location of a particular user. Obviously, the question of reliability, security, and confidentiality of the given information arises. Many online shops provide effective privacy protection for the users by using specialized algorithms. Although in the U.S there is no unifying framework in place, U.S federal trade commission has recommended that Congress develop legislation to standardize the personal information of websites.

Dependency on Usage Data Alone

Recommender systems based on WUM suggest items using historical data found on the server due to which they miss certain objects added to a site recently which have not been consumed by a user. Moreover, without using the semantic knowledge they cannot recommend certain complex objects using their properties and attributes. For instance, let us consider an e-learning recommender system having few tables in its database namely course, instructor, book and so on. In the absence of semantic knowledge or ontological structure of the site, we might end up recommending objects only on the basis of features of the entities (tables). Moreover, we might also miss some important relationships which exist among those tables and have the potential to improve the quality of recommendations. However, incorporation of semantic information in the different phases of the web usage mining process allows a recommender system to infer new semantic relationships which were not obvious otherwise. For instance, in "pure usage-based e-learning system", a learner could be recommended "Advance operating system" course based on his profile. However, with the integration of semantic information with usage data we could suggest learner some "introductory level operating system course" in order to enable the learner to ac-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

quire some basic knowledge of the course before opting for the "advanced operating system" course or could also suggest him a course instructor. Such type of semantic relationship could only be discovered by enriching usage level transactions with semantic information.

Usage-based recommendation systems also cannot explain the reason behind a particular set of recommendations. However semantic web has successfully overcome this limitation by providing the domain knowledge representation language such as DAML+OIL (Horrocks, 2002), SHIP (DL) and description logic (Horrocks et al., 2001).

Pre-Processing of Click-Stream Data

Another issue with the web usage mining is pre-processing of users' click-stream data which ensures the quality of data before mining. It involves applying different algorithms and heuristics such as data cleaning, session identification, path completion, pageview identification among others to raw data (Cooley et al., 1999). We also found many data pre-processing tools such as WEKA (Waikato environment for knowledge analysis), DBMiner, Delta miner among others which can be used effectively and efficiently for the data preparation tasks (Goebel et al., 1999). Among all these tools, WEKA is the most popular one due to its user-friendly interface and open source nature.

Semantic Web

The Semantic web is the extension of the web 2.0 in a sense that adding semantic to the web makes the information meaningful. It can be used in a variety of applications areas such as data integration where data in various locations and formats can be integrated in one seamless application; In resource delivery and classification to provide better domain-specific search engine capabilities; In cataloguing for describing the content and its relationship available at a particular website; In describing content rating; In explaining collection of pages that represent a single logical document; for illustrating intellectual property rights of web pages among others. Moreover, semantic web helps improve the existing services available on the web. For example, many e-portals, such as e-learning, e-government, and e-commerce were in need of semantic structure and intelligent logic. This goal is achieved by integrating two fast-developing research areas such as semantic web and web mining. Currently, they are among the most discussed topics in the research community. They can be combined successfully in order to achieve the common objective of understanding huge data in an automated way and also discovering and obtaining meaningful results. Semantic tools such as web services and semantic web agents can be used to achieve the goal of personalization of services over current e-learning portals and course management system (Mustapasa et al., 2010).

Semantic Web Mining (SWM)

The semantic web is proposed by Tim Berners-Lee the inventor of the World Wide Web. In the words of Tim Berner-Lee "The semantic web is not a separate Web but an extension of the current web, in which information is given well -defined meaning, better-enabling computers and people to work in cooperation. Generally semantic refers to the meaning of a sentence. Semantic web mining involves the integration of domain knowledge into the web mining process. The Majority of research, in SWM, has focused on application areas such as web content and structure mining (Stumme et al., 2002).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Semantic Web Usage Mining (SWUM)

Semantic web aims to make the web content understandable to humans as well as machines, hence allowing the software agent to search for desired content, share information and knowledge in a format that other software agent can understand.

Recommender systems based on semantic web usage mining require various inputs such as "server log", "web page" and "domain information". The server log and web pages are fed to the usage processing step which outputs "semantic user transactions". Ontology processing is also performed which results in domain ontology and acts as input to the subsequent phases such as pattern discovery and online recommendations. "Semantic user transactions" and "domain ontology" are fed to the pattern discovery phase which outputs "semantic usage patterns" which are finally inputted to online recommendation engine for generating recommendations to the active user. Table 2 presents some of the most recommendation system based on well-known semantic web usage mining techniques.

Issues of Semantic Web Usage Mining

For semantic web usage mining to be successful, it has to address the following issues:

Table 2. E-learning recommendation systems based on semantic web usage mining

S.No	Methods and Algorithms used	Results
1	Domain ontology and aggregate usage profiles(Mobasher et al.,2002).	To recommend personalized learning.
2	Multi-agent system based on ontology(Rani et al.,2015).	Formal description of learning Content and Process.
3	An e-learning recommender system using ontology, web ontology language (OWL) and rule filtering(Shishehchi et al.,2010).	Help the learners choosing the suitable material.
4	Web-based learning management system based on semantic web technologies(Begam et al.,2013).	To recommend varied resources to different user based on their need.
5	A combination of Ontology, information retrieval, and data mining (Faoud et al.,2012).	To discover, categorize and retrieve of personalized resources.
6	Integration of ontology and web usage mining(Zhudhdar et al.,2009).	To improve the accuracy of recommendations.
7	E-learning system based on RDF/XML(Shah et al.,2011).	Provide many semantic web services such as registration, personalization, notification and ontology mapped tutorial pages.
8	E-learning system using the combination of the semantic web and web mining (Jain et al.,2012).	Personalization of e-learning services.
9	Personalized learning system based on ontologies and semantic web(Vargas et al.,2008)	To manage the personal profiles and identities of learners.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

How to Represent Domain Knowledge as Content Feature

Many sources of domain knowledge can be utilized in the mining and personalization phase such as textual data of a website. The semantic of a site to some extent depends on the content features which are related to items. This feature needs keywords, phrase, category name, etc. Content pre-processing is essential before domain knowledge can be represented using domain representation languages. It is defined as extracting relevant features from text and meta-data. If features are extracted from text then features are assigned different weights which show the relative importance of features, if they are retrieved from meta-data then weights are derived automatically. The integration of content features with personalization process is required when the site is rich in text. However, the difficulty with this approach is that it is not able to find the complex relationship that may exist among objects at a deeper semantic level due to inherent properties possessed by these objects. Machine learning algorithms can also be used to further pre-processing of content features. For instance, content feature's classification with concept hierarchy may be used to limit the discovered patterns to those having page views about a particular subject. Similarly using clustering, formal concept analysis or association rule mining in the feature space can lead to composite features representing concept categories or hierarchies. For a system to be able to recommend the various types of complex objects the system must use ontology and not just keywords for the objects (Clerkin et al., 2001).

How to Represent Domain Knowledge as Structured Data

In web usage mining, the focus is on a web transaction or a user profile which is usually composed of a group of page view names and query strings (extracted from Web server logs). Though these features do not possess the semantic information about a transaction, therefore a mapping needs to be done between these features and the objects they represent. There are many e-commerce sites which allow generating web pages by querying an operational database from which semantic information can be easily obtained. Machine learning techniques can be used when structured data is not available. Moreover, to perform further pre-processing the acquired knowledge should be easily understandable. So, a standard language is needed to represent extracted knowledge. DAML+OIL (Darpa agent markup language and Ontology inference layer) is an example of an ontology language that combines the Web standards from XML (Extensible markup language) and RDF (Resource description framework), with the reasoning capabilities from SHIP (DL) description logic (Horrocks, 2002; Horrocks et al., 2001; Unadkar, 2015). Relational models and probabilistic models can also be integrated in order to enhance web personalization process with domain knowledge. In (Anderson et al., 2002) they have used Relational Markov Model for personalization. In (Horrocks et al., 2001) they adopted the syntax and semantics of another ontology representation framework and SHOQ (D) (Description logic) to represent domain ontologies.

How to Create Mapping between Usage Level and Domain Level Instances

The mapping process involves assigning usage level patterns like URL or page view name to the instances in the knowledge base. For example instead of describing a user's navigational path as: "u1, u2, ..., un" (where ui is a URL (Uniform resource locator) pointing to a Web resource), we need to represent it using the instances containing in the knowledge base, like: "movie(name = The Hangover), movie(name = Ghostbusters), movie(name = Austin Powers).Moreover, with the help of pre-acquired concept hierar-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

chy, we may be able to infer that current user's interest is in the comedy category. This is also termed as semantic user profile which can further be used for the purpose of semantic pattern discovery and online recommendations.

Using Query Strings

We generally ignore the huge amount of information accumulated in the database or semi-structured documents in a site. Large database server frequently provides content integrated from various underlying servers and database. The dynamically created pages on such servers are based on queries with several parameters linked to the URL matching to the underlying scripts. By means of the web server query string recorded in the server log files, it is likely to recreate the response page. Usually, in well-designed sites, there is a semantic mapping between query parameters and objects (such as products and categories). When a user clicks on a certain item on the product page of a site, this request is sent and recorded in the web server. It also appears in the address bar of the browser as query strings. The item details can be retrieved from the database using item_no, item category, and the backend database. For example, two query strings from a hypothetical bookstore are given below. The name value pair which begins right after question mark containing item_no and item category can be utilized to learn the interest of a customer in various books. Based on this a recommender system can update the profile of a user which can further be used to suggest more books to a customer. The first string below shows the item_no=7859645 and item category=3456 which can be queried to the database and the corresponding item details retrieved from the server can be displayed on a response page.

1. http://www.xyz.com/app.cgi?action=viewitem&item=7859645&category=3456.
2. http://www.xyz.com/app.cgi?action=order&item=1234567&category=1234&coupanid=3456.

Table 3 and Table 4 shows that semantic web usage mining (SWUM) is frequently used in e-learning domain which is due to the availability of many open source e-learning systems such as Moodle, ATutor, Eliademy, Formal LMS, and OLAT among others (Cavus et al., 2014) which provide rich features and offer several functionalities to their users. Moreover, due to their user-friendliness and less cost in implementing these systems, many universities, institutions, and schools are adopting them in order to improve their learning process. The database schema of these systems is easily accessible and the emergence of several semantic web technologies such as DAML+OIL, SHIP(DL) and a wide variety of open source ontology tools makes it easy to extract and process domain knowledge. Moreover, semantic web usage mining requires usage data (extracted from the log file of a web server) and domain knowledge (part of content and structure of a site) which is easily available in these systems. Table 3 and Table 4 also presents a summary of various traditional and non-traditional recommendation approaches their benefits, limitations, along with real applications which are based on these approaches and corresponding domain where they have been used often.

FURTHER RESEARCH IN RECOMMENDER SYSTEM

The emergence of a plethora of social networking sites such as Facebook, tweeter, flicker, YouTube, LinkedIn, Instagram among others provide opportunities for researchers to exploit the additional informa-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Summary of recommender system approaches

Approaches	Benefits	Limitations	Application Examples	Frequently used Domain
Collaborative Filtering	No domain information required	Cold-start, New user	• Altered Vista(Recker et al.,2003) • LARGO(Pinkwart et al.,2006) • RACOFI(Rule-Applying collaborative filtering(Anderson et al.,2003)	E-resource
Content-Based Filtering	No domain information required	Sparsity, Cold start problem	• Pixed (Project Integrating Experience in Distance Learning (Heraud et al.,2004). • AACS (Attribute -Based Ant Colony System) (Yang et al.,2009).	E-resource/E-group activity
Hybrid Filtering	Improve item-user cold start problem	Slow performance, Time complexity	• Fab (Balabanovic et al.,1997). • CYCLADES(Renda et al.,2005)	E-resource/E-tourism
Knowledge-Based Filtering	Sensitive to preference change	Difficulty in Knowledge acquisition	• An e-learning recommender agent(Zaiane,2002) • Course Agent(Farzan et al.,2006)	E-learning/E-tourism
Fuzzy Linguistic Modelling	Strong mathematical support for range of operations in RS	Difficult to interpret results due to their inability to express them linguistically	• University digital library Recommender system(Porcel et al.,2010) • Research resource(Porcel et al.,2009)	E-government

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Summary of recommender system approaches

Approaches	Benefits	Limitations	Application Examples	Frequently used Domain
K-nearest classification	Faster and more accurate recommend ation	High computation cost due to comparison of test instance with each training instance	• Foxtrot(Middleton et al.,2004).	E-resource
Multi-Criteria Decision Making	Ease of use ,Scalable, Easily adjustable in size due to hierarchical structure	Incredible amount of input is required at each step of the process.	• Primitive cognitive network process(P-CNP) (Yuen,2012) • MCEQLS(Multipl e criteria evaluation of the quality of software) (Kurlovas et al.,2013)	E-resource
Web Usage Mining	Overcomin g deficiencies of traditional approaches such as collaborativ e filtering.	New item problem, Sparsity, scalability, overspecializa tion problem, security and privacy issues.	• LMOS (Learning material organization system)(Tsai et al.,2006). • LLA (Learner learning assessment system)(Hwang et al.,2003). • AHA!(Romero et al.,2009).	E-learning
Semantic Web Usage Mining(SWU M)	More relevant information, exchange of information becomes easier between program and sites.	Representatio n of domain knowledge, the mapping between usage data to domain level instances.	• SEWeP(Eirinaki et al.,2003). • Web-page recommender (Nguyen et al.,2013).	E-learning

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

tion associated with user's profile on these sites such as contextual information, tag, and annotation in order to improve the quality of recommendations. The additional information is labeled by users while uploading their resources such as images, videos among others which help to enrich the profile of a user (Golder et al., 2005). Moreover, tagging information could also help recommender systems to learn more accurately about their user's preferences. The integration of tag information with user preferences results in improving the accuracy of prediction, hence more relevant items could be recommended to a user. Collaborative tagging could further be explored to learn new ways to extract user's preferences effectively and efficiently. Another emerging area is "semantic web usage mining" (Berendt et al., 2002) which holds a lot of potentials to assist recommender systems in recommending more accurate items. A considerable amount of information about user and item is underlying a website which is hidden and unexplored (Third, 2012). Moreover, this information can be extracted from the content and structure of a site using tools from semantic web to discover the semantic relationship among objects which is usually missing in usage-based recommendation system (Stumme et al., 2006). Furthermore, more intelligent recommender systems could be developed by exploiting and integrating the semantic knowledge in different stages of the web usage mining process.

Recommender systems based on multi-criteria decision making needs to be investigated further. They learn user's preferences by asking for information on multiple attributes of an item in contrast to single criteria recommender systems. However, multi-criteria systems present several issues such as the issue of intrusiveness. Instead of interrupting a user by getting information on multiple attributes of an item, it would be better to collect information implicitly which reduces user involvement. New methods based on multi-criteria decision-making need to be developed in order to implicitly collect data from users.

Furthermore, construction of multi-criteria (an item having multiple attributes) in terms of identifying and specifying functions that represent preferences of users could further be explored in multi-criteria recommender systems(). For example, in a "book recommender system" a book is treated as an item which could have various properties such as "type of book" (fiction or non-fiction), "book language" (French, English) type of author(national, international) among others and non-fiction attribute might have its own subcategory such as "non-fiction by discipline" or "non-fiction by topic" or "non-fiction by writers". Furthermore, non-fiction by topic could further have another level of attributes such as "a book about the arts" or "business books." Hence, there could be a hierarchy of attributes and each level could have one subcategory of attributes. These attributes should be selected carefully so as to have a meaningful relationship with them which finally leads to improving the quality of recommendation. Overall, we can say that by exploiting additional sources of information leads to the development of next generation of recommender systems.

CONCLUSION

We have discussed various recommendation approaches especially web usage mining and semantic web usage mining, their algorithms, issues and additional data sources for improving the quality of recommendation system. While several attempts have been made earlier to address the issues of traditional recommender systems, these systems still have not been able to fully exploit additional information. Moreover, the domain knowledge could be represented either as the content feature or structured data depending on the application. Various classification algorithms were also reviewed for creating map-

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

pings from usage data to domain level instances. However, other machine learning algorithms could also be investigated in this context. Moreover, query strings can also be utilized to exploit enormous information stored in a database or semi-structured documents associated with a site in order to learn more about a user's preferences. The recommender systems based on above-discussed approaches can be made smarter and intelligent by incorporating additional information such as tag, annotation and domain knowledge in the recommendation process.

REFERENCES

Adomavicius, G., & Kwon, Y. (1999). Multi-criteria recommender systems. In Recommender Systems Handbook, 847–880.

Adomavicius, G., & Kwon, Y. (2007). New Recommendation Techniques for Multi-Criteria Rating Systems. *IEEE Intelligent Systems*, *22*(3), 48–55. doi:10.1109/MIS.2007.58

Adomavicius, G., Sankaranarayanan, R., Sen, S., & Tuzhilin, A. (2005). Incorporating contextual information in recommender systems using a multidimensional approach. *ACM Transactions on Information Systems*, *23*(1), 103–145. doi:10.1145/1055709.1055714

Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, *17*(6), 734–749. doi:10.1109/TKDE.2005.99

Aggarwal, C. C., Wolf, J. L., & Wu, K.-L., & Yu. P. S. (1999). Horting hatches an egg: A new graph theoretic approach to collaborative filtering. In *Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 10.1145/312129.312230

Aggarwal, R., & Srikant, R. (1994). Fast algorithms for mining association rules in large databases. In *Proceedings of the 20th International Conference on Very Large Data Bases*.

Aha, D., Kibler, D., & Albert, M. K. (1991). Instances based learning algorithms. *Machine Learning*, *6*(1), 37–66. doi:10.1007/BF00153759

Aher, S., & Lobo, L. (2013). Combination of machine learning algorithms for recommendation of courses in E-learning system based on historical data. *Knowledge-Based Systems*, *51*, 1–14. doi:10.1016/j.knosys.2013.04.015

Alphy, A., & Prabhakaran, S. (2015). A dynamic recommender system for improved web usage mining and CRM using Swarm Intelligence. *The Scientific World Journal*, *2015*, 1–16. doi:10.1155/2015/193631 PMID:26229978

Alvarez, S. A., Ruiz, C., Kawato, T., & Kogel, W. (2011). Neural expert networks for faster combined collaborative and content-based recommendation. *Journal of Computational Methods in Sciences and Engineering*, *11*(4), 161–172.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Anderson, C. R., Domingos, P., & Weld, D. S. (2002). Relational markov models and their application to adaptive web navigation. In *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 143-152). 10.1145/775047.775068

Anderson, M., Ball, M., Boley, H., Greene, S., Howse, N., Lemire, D., & McGrath, S. (2003) RACOFI: a rule-applying collaborative filtering system. In A. Ghorbani & S. Marsh (Eds.), *Proceedings of IEEE/ WIC COLA'03* (pp. 53–72). National Research Council Canada.

Arfi, B. (2005). Fuzzy decision making in politics. A linguistic fuzzy-set approach (LFSA). *Political Analysis*, *13*(1), 23–56. doi:10.1093/pan/mpi002

Balabanovic, M., & Shoham, Y. (1997). Fab: Content-based, collaborative recommendation. *Communications of the ACM*, *40*(3), 66–72. doi:10.1145/245108.245124

Basu, C., Hirish, H., & Cohen, W. (1998). Recommendation as Classification: Using Social and Content-based Information in Recommendation. In *Proc of the 15th National Conference on Artificial Intelligence* (pp. 714-720).

Begam, F., & Ganapathy, G. (2013). Adaptive learning management system using semantic web technologies. *International Journal on Soft Computing*, *4*(1).

Bekele, R., & Menzel, W. (2005). A Bayesian Approach to Predict Performance of a Student (BAPPS): A Case with Ethiopian students. In *Proceedings of the international conference on Artificial intelligence and Application* (pp. 189-194).

Berendt, B., Hotho, A., & Stumme, G. (2002). Towards Semantic Web Mining. In *Proceedings of First International Semantic Web Conference (ISWC)* (pp. 264-278).

Bezdek, J. C. (1994). What is computational intelligence? In J. M. Zurada, R. J. Marks II, & C. J. Robinson (Eds.), *Computational Intelligence, Imitating Life*.

Breese, J. S., Heckerman, D., & Kadie, C. (1998). Empirical analysis of predictive algorithms for collaborative filtering. In *Proc. of the 14th Annual Conf. on Uncertainty in Artificial Intelligence* (pp. 43-52).

Bridge, D., Goker, M., McGinty, L., & Smyth, B. (2006). Case-based recommender systems. *The Knowledge Engineering Review*, *20*(3), 315–320. doi:10.1017/S0269888906000567

Burke, R. (2002). Hybrid recommender systems: Survey and Experiments. *User Modeling and User-Adapted Interaction*, *2*(4), 331–370. doi:10.1023/A:1021240730564

Burke, R. (2007). Hybrid web recommender systems. In The Adaptive Web (pp. 377–408). doi:10.1007/978-3-540-72079-9_12

Burke, R., Hammond, K., & Young, B. (1996). Knowledge-based navigation of complex information spaces. In *Proceedings of the 13th National Conference on Artificial Intelligence* (pp. 462-468).

Caglayan, A., Snorrason, M., Jacoby, J., Mazzu, J., Jones, R., & Kumar, K. (1997). Learn Sesame a learning agent engine. *Applied Artificial Intelligence*, *11*(5), 393–412. doi:10.1080/088395197118109

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Cavus, N., & Zabadi, T. (2014). A Comparison of Open Source Learning Management Systems. *Procedia social and behavioral sciences*, *143*, 521-526.

Chen, Y., Wu, C., Xie, M., & Guo, X. (2011). Solving the Sparsity Problem in Recommender System Using Association Retrieval. *Journal of Computers*, *6*(9). doi:10.4304/jcp.6.9.1896-1902

Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computer Education*, *48*(3), 409–426. doi:10.1016/j.compedu.2005.02.004

Cho, Y. H., Kim, J. K., & Kim, S. H. (2002). A personalized recommender system based on web usage mining and decision tree induction. *Expert Systems with Applications*, *23*(3), 329–342. doi:10.1016/S0957-4174(02)00052-0

Claypool, M., Gokhale, A., & Miranda, T. (1999). Combining content based and collaborative filters in an online newspaper. In *Proceedings of the ACM SIGIR '99 Workshop on Recommender Systems: Algorithms and Evaluation*.

Clerkin, P., Cunningham, P., & Hayes, C. (2001). Ontology discovery for the semantic web using hierarchical clustering. In *Semantic Web Mining Workshop*.

Cooley, R., Mobasher, B., & Srivastava, J. (1999). Data Preparation for Mining World Wide Web Browsing Patterns. *Knowledge and Information Systems*, *1*(1), 5–32. doi:10.1007/BF03325089

Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, *13*(1), 21–27. doi:10.1109/TIT.1967.1053964

Degani, R., & Bortolan, G. (1998). The Problem of Linguistic Approximation in Clinical Decision Making. *International Journal of Approximate Reasoning*, *2*(2), 143–162. doi:10.1016/0888-613X(88)90105-3

Degemmis, M., Lops, P., & Semeraro, G. (2007). A content-collaborative recommender that exploits wordnet-based user profiles for neighborhood formation. *User Modeling and User-Adapted Interaction*, *17*(3), 217–255. doi:10.100711257-006-9023-4

Devi, M. K. K., Samy, R. T., Kumar, S. V., & Venkatesh, P. (2010). Probabilistic neural network approach to alleviate sparsity and cold start problems in collaborative recommender systems. In *Proceedings of the IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*. 10.1109/ICCIC.2010.5705777

Eirinaki, M., Vazirgiannis, M., & Varlamis, I. (2003). SEWeP: Using Site Semantics and a Taxonomy to Enhance the Web Personalization Process. In *Proc. of the 9th SIGKDD Conference*. 10.1145/956750.956765

Elorriaga, J. A., & Fernandez-Castro, I. (2000). Using case-based reasoning in instructional planning: Towards a hybrid self-improving instructional planner. *International Journal of Artificial Intelligence in Education*, *11*, 416–449.

Ericson, K., & Pallickara, S. (2013). On the performance of high dimensional data clustering and classification algorithms. *Future Generation Computer Systems*, *29*(4), 1024–1034. doi:10.1016/j.future.2012.05.026

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Facca, F. M., & Lanz, P. i. (2005). Mining interesting knowledge from web logs: A survey. *Journal of Data and Knowledge Engineering*, *53*(3), 225–241. doi:10.1016/j.datak.2004.08.001

Faoud, K.M., Khalifa, A.R., Nagdy, N.M., & Harb, H.M. (2012). Web based semantic and Personalized information retrieval. *International Journal of Computer Science Issues*, *9*(3), 266-276.

Farzan, R., & Brusilovsky, P. (2006). Social navigation support in a course recommendation system. In V. Wade, H. Ashman & B. Smyth (Eds.), Adaptive Hypermedia and Adaptive Web-Based Systems (pp. 91-100). doi:10.1007/11768012_11

Felfering, A., & Burke, R. (2008). Constraint based recommender system: technologies and research issues. In *Proceedings of 10th international conference on electronic commerce*.

Felfernig, A., & Shchekotykhin, K. (2006). Debugging user interface descriptions of knowledge-based recommender applications. In *Proceedings of the 11th International Conference on Intelligent User Interfaces* (pp. 234-241). 10.1145/1111449.1111499

Fetaji, B., Fetaji, M., & Ebibi, M. (2011). E-Learning Systems Comparative Analysis. In *Proceedings of the 5th International Conference on Informatics and Information Technology* (pp. 55-69).

Frank, E., Hall, M., & Pfahringer, B. (2003). Locally weighted Naïve Bayes. *In 19th conference in Uncertainty in Artificial Intelligence* (pp. 249-256).

Garcia, E., Romero, C., Ventura, S., & Castro, C.D. (2009). An architecture for making recommendations to courseware authors using association rule mining and collaborative filtering. *User modelling and user adapted interaction: The journal of Personalization Research*, *19*(1), 99-132.

Gehringer, E. F. (2001). Electronic peer-review and peer grading in computer-science courses. In *Proceedings of the 32nd SIGCSE technical symposium on computer science education* (pp. 139-143). 10.1145/364447.364564

Gemmis, D.M., Iaquinta, L., Lops, P., Muto, C., Narducci, F., & Semeraro, G.(2009).Preferences learning in Recommender Systems. *Preferences Learning*,41.

Georgiou, O., & Tsapatsoulis, N. (2010). Improving the Scalability of Recommender Systems by Clustering Using Genetic Algorithms. In *Proc. of the 20th International conference on Artificial Neural Networks* (pp. 442–449). 10.1007/978-3-642-15819-3_60

Ghahraman, Z., & Jordan, M. I. (1994). Supervised learning from incomplete data via an EM approach. In*Advances in Neural Information Processing Systems* (pp. 120–127).

Goebel, M., & Gruenwald, L. (1999). A survey of data mining and knowledge discovery software tools. *ACM SIGKDD*, *1*(1), 20–33. doi:10.1145/846170.846172

Golder, A., & Huberman, A. (2005). The structure of collaborative tagging systems. *HPL Technical Report*.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Good, N., Schafer, J. B., Konstan, J. A., Borchers, A., & Sarwar, B., Herlocker, & J., Riedl, J. (1999). Combining collaborative filtering with personal agents for better recommendations. In *AAAI '99/IAAI '99: Proc. of the 16th National Conf. on Artificial Intelligence* (pp. 439–446). American Association for Artificial Intelligence.

Hand, D., Mannila, H., & Smyth, P. (2001). *Principles of data mining*. Cambridge: The MIT Press.

Hariri, N., Castro-Herrera, C., Mirakhorli, M., Cleland-Huang, J., & Mobasher, B. (2013). Supporting domain analysis through mining and recommending features from online product listings. *IEEE Transactions on Software Engineering*, *39*(12), 1736–1752. doi:10.1109/TSE.2013.39

Heraud, J., France, L., & Mille, A. (2004) Pixed: an ITS that guides students with the help of learners' interaction log. In *Proceedings of International conference on intelligent tutoring system* (pp. 57–64).

Herlocker, J. L., & Konstan, J. A., Borchers., & Riedl, J. *(1999).* An Algorithmic Framework for Performing Collaborative Filtering. In *Proc on 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 230-237). 10.1145/312624.312682

Hernandez, F., & Gaudioso, E. (2008). Evaluation of recommender systems: A new approach. *Expert Systems with Applications*, *35*(3), 790–804. doi:10.1016/j.eswa.2007.07.047

Herrera, F., Herrera-Viedma, E., & Martınez, L. (2000). A Fusion Approach for Managing Multi-Granularity Linguistic Term Sets in Decision Making. *Fuzzy Sets and Systems*, *114*(1), 43–58. doi:10.1016/S0165-0114(98)00093-1

Herrera, F., Herrera-Viedma, E., & Verdegay, J. L. (1996). Direct approach processes in group decision making using linguistic OWA operators. *Fuzzy Sets and Systems*, *79*(2), 175–190. doi:10.1016/0165-0114(95)00162-X

Herrera, F., & Martınez, L. (2000). A 2-tuple fuzzy linguistic representation model for computing with words. *IEEE Transactions on Fuzzy Systems*, *8*(6), 746–752. doi:10.1109/91.890332

Horrocks, I. (2002). DAML+OIL: A description logic for the semantic web. *A Quarterly Bulletin of the Computer Society of the IEEE Technical Committee on Data Engineering*, *25*(1), 4–9.

Horrocks, I., & Sattler, U. (2001). Ontology reasoning in the SHOQ (D) description logic. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence* (pp. 199-204).

Hsu, M. (2006). A personalized English learning recommender system for ESL student. *Expert Systems with Applications*, *34*(1), 683–688. doi:10.1016/j.eswa.2006.10.004

Hsu, S. H., Wen, M. H., Lin, H. C., Lee, C. C., & Lee, C. H. (2007). AIMED- A Personalized TV Recommendation System. In P. Cesar, K. Chorianopoulos, & J. F. Jensen (Eds.), *Interactive TV: a Shared Experience. EuroITV 2007*, LNCS (Vol. 4471). Berlin, Heidelberg: Springer. doi:10.1007/978-3-540-72559-6_18

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Huang, H. Z., Bo, R., & Chen, W. (2006). An integrated computational intelligence approach to product concept generation and evaluation. *Mechanism and Machine Theory*, *41*(5), 567–583. doi:10.1016/j.mechmachtheory.2005.07.006

Huang, Y.-M., Huang, T.-C., Wang, K.-T., & Hwang, W.-Y. (2009). A Markov-based Recommendation Model for Exploring the Transfer of Learning on the Web. *Journal of Educational Technology & Society*, *12*(2), 144–162.

Hwang, J., Hsiao, L., & Tseng, R. (2003). A computer-assisted approach to diagnosing student learning problems in science courses. *Journal of Information Science and Engineering*, *19*, 229–248.

Iaquinta, L., & Semeraro, G. (2011). Lightweight Approach to the Cold-Start Problem in the Video Lecture Recommendation. In *Proc of ECML-PKDD Discovery Challenge Workshop* (pp. 95-101).

Jain, S., Jain, D.K., Bhojak, H., Bhilwar, A., & Mamatha, J. (2012). Personalization of e-learning services using web mining and semantic web. *International journal of machine learning and computing*, *2*(5).

John, G. C., & Leonard, E. T. (1995).An Instance Based Learner Using an Entropy Distance Measure. In *Proceedings of the 12th International Conference on Machine Learning* (pp. 108-114).

Khirbi, M.K., Jemni, M., & Nasraoui, O. (2008). Automatic recommendations for E-learning Personalization Based on Web Usage Mining Techniques and Information Retrieval. In *Proceedings Of eight international conference on advanced learning technologies*,30-42.

Kim, H. N., Alkhaldi, A., Saddik, A. E., & Jo, G. S. (2011). Collaborative user modeling with user generated tags for social recommender Systems. *Expert Systems with Applications*, *38*(7), 8488–8496. doi:10.1016/j.eswa.2011.01.048

Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., & Riedl, J. (1997). Grouplens: Applying collaborative filtering to usenet news. *Communications of the ACM*, *40*(3), 87. doi:10.1145/245108.245126

Koren, Y., & Bell, R. *(2011)*. Advances in Collaborative Filtering. In Recommender Systems Handbook (pp. 145–186).

Koren, Y., Bell, R. M., & Volinsky, C. (2009). Matrix factorization techniques for recommender systems. *IEEE Computer*, *42*(8), 30–37. doi:10.1109/MC.2009.263

Kotsiantis, S. B., Pierrakeas, C. J., & Pintelas, P. E. (2003). Preventing Student Dropout in Distance Learning Using Machine Learning Techniques. In Knowledge-Based Intelligent Information and Engineering Systems KES '03, LNCS (Vol. 2774). Springer, Berlin, Heidelberg. doi:10.1007/978-3-540-45226-3_37

Kumar, A. (2005). Rule-based adaptive problem generation in programming tutors and its evaluation. In *Proceedings of the 12th International conference on artificial intelligence in education* (pp. 36-44).

Kurilovas, E., & Serikoviene, S. (2013). New MCEQLS TFN method for evaluating quality and reusability of learning objects. *Technological and Economic Development of Economy*, *19*(4), 706–723. doi:10.3846/20294913.2013.837112

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Li, Q., Wang, C., & Geng, G. (2008). Improving Personalized Services in Mobile Commerce by a Novel Multicriteria Rating Approach. In *Proc. of the 17th International World Wide Web Conference*. 10.1145/1367497.1367743

Lin, W. (2000). *Association rule mining for collaborative recommender systems*. Worcester Polytechnic Institute.

Liu, B. (2011). *Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data* (2nd ed.). Berlin: Springer. doi:10.1007/978-3-642-19460-3

Lu, J. (2004). A personalized e-learning material recommender system. In *Proceedings of the 2nd International Conference on Information Technology for Application* (pp. 374-379).

Lu, J., Wu, D., Mao, M., Wang, W., & Zhang, G. (2015). Recommender system application developments: A survey. *Decision Support Systems*, *74*, 12–32. doi:10.1016/j.dss.2015.03.008

Lucas, J. P., Segrera, S., & Moreno, M. N. (2012). Making use of associative classifiers in order to alleviate typical drawbacks in recommender systems. *Expert Systems with Applications*, *39*(1), 1273–1283. doi:10.1016/j.eswa.2011.07.136

Manouselis, N., & Costopoulou, C. (2007). Experimental Analysis of Design Choices in Multi-Attribute Utility Collaborative Filtering. *International Journal of Pattern Recognition and Artificial Intelligence*, *21*(2), 311–332. doi:10.1142/S021800140700548X

Mathes, A. (2004). Folksonomies-cooperative classification and communication through shared metadata. *Comput Mediat Commun*, *47*(10).

Michael, P., & Barry, S. (2007). A Recommender System for On-line Course Enrolment: An Initial Study. In *Proceedings of the ACM Conference on Recommender System* (pp. 133-136).

Mobasher, B., Dai, H., Luo, T., & Nakagawa, M. (2002). Discovery and evaluation of aggregate usage Profiles for web personalization. *Data Mining and Knowledge Discovery*, *6*(1), 61–82. doi:10.1023/A:1013232803866

Mustapasa, O., Karahoca, D., Karahoca, A., Yucel, A., & Uzunboylu, H. (2010). *Implementation of semantic web mining on E-learning. Procedia social and behavioral science, 2(2), 5820-5823*.

Nguyen, T., Lu, H., & Lu, J. (2013). Web-page recommendation based on web usage and domain knowledge. *IEEE Transactions on Knowledge and Data Engineering*, 26(10), 1041–4347.

Owoc, M., & Weichbroth, P. (2012). Validation Model for Discovered Web User Navigation Patterns. In *International workshop on Artificial intelligence for knowledge management* (pp. 38-52).

Pazzani, M. (1999). A framework for collaborative, content-based and demographic filtering. *Artificial Intelligence Review*, *13*(5/6), 393–408. doi:10.1023/A:1006544522159

Pazzani, M., & Billsus, D. (1997). Learning and Revising User Profiles: The Identification of Interesting Web Sites. *Machine Learning*, *27*(3), 313–331. doi:10.1023/A:1007369909943

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Pinkwart, N., Aleven, V., Ashley, K., & Lynch, C. (2006) Toward legal argument instruction with graph grammars and collaborative filtering techniques. In *Proceedings of the 8th international conference on intelligent tutoring systems, LNCS* (Vol. 4053, pp. 227-236). 10.1007/11774303_23

Popescul, A., Ungar, L. H., Pennock, D. M., & Lawrence, S. (2001). Probabilistic models for unified collaborative and content - based recommendation in sparse-data environments. In *Proc of the 17th conference on uncertainty in artificial intelligence* (pp. 437-444).

Porcel, C., & Herrera-Viedma, E. (2010). Dealing with incomplete information in a fuzzy linguistic recommender system to disseminate information in university digital libraries. *Knowledge-Based Systems*, *23*(1), 32–39. doi:10.1016/j.knosys.2009.07.007

Porcel, C., López-Herrera, A. G., & Herrera-Viedma, E. (2009). A recommender system for research resources based on fuzzy linguistic modeling. *Expert Systems with Applications*, *36*(3), 5173–5183. doi:10.1016/j.eswa.2008.06.038

Portugal, I., Alencar, P., & Cowan, D. (2015). The use of machine learning algorithms in recommender systems: A systematic review. In *Proceedings of International conference on Hybrid Artificial Intelligence Systems* (pp. 600-612).

Powell, M. J. D. (1981). *Approximation Theory and Methods*. Cambridge University Press.

Rafter, R. (2010). Evaluation and Conversation in Collaborative Filtering [PhD Thesis]. University College Dublin, College of Engineering Mathematical and Physical Sciences.

Rani, M., Nayak, R., & Vyas, O. P. (2015). An ontology-based adaptive personalized e-learning system, assisted by software agents on cloud storage. *Knowledge-Based Systems*, *90*, 33–48. doi:10.1016/j.knosys.2015.10.002

Recker, M., & Walker, A. (2003). Supporting 'word-of-mouth' social networks via collaborative information filtering. *Journal of Interactive Learning Research*, *14*, 79–98.

Renda, M. E., & Straccia, U. (2005). A personalized collaborative digital library environment: A model and an application. *Information Processing & Management*, *41*(1), 5–21. doi:10.1016/j.ipm.2004.04.007

Resnick, P., Iakovou, N., Sushak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: An open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 Computer Supported Cooperative Work Conference*. 10.1145/192844.192905

Resnick, P., & Varian, H. R. (1997). Recommender systems. *Communications of the ACM*, *40*(3), 56–58. doi:10.1145/245108.245121

Rich, E. (1979). User Modeling via Stereotypes. *Cognitive Science*, *3*(4), 329–354. doi:10.120715516709cog0304_3

Romero, C., Ventura, S., Zafra, A., & Bra, P. D. (2009). Applying web usage mining for personalizing hyperlinks in web-based adaptive educational systems. *Computers & Education*, *53*(3), 828–840. doi:10.1016/j.compedu.2009.05.003

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ruiz-Iniesta, A., Jimenez-Diaz, G., & Gomez-Albarran, M. (2012). A framework for the rapid prototyping of knowledge based recommender systems in the learning domain. *Journal of Research and Practice in Information Technology*, *44*(2), 167.

Sahoo, N., & Krishnan, R., Duncan, & G., Callan, J.P. (2006). Collaborative Filtering with Multicomponent Rating for Recommender Systems. In *Proc. of the 16th Workshop on Informatio Technologies and Systems*.

Salton, G. (1989). *Automatic Text Processing*. Addison-Wesley.

Sarwar, B. M., Konstan, J. A., Borchers, A., Herlocker, J., Miller, B., & Riedl, J. (1998) Using filtering agents to improve prediction quality in the GroupLens research collaborative filtering system. In *Proceedings of the 1998 ACM conference on computer supported cooperative work* (pp. 345–354). 10.1145/289444.289509

Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys*, *34*(1), 1–47. doi:10.1145/505282.505283

Shah, K., & Gadge, J. (2011). Semantic web services for E-learning: Engineering and Technology domain. *International journal of computer theory and engineering, 3*(6).

Shishehchi, S., Banihashem, S.Y., & Zin, N.A.M. (2010). A proposed semantic recommendation system for e- learning: A rule and ontology based e-learning recommendation system. In *Proceedings of the international symposium in Information Technology*.

Silva, N. B., Tsang, I. R., Cavalcanti, G. D. C., & Tsang, I. J. (2010). A graph based friend recommendation system using Genetic Algorithm. In *Evolutionary computation*. CEC. doi:10.1109/CEC.2010.5586144

Soboroff, I., & Nicholas, C., Nicholas, & C. K. (1999). Combining Content and Collaboration in Text Filtering. In *Proceedings of the IJCAI'99 Workshop on Machine Learning for Information Filtering* (pp. 86–91).

Srivastava, J., Cooley, R., Deshpande, M., & Tan, P. N. (2000). Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data. *SIGKDD Explorations*, *1*(2). doi:10.1145/846183.846188

Stumme, G., Hotho, A., & Berendt, B. (2002). Usage Mining for and on the Semantic Web. In *National Science Foundation Workshop on Next Generation Data Mining*.

Stumme, G., Hotho, A., & Berendt, B. (2006). Semantic web mining: State of the art and future directions. *Journal of Web Semantics*, *4*(2), 124–143. doi:10.1016/j.websem.2006.02.001

Sungjune, P. S., Suresh, N. C., & Jeong, B. K. (2008). Sequence - based clustering for Web usage mining: A new Experimental framework and ANN-enhanced K-means algorithm. *Data & Knowledge Engineering*, *65*(3), 512–543. doi:10.1016/j.datak.2008.01.002

Takacs, G., Pilaszy, I., Nemeth, B., & Tikk, D. (2008). Investigation of various matrix factorization. methods for large recommender systems. In *Proc. of the 2nd KDD Workshop on Large Scale Recommender Systems and the Netflix Prize Competition*. 10.1145/1722149.1722155

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Third, A. (2012). Hidden semantics: what can we learn from the names in an ontology? In *Proc. of INLG* (pp. 67–75).

Tiemann, M., & Pauws, S. (2007). Towards ensemble learning for hybrid music recommendation. In *RecSys '07: Proceedings of the 2007 ACM conference on Recommender systems* (pp. 177–178). 10.1145/1297231.1297265

Trewin, S. (2000). Knowledge Based Recommender Systems. Encyclopedia of Library and Information Science, 69(32).

Tsai, K. H., Chiu, T. K., Lee, M. C., & Wang, T. I. (2006). A learning objects recommendation model based on the preference and ontological approaches. In *Proc. of the 6th International Conference on Advanced Learning Technologies (ICALT'06)* (pp. 36–40). 10.1109/ICALT.2006.1652359

Unadkat, R. (2015). Survey Paper on Semantic Web. *International Journal of Advanced Pervasive and Ubiquitous Computing*, *7*(4), 13–17. doi:10.4018/IJAPUC.2015100102

Vargas, M., & Lytras, M. (2008). Personalized learning using ontologies and semantic web ontologies. In *World summit on knowledge society* (pp. 177-186).

Walker, A., Recker, M., Lawless, K., & Wiley, D. (2004). Collaborative information filtering: A review and an educational application. *International Journal of Artificial Intelligence in Education*, *14*(1).

Yager, R. (1988). On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Transactions on Systems, Man, and Cybernetics*, *18*(1), 183–190. doi:10.1109/21.87068

Yager, R. R. (2003). Fuzzy logic methods in recommender systems. *Fuzzy Sets and Systems*, *136*(2), 133–149. doi:10.1016/S0165-0114(02)00223-3

Yang, Y., & Wu, C. (2009). Attribute-based ant colony systems for adaptive learning object recommendation. *Expert Systems with Applications*, *36*(2), 3034–3047. doi:10.1016/j.eswa.2008.01.066

Yeung, A. A. (2010). Matrix Factorization: A Simple Tutorial and Implementation in Python [Web log post]. Retrieved January 5, 2018, from http://www.quuxlabs.com/blog/2010/09/matrix-factorization-a-simple-tutorial-and-implementation-in-python/

Yuen, K. K. F. (2012). A multiple criteria decision making approach for E-learning platform selection: the Primitive Cognitive Network Process. In *Computing, Communications and Applications Conference (ComComAp)* (pp. 294–298). 10.1109/ComComAp.2012.6154860

Zadeh, L. A. (1975). The concept of a linguistic variable and its applications to approximate reasoning. *Information Sciences*, *9*, 43–80. doi:10.1016/0020-0255(75)90017-1

Zaiane, R. (2002). Building a recommender agent for e-Learning Systems. In *Proceedings of the International conference on Computers in Education* (pp. 55-59). 10.1109/CIE.2002.1185862

Zanker, M., Fuchs, M., Hopken, W., Tuta, M., & Muller, N. (2008). Evaluating Recommender Systems in Tourism - A Case Study from Austria. In *Proceedings of the International Conference on Information and Communication Technologies in Tourism* (pp. 24–34). 10.1007/978-3-211-77280-5_3

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Zenebe, A., & Norcio, A. F. (2009). Representation, similarity measures and aggregation methods using fuzzy sets for content-based recommender systems. *Fuzzy Sets and Systems*, *160*(1), 76–94. doi:10.1016/j.fss.2008.03.017

Zhang, D., Zhao, J. L., Zhou, L., & Nunamaker, J. (2004). Can e-learning replace traditional classroom learning—evidence and implication of the evolving e-learning technology. *Communications of the ACM*, *47*(5), 75–79. doi:10.1145/986213.986216

Zhang, Z., Lin, H., Liu, K., Wu, D., Zhang, G., & Lu, J. (2013). A Hybrid Fuzzy-Based Personalized Recommender System for Telecom Products/Services. *Information Sciences*, *235*, 117–129. doi:10.1016/j.ins.2013.01.025

Zhang, Z., Liu, C., Zhang, Y.-C., & Zhou, T. (2010). Solving the cold start problem in recommender systems with social tags. *EPL*, *92*(2), 28002. doi:10.1209/0295-5075/92/28002

Zhudhdar, L., & Nasraoui, O. (2009). Automatic discovery, categorization and retrieval of personalized semantically enriched e-learning resources. In *Proceedings of the International conference on semantic computing* (pp. 414-419).

This research was previously published in the International Journal of Advanced Pervasive and Ubiquitous Computing (IJAPUC), 10(1); edited by Tao Gao; pages 51-76; copyright year 2018 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 76
Combining Artificial Intelligence and NetMedicine for Ambient Assisted Living: A Distributed BDI-based Expert System

Paolo Sernani
Università Politecnica delle Marche, Italy

Andrea Claudi
Università Politecnica delle Marche, Italy

Aldo Franco Dragoni
Università Politecnica delle Marche, Italy

ABSTRACT

World population is shifting towards older ages: according to recent estimates there will be 1.5 billion people over 65 years old in 2050. Local governments, international institutions, care organizations and industry are fostering the research community to find solutions to face the unprecedented challenges raised by population ageing. A combination of Artificial Intelligence and NetMedicine could be ideal to face these challenges: they provide the means to develop an intelligent system and simultaneously to distribute it over a network, allowing the communication over the internet, if needed. Hence, the authors present a Multi-Agent Architecture for Ambient Assisted Living (AAL): it is the model for a system to manage a distributed sensor network composed by ambient and biometric sensors. The system should analyse data and pro-actively decide to trigger alarms if anomalies are detected. The authors tested the architecture implementing a prototypical Multi-Agent System (MAS), based on Belief-Desire-Intention (BDI) paradigm: the Virtual Carer.

DOI: 10.4018/978-1-7998-0951-7.ch076

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

INTRODUCTION

The lengthening of life expectancy, jointly to the decreasing of fertility, will triple the number of people aged 65 or older from 524 million in 2010 (8 percent of the world population) to 1.5 billion in 2050 (16 percent of the world population) (World Health Organization, 2011). The increasing median age of the world population has significant social and economic implications: the consequent rise of chronic diseases will result in a dramatic increase of emergency situations and missions within the coming years (Kleinberger, Becker, Ras, Holzinger, & Müller, 2007). This is becoming a substantial issue in Europe, USA and Japan owing to their high dependency ratio[1] (Christensen, Doblhammer, Rau, & Vaupel, 2009). Thus national governments, no-profit organizations as well as industry are fostering the research community to contribute to the Ambient Assisted Living (AAL) field. AAL aims at extending the time older people can live in their home environment, assisting them with the activities of daily living, promoting the use of intelligent products and Information Technology (IT) tools to provide remote care services (Sun, De Florio, Gui, & Blondia, 2009).

In this paper we propose a multi-agent architecture for an AAL system. The cornerstone of this idea is the combination of Artificial Intelligence and NetMedicine in order to develop a system able to manage a distributed sensor network, composed by ambient and biometric sensors; it should be, at the same time, an interface layer between the network, the assisted person, his relatives and the medical staff. The system, named Virtual Carer, models actions and behaviours of a human caregiver, to help a patient in his daily activities inside his home environment, simultaneously monitoring his health conditions and ensuring his security. The Virtual Carer is able to analyse data coming from the available sensors and, thanks to its knowledge base, to infer new knowledge and to decide if detected values should be notified, possibly triggering an alarm.

The need for an intelligent system controlling a distributed network clearly suggests an integration between Artificial Intelligence and NetMedicine. Artificial Intelligence provides technologies to develop a proactive expert system, while NetMedicine offers the possibility to access data and alarms triggered by the system through the web. The term NetMedicine simply denotes every medicine activity which is carried on through the Internet (Dragoni, 2009). In other words NetMedicine has to be intended as the evolution of TeleMedicine, due to the digitalization of information and the emergence of internet: the fast improvement of the ICT technologies is greatly moving the demand for knowledge by the physicians towards the cyberspace (Nakic, & Loškovska, 2012). Moreover, social networking fosters the constitution of communities sharing similar medical interests, even enhancing primary care practice (Hawn, 2009).

Multi-Agent Systems (MAS) represent one of the possible points of contact between Artificial Intelligence and NetMedicine; indeed the main advantages of a MAS approach (Palazzo, Rossi, Dragoni, Claudi, Dolcini, & Sernani, 2013), essential for the proposed system, are:

- Distribution;
- Modularity;
- Robustness.

The MAS architecture allows to distribute software agents and the platforms that manage them over one or more networks; each sensor can be associated with an agent, making the system resilient to sensor failures and disconnections. The adoption of the Foundation for Intelligent Physical Agents (FIPA) standards as the Agent Communication Language (ACL) allows to add and remove sensors and devices

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

in a transparent way, simply wrapping them with agents. Furthermore agent-based systems can provide interfaces to access through the internet.

In addition to the listed features, pro-activeness plays a prominent role, making the system able to autonomously take the initiative, guided by goals. Thus the reasoning module of the system, i.e. the Virtual Carer Agent (VCA), is based on Belief-Desire-Intention (BDI) paradigm. The VCA models logical structures, reasoning and behaviours similar to those of a human caregiver, transforming data from sensors in logical predicates representing its knowledge base. To perform actions on the monitored environment and interact with the assisted person, VCA communicates (through FIPA messages) with agents controlling actuators.

The listed features meet the requirements on modularity and adaptability of AAL systems (Ayala, Amor, & Fuentes, 2012): MAS approach and BDI paradigm make the proposed system capable to autonomously adapt to context changes.

The rest of the paper is organized as follows. The next section gives the needed background on agent technology, BDI paradigm and AAL in general. Then some related works are outlined. In the further section a MAS general architecture for AAL is described. Implementation section gives details about the proposed system, highlighting some features of the Virtual Carer and describing its main advantages by means of a simple simulation scenario. Finally, results are highlighted and conclusions drawn.

BACKGROUND

The guidelines of AAL indicates the need of unobtrusive, adaptable and dynamic intelligent environments (Sadri, 2011); AAL systems have to adapt themselves over time to cope with the changing needs and health conditions of the assisted person, providing technical solutions flexible and adaptable enough to respond to this dynamic requirement (Wolf et al., 2010). Moreover, they need to perceive variation in habits that can indicate health-related problems or stressful situations. Applying Artificial Intelligence techniques on this domain seems to be a promising direction for research (Pollack, 2005; Cook, Augusto, & Jakkula, 2009; O'Grady, Muldoon, Dragone, Tynan, & O'Hare, 2010).

The next two subsections describe agent technology and the AAL field, clarifying that MAS and BDI agents with an appropriate knowledge base are suitable to meet the listed guidelines.

Agents

According to Russell & Norvig (2009), an agent is an entity that can perceive the environment through sensors and act on the environment through actuators. Wooldridge (1995) defines the main properties of an agent: autonomy, social ability, reactivity and pro-activeness indicate that agents operate without external intervention, interact with other agents using an Agent Communication Language (ACL), perceive the environment answering to changes and are able to autonomously exhibit goal-directed behaviours.

The human practical reasoning described by Bratman (1987) is the theoretical foundation for the BDI paradigm: the intersection between beliefs, i.e. the agent knowledge about the environment, and desires, i.e. the set of states representing the environment as desired by the agent, generates agent's intentions, i.e. the set of desires that an agent wants to achieve; in other words intentions constitute the goals pursued by an agent.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

An agent-based system usually consists of multiple agents (Bellifemine, Caire, & Greenwood, 2007): these MAS are suitable to model complex systems; the agents composing a MAS may interact with each other both indirectly (by acting on the environment) or directly (via communication and negotiation). The interaction between agents may be cooperative (they cooperate to achieve mutual benefits) or competitive (they vie for their own interests).

Ambient Assisted Living

Through the AAL Joint Programme, the European Union aims to foster the emergence of AAL services and systems for ageing well at home, in the community, and at work (The AAL Joint Programme, 2004). The AAL Joint Programme defines these key objectives for AAL:

- To extend the time people can live in their preferred environment by increasing their autonomy, self-confidence and mobility;
- To support maintaining health and functional capability of the elderly individuals;
- To promote a better and healthier lifestyle for individuals at risk;
- To enhance the security, to prevent social isolation and to support maintaining the multifunctional network around the individual;
- To support carers, families and care organizations;
- To increase the efficiency and productivity of used resources in the ageing societies.

Home Automation and, specifically, Ambient Intelligence (AmI) provide pervasive devices and technologies for AAL systems, in order to build a safe environment for the assisted person (Sun et al., 2009). AmI builds around people an unobtrusive, interconnected, adaptable, dynamic, embedded, and intelligent environment (Sadri, 2011), with the leading goal to help them in their daily activities. One of the main commitments of AmI is to allow the user to interact with an AmI system using the most natural ways, as he would do with any other human (Kleinberger et al., 2007). However assistive technologies can supplement human caregiving and cannot substitute it, having the potential to improve the quality of life for both older adults and their caregivers (Pollack, 2005).

The need of a smart interaction between AAL and AmI systems and the elderly is guiding the research community towards the exploration of Human-Computer Interaction (HCI) and Human-Robot Interaction (HRI) solutions. In such a context elderly people should be considered the most prominent stakeholders for IT developments (Holzinger, Searle, Auinger, & Ziefle, 2011).

RELATED WORKS

In recent years the MAS approach to AAL and healthcare has been exploited in several works. Reinisch & Kastner (2011) point out the importance of pro-activeness in a smart home characterized by a dynamic, non-deterministic, complex environment: being capable of reasoning to find out which actions to execute, BDI agents are better suited to such an environment than reactive agents; in the use case proposed by authors a small society of BDI agents is designed to cope with the energy efficiency of a smart home when different devices are available. To monitor the patient status, BDI agents are implemented also in the work described by Borowczyk, Gawinecki, & Paprzycki (2008): agents cooperate to carry out

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

specific tasks such as evaluating the patient's heartbeat and deciding to trigger an alarm if necessary. Nefti, Manzoor, & Manzoor (2010) propose a MAS performing risk assessment tasks for patients with dementia: a Risk Assessment Agent, using its knowledge base (to load the patient's history) and information provided by agents responsible for various ambient sensors (such as a Gas Detector Agent or a Temperature Analyzer Agent), employs a method based on fuzzy logic to predict risk situations and to send appropriate alarms. Beside agent-based works, scientific literature is rich of contributions on the intelligent interpretation of available data. In (Mileo, Merico, Pinardi, & Bisiani, 2009), to monitor the posture and the health conditions of an elderly, a reasoning component based on answer set programming analyses data gathered by a hierarchy of wireless sensor networks in a home environment. Pinardi, & Bisiani (2010) describe the recognition of the patient's movements using a machine learning approach. A smart home should also include technologies to allow an easy interaction between the elderly and the system: speech recognition is deeply investigated, establishing new challenges for the analysis of ageing voices (Vacher, Portet, Fleury, & Noury, 2011).

To implement an AAL system with an effective impact on daily life the design phase and, at a later stage, the evaluation phase should include several perspectives (Kleinberger, Jedlitschka, Storf, Steinbach-Nordmann, & Prueckner, 2009): the points of view of end users (elderly persons and their relatives), physicians, medical caregivers and research developers (software and systems engineers) are all characterized by the same importance to assess the effectiveness of an AAL system. Participatory design (Serrano, 2010) and evaluation (Demiris, Oliver, Dickey, Skubic, & Rantz, 2010) involve all the stakeholders, taking into account different (and sometimes conflicting) needs and opinions.

ARCHITECTURE

Our multi-agent architecture constitutes the ground for an AAL (IT) system, characterized by the following fundamental purposes:

- To communicate with an assisted person;
- To monitor his health conditions;
- To control the environment around him.

The system is purely an Information Technology (IT) system: it exclusively use technologies provided by Computer Science and Artificial intelligence to model a distributed, reliable and modular sensor network composed by biometric and ambient sensors. To cope with highly variable environments typical of AAL, the prosed agency integrates a BDI core agent: the Virtual Carer Agent (VCA). It represents reasoning mechanisms and behaviours similar to those of a human caregiver. This feature is the core of the prototypical system "Virtual Carer" based on the defined multi-agent architecture.

The large amount of information used by the Virtual Carer system is represented by logical predicates, forming its Knowledge Base (KB). It chooses the right actions to perform on the environment with an inference engine applied on its KB. The Virtual Carer works as follows:

- It analyses data provided by sensors and devices forming the system (i.e. by agents controlling them) and updates its KB;
- When its KB is updated, Virtual Carer generates new knowledge using its inference engine;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

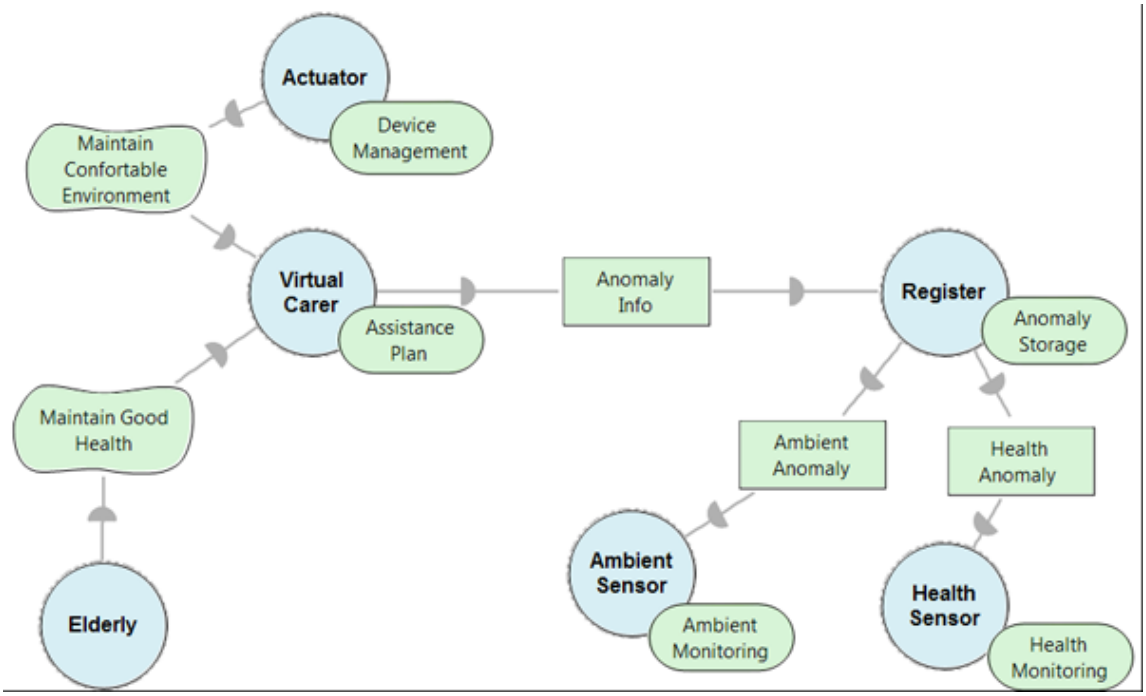
- With backward reasoning rules applied on its KB, the Virtual Carer selects the main goal and chooses (through backtracking) the plan to satisfy it;
- The Virtual Carer carries out the actions of the chosen plan, collaborating with agents responsible for actuators, in order to directly act on the environment.

We modelled the system using the *i** notation (Yu, Giorgini, Maiden, & Mylopoulos 2011). Figure 1 illustrates a high-level Strategic Dependency Model; the system is composed by various actors: Virtual Carer, Register, Actuator, Ambient Sensor, Health Sensor and the elderly person who needs assistance. To preserve his health conditions, the elderly depends on plans executed by the Virtual Carer; also Actuators depend on command and plan decided by the Virtual Carer, in order to maintain the environment comfortable. A Register Actor is responsible for the storage of detected anomalies, offering the databases to the Virtual Carer. Thus, the Register depends on Ambient Sensor and Health Sensor for the acquisition of values.

Agency

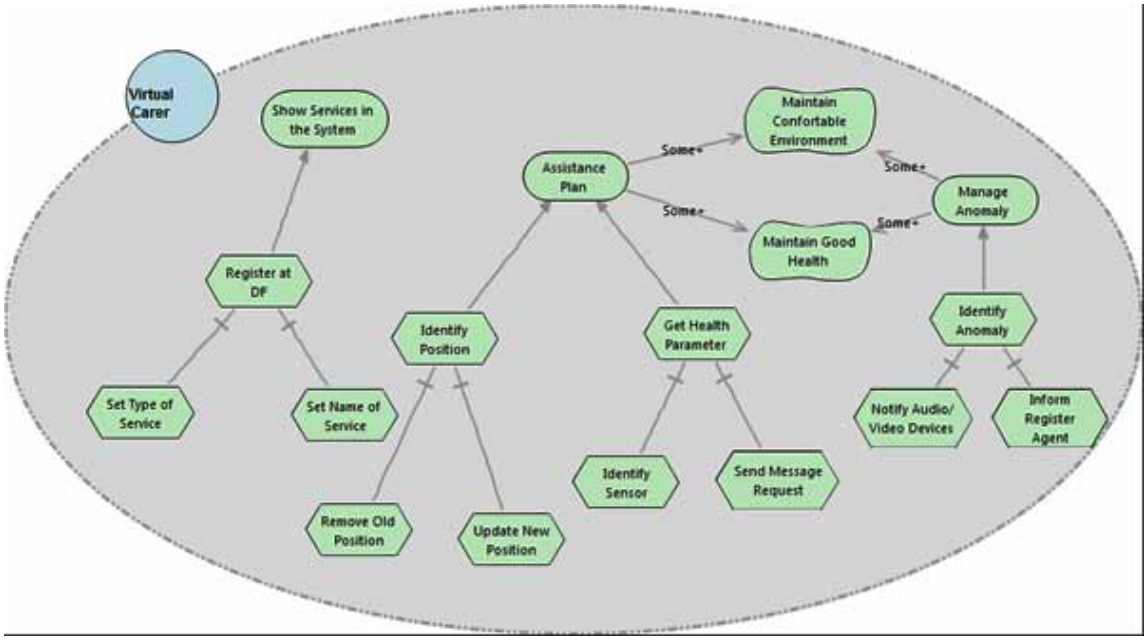
Each actor in the Strategic Dependency Model, except for the Elderly, corresponds to an agent in the defined architecture. Hence, there are a Virtual Carer Agent (VCA), a Register Agent, Ambient and Health Agents and Actuator Agents. The VCA is responsible for the execution of plans on the basis of the context (its belief base), according to the BDI paradigm. Figure 2 shows the Actor Model developed for the VCA. To achieve its goals the VCA performs several tasks: it has to identify the patient's health

Figure 1. i Strategic Dependency Model for the whole system. The Virtual Carer should cooperate with other actors to allow the elderly to "mantain a good health"*



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. i Actor model for the Virtual Carer: to achieve its goals the Virtual Carer uses data gathered by health and ambient sensor*



condition, his position and to detect dangerous or anomalous parameters using available sensors. To complete these tasks the VCA has to publish its services for other agents: this goal can be achieved by registering to the Directory Facilitator[2] (DF) of the whole agency.

Sensor Agents are responsible for environment and health monitoring. Hence, they can be distinguished in two types: Ambient Agents and Health Agents. Ambient Agents read the value of ambient sensors, as the temperature of a specific room. If the value is outside of a predetermined range, the Ambient Agent communicates the anomaly to the VCA. Also presence sensors (as Passive Infrared) are managed the Ambient Agents: they directly inform the VCA about the presence of someone in the monitored room. Similarly, Health Agents manage sensors measuring parameters related to the health conditions of the assisted person. An example could be a Health Agent to control a heartbeat sensor.

Finally, Actuator Agents are responsible for the management of the various devices composing the system such as lights, shutters, speakers, monitors, etc. They must be able to receive request messages (from the VCA) in order to execute on/off commands on the devices.

VIRTUAL CARER IMPLEMENTATION

Two frameworks were adopted to implement the agents composing the Virtual Carer: JADE (Bellifemine et al., 2007) and JASON (Bordini, Hübner, & Wooldridge, 2007): both are FIPA compliant. Firstly the frameworks are fully interoperable; moreover modularity of the system is ensured, since the communication takes advantage of the adoption of FIPA-ACL messages: different devices can be added any time simply integrating more agents. JASON and the AgentSpeak language were used for BDI agents,

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

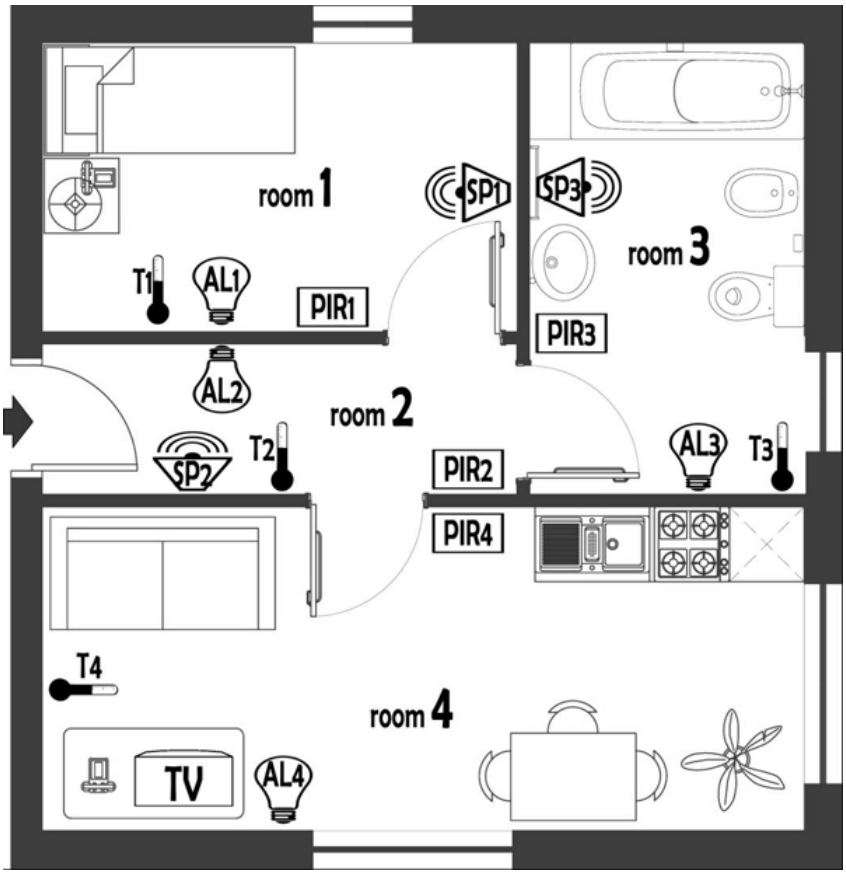
i.e. the VCA and the Register Agent, making it capable to store beliefs directly in the database. JADE, instead, was employed for the implementation of Sensor and Actuator Agents.

To show the effectiveness of the proposed agency a simulation scenario were built. In addition, to highlight the capability to integrate devices in a transparent way, a mobile application for the STMicroelectronics BodyGateWay sensor was developed and integrated into the agency using the JADE-LEAP add-on (Bellifemine et. Al, 2007).

Agency Operating Mode

In the typical interaction among the agents composing the Virtual Carer, the VCA receives alarms from Sensor Agents when a value is out of the predetermined safe range; otherwise, if the current plan needs data to satisfy the goal, the VCA can decide to require to read a value on his own initiative. The Register Agent receives and stores in a database values from Sensor Agents and warnings or anomalies reported by VCA. Similarly, the VCA can require values from the database to the Register Agent. The VCA sends request to the Actuator Agents in order to execute actions corresponding to the adopted plan[3].

Figure 3. Fictional home map for the simulation scenario of the Virtual Carer



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Simulation Scenario

To test the Virtual Carer, software simulations were carried out: figure 3 presents the fictional home map used as a simulation scenario, highlighting the arrangement of the available devices in the environment. There are 4 rooms: the bedroom (room 1), the hallway (room 2), the bathroom (room 3) and the living room (room 4). For each one, a temperature sensor (T1, T2, T3, T4), a PIR sensor (PIR1, PIR2, PIR3, PIR4) and a device to turn the light on/off (AL1, AL2, AL3, AL4) are provided. In addition there is a speaker (SP1, SP2, SP3) for each one of the first three rooms, and a monitor to visually communicate with the assisted person in room 4. Pressure sensors are supposed to be in each place where the person can sit or lie down (e.g. couch, bed, seats), and contact sensors are supposed to be placed on every door and window, to monitor their open/close state. Body sensors are used to monitor the vital parameters (e.g. heartbeat, body temperature) of the assisted person. For each sensor and device in the described scenario the appropriate Ambient, Health and Actuator Agents are created.

To run the simulations a BDI agent, i.e. the Elderly Agent, was integrated in the agency, modelling plausible behaviours of the person assisted by the Virtual Carer system. The Elderly Agent sends messages to other agents of the system: for example, to simulate the detection by a PIR sensor, it informs the Ambient Agent when it moves in the controlled room; moreover it can send requests directly to Health Sensors to know the value of specific health parameters.

Figure 4 shows an emergency plan example for the VCA: it is executed when an Ambient Agent, responsible for a temperature sensor, reads a value out of the predetermined range. In this case the VCA communicates the anomaly to the Register Agent and activates the plan “notify”, in order to evaluate if the monitored person has to be notified with some alarms. The plan in figure 5 is activated when the body temperature of the assisted person is out of a security range. This plan is similar to the previous one.

In addition, the VCA needs to evaluate if a combination of values provided by Sensor Agents is dangerous. Thus the VCA’s KB includes specific rules to perform this task.

Figure 4. Emergency plan for the Virtual Carer Agent. The plan responds to a room temperature anomaly

```
+emergency1(V)[source(A)] : true <-
        .print("Room temperature out of range");
        .print("from sensor ", A);
        .print("value ", V);
        .time(H,M,S);
        .send(register,tell,anomaly(A,V,H,M,S));
        .print("Information sent to Register
Agent");
        ?em1(N);
        New = N + 1;
        -+em1(New);
        !notify;
```


EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 5. Emergency plan for the Virtual Carer Agent. The plan responds to a body temperature anomaly

```
+emergency2(V)[source(A)] : true <-
        .print("Body temperature out of range");
        .print("from sensor ", A);
        .print("value ", V);
        .time(H,M,S);
        .send(register,tell,anomaly(A,V,H,M,S));
        .print("Information sent to Register
Agent");
        ?em2(N);
        New = N + 1;
        -+em2(New);
        !notify;
```

Figure 6. An example of the rules in the VCA's KB. If the heartbeat is over 120 bpm and the patient is not moving, it could be tachycardia

```
tachycardia :- heartbeat(X) & dangerousHeartbeat(X)
& restfulActivity.

dangerousHeartbeat(X) :- X >= 120.

restfulActivity :- accelerazione(Y) & Y < 0.3.
```

Figure 6 contains an example[4] from the VCA's KB. The rules intuitively indicate that a scenario where the patient has an heartbeat over 120 bpm and is not in a stressful situation (i.e. he is not moving) could be considered tachycardia: the VCA has to perform some actions to solve this emergency without stressing the elderly.

To simulate the different behaviours of the assisted person, the Elderly Agent was provided with different plans. Figure 7 shows a plan including the transition from the hallway to the bedroom the plan models the situation in which the assisted person goes to bed and measures its body temperature. This simulation works as follows. At first, the Elderly Agent initializes the VCA, providing it the information about its initial position; then, after 10 seconds (i.e. the time necessary to get to bedroom), it informs the PIR Agent of its presence in the final position (this simulates the detection of the elderly by the PIR sensor). The PIR Agent controlling the bedroom notifies the VCA about the presence of a person in the bedroom; thus the VCA has a new belief (i.e. the presence of the assisted person in room 1), and activates the plan to turn the light on in the bedroom and to turn it off in the hallway. Finally, the Elderly Agent sends a message to the VCA to know its body temperature value. The VCA activates a plan that includes all the actions necessary to require data from the proper Health Agent and to communicate it to the Elderly Agent.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

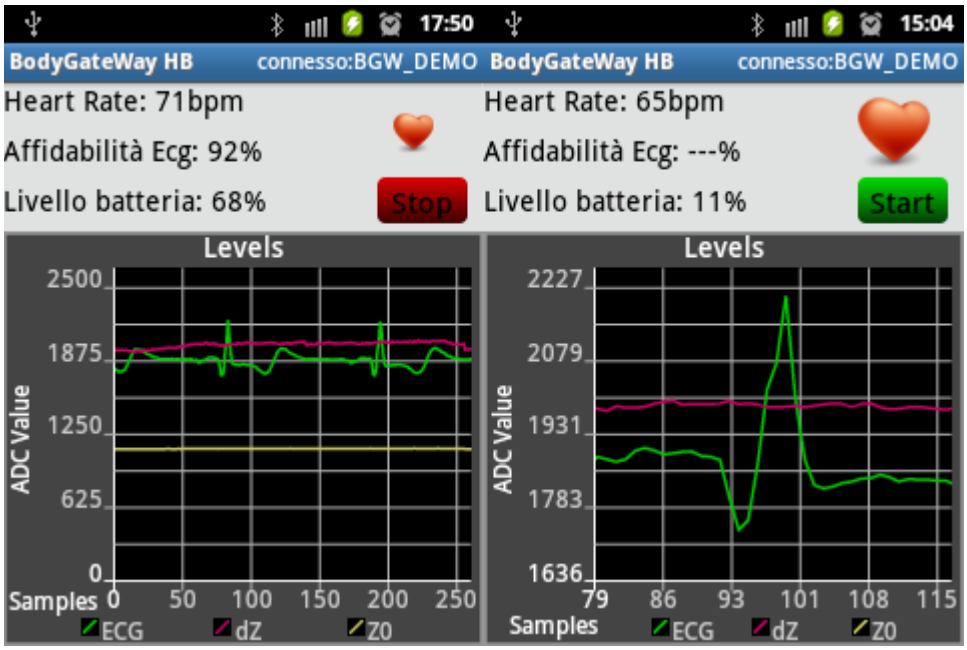
Figure 7. A plan for the Elderly Agent

```
+-scenario5 : true <-
        .send(carer,tell,init(hallway));
        .wait(10000);
        .send(pir1,tell,at);
```

Beside plans similar to the one just described, we simulated plans without an explicit request by the Elderly Agent, in which the Virtual Carer had to infer the right actions. For example, when a window remains open in the bedroom and the assisted person is sleeping, the decreasing of room and body temperature should activate the plan to wake up the person and ask him to close the window. In this scenario, the Ambient Agent associated to the bedroom window informs the VCA that the window state is "open". At a later stage the Ambient Agent responsible for T1 sensor communicates to the VC Agent that the detected value is out of the predetermined range; a similar message is sent by Health Agent. VCA updates its KB with these new beliefs and activates a plan ending with a message to the Elderly Agent, asking it to close the window.

To highlight the power of the MAS approach we developed a mobile application for the STMicroelectronics BodyGateWay, a wearable electronic patch integrating three accelerometers with sensors to detect the ECG and the bioelectrical impedance. Figure 8 shows a screenshot from the Android application. We tested the application separately, on human patients. Then we integrated the mobile application in the agency, using the JADE-LEAP add-on and simulating the message exchanges with the Virtual Carer Agent.

Figure 8. Screenshots from the mobile application for the STMicroelectronics BodyGateWay



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

RESULTS

To fully validate the proposed approach, some tests should be carried out in a real home environment, equipped with existing sensors and devices. However, the implemented simulations are suitable to underline the advantages of the application of MAS and BDI paradigm to AAL.

JADE and JASON (i.e. the frameworks we used for implementation) provide fault tolerant techniques to implement agent platform, such as the Main Container Replication Service (MCRS) and the Directory Facilitator Persistence: the first prevents that the agents creation and management can be single points of failure, the latter allows the traceability of agent services.

More in general MAS approach ensures the distribution of agents over a network; the adoption of FIPA standards guarantees modularity to the system: as tested in the simulations, devices can be added without specific efforts. Furthermore agents, using FIPA-ACL messages for communication, can provide information over the internet, publishing services for the authorized clients.

Thanks to the BDI paradigm, data from ambient and health sensors update the VCA's belief base; thus the VCA can infer new knowledge and desires, being able to decide to trigger an alarm if an anomaly is detected. Hence, the VCA quickly responds to changes in a dynamic environment.

CONCLUSION

We presented a MAS approach to AAL to face a dynamic environment, monitoring the health conditions of an elderly or disable person. The essential idea of the work was to combine Artificial Intelligence and NetMedicine through the definition of a multi-agent architecture and the adoption of BDI paradigm, to lay the foundations for a distributed expert system. A prototype system, i.e. the Virtual Carer, was implemented to highlight the benefits of the proposed architecture. The reasoning core of the system is a BDI agent: the Virtual Carer Agent. It collaborates with other agents controlling various sensors and devices (Actuator Agents) to facilitate the daily activities of the assisted person and to trigger alarms when an anomaly is detected among the parameters regarding the patient's health conditions.

The system flexibility is ensured by the adoption of FIPA standards: new devices can be added simply integrating agents in the proposed architecture, making the system ready for the inclusion of new technologies. The adopted frameworks natively provide services addressed to fault-tolerance: the MCRS and the DF persistence make the agent-based architecture resilient to single point of failures for the agent management and for the traceability of agents' services.

The system was tested through software simulations in a fictional scenario; the assisted person was represented by another BDI agent: the Elderly Agent. Several plans were taken into account, modelling different actions performed by the Elderly Agent (e.g. movements in the home environment, direct request of health parameters). Tests show that MAS and BDI paradigm improve modularity and adaptability of AAL and AmI systems. The MAS approach guarantees the distribution of the system, making it able to share information with authorized clients, if needed; the BDI paradigm allows to respond to changes in the environment and in the needs of the assisted person. Thus MAS, as the point of contact between Artificial Intelligence and NetMedicine, can play a crucial role in the development of systems to extend the time in which people can live in their preferred environment.

Future works include the execution of more insightful tests; nevertheless a daily-used home environment, equipped with devices and sensors, where an assisted person can actually live, or an accurated

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

simulator (Sernani, Claudi, Calvaresi, Accattoli, Tofani, & Dragoni, 2014) is beyond our possibilities at this stage of the work, but it is the only way to fully validate the proposed approach.

ACKNOWLEDGMENT

The authors would like to thank STMicroelectronics for providing the BodyGateWay sensor for tests.

REFERENCES

Ayala, I., Amor, M., & Fuentes, L. (2013). Self-configuring agents for ambient assisted living applications. *Personal and Ubiquitous Computing*, *17*(6), 1159–1169. doi:10.100700779-012-0555-9

Bellifemine, F. L., Caire, G., & Greenwood, D. (2007). *Developing multi-agent systems with jade*. John Wiley & Sons. doi:10.1002/9780470058411

Bordini, R. H., Hübner, J. F., & Wooldridge, M. (2007). *Programming multi-agent systems in agentspeak using jason*. John Wiley & Sons.

Borowczyk, A., Gawinecki, M., & Paprzycki, M. (2008). BDI agents in a patient monitoring scenario. In *Pervasive Computing Technologies for Healthcare, 2008. PervasiveHealth 2008. 2nd international conference on* (pp. 82-85). Tampere, FIN: IEEE.

Bratman, M. (1987). *Intentions, Plans, and Practical Reason*. Cambridge, MA, USA: Harvard University Press.

Christensen, K., Doblhammer, G., Rau, R., & Vaupel, J. W. (2009). Ageing populations: The challenges ahead. *Lancet*, *374*(9696), 1196–1208. doi:10.1016/S0140-6736(09)61460-4 PMID:19801098

Cook, D. J., Augusto, J. C., & Jakkula, V. R. (2009). Ambient intelligence: Technologies, applications, and opportunities. *Pervasive and Mobile Computing*, *5*(4), 277–298. doi:10.1016/j.pmcj.2009.04.001

Demiris, G., Oliver, D. P., Dickey, G., Skubic, M., & Rantz, M. (2008). Findings from a participatory evaluation of a smart home application for older adults. *Technology and Health Care*, *16*(2), 111–118. PMID:18487857

Dragoni, A. F. (2009). Stable communities and social networking for the netmedicine of tomorrow (or today). *Acta Informatica Medica*, *17*(3), 161–164.

Hawn, C. (2009). Take two aspirin and tweet me in the morning: How Twitter, Facebook, and other social media are reshaping health care. *Health Affairs (Project Hope)*, *28*(2), 361–368. doi:10.1377/hlthaff.28.2.361 PMID:19275991

Holzinger, A., Searle, G., Auinger, A., & Ziefle, M. (2011). Informatics as semiotics engineering: lessons learned from design, development and evaluation of ambient assisted living applications for elderly people. In. Lecture Notes in Computer Science: Vol. 6767. *Universal access in human-computer interaction. Context diversity* (pp. 183–192). Springer Berlin Heidelberg. doi:10.1007/978-3-642-21666-4_21

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Kleinberger, T., Becker, M., Ras, E., Holzinger, A., & Müller, P. (2007). Ambient intelligence in assisted living: enable elderly people to handle future interfaces. In. Lecture Notes in Computer Science: Vol. 4555. *Universal access in human-computer interaction. Ambient interaction* (pp. 103–112). Springer-Verlag Berlin Heidelberg. doi:10.1007/978-3-540-73281-5_11

Kleinberger, T., Jedlitschka, A., Storf, H., Steinbach-Nordmann, S., & Prueckner, S. (2009). An approach to and evaluations of assisted living systems using ambient intelligence for emergency monitoring and prevention. In. Lecture Notes in Computer Science: Vol. 5615. *Universal access in human-computer interaction. Intelligent and ubiquitous interaction environments* (pp. 199–208). Springer Berlin Heidelberg. doi:10.1007/978-3-642-02710-9_23

Mileo, A., Merico, D., Pinardi, S., & Bisiani, R. (2009). A logical approach to home healthcare with intelligent sensor-network support. *The Computer Journal*, *53*(8), 1257–1276. doi:10.1093/comjnl/bxn074

Nakic, D., & Loškovska, S. (2012). Creating and Using the Knowledge Archive in the Internet Medical Consultant for Decision Support at the Point of Care. [IJEHMC]. *International Journal of E-Health and Medical Communications*, *3*(3), 72–85. doi:10.4018/jehmc.2012070106

Nefti, S., Manzoor, U., & Manzoor, S. (2010). Cognitive agent based intelligent warning system to monitor patients suffering from dementia using ambient assisted living. In *Information society (i-society), 2010 international conference on* (pp. 92–97). London, UK: IEEE.

O'Grady, M. J., Muldoon, C., Dragone, M., Tynan, R., & O'Hare, G. M. (2010). Towards evolutionary ambient assisted living systems. *Journal of Ambient Intelligence and Humanized Computing*, *1*(1), 15–29. doi:10.100712652-009-0003-5

Palazzo, L., Rossi, M., Dragoni, A. F., Claudi, A., Dolcini, G., & Sernani, P. (2013). A multi-agent architecture for health information systems. In Frontiers in artificial intelligence and applications, Vol. 252: advanced methods and technologies for agent and multi-agent systems (pp. 375-384). IOS Press.

Pinardi, S., & Bisiani, R. (2010). Movement Recognition with Intelligent Multisensor Analysis, a Lexical Approach. In: *Ambient Intelligence and Smart Environments, Vol. 8: Workshop Proceedings of the 6th International Conference on Intelligent Environments* (pp. 170-177). IOS Press.

Reinisch, C., & Kastner, W. (2011). Agent based control in the smart home. In *IECON 2011 - 37th Annual Conference on IEEE Industrial Electronics Society* (pp. 334-339). Melbourne, VIC, AU: IEEE. 10.1109/IECON.2011.6119275

Russell, S., & Norvig, P. (2009). *Artificial Intelligence: A Modern Approach* (3rd ed.). Prentice Hall.

Sadri, F. (2011). Ambient intelligence: A survey. *ACM Computing Surveys*, *43*(4), 1–66. doi:10.1145/1978802.1978815

Sernani, P., Claudi, A., Calvaresi, P., Accattoli, D., Tofani, R., & Dragoni, A. F. (2014). Using 3D simulators for the Ambient Assisted Living. In 3rd International Workshop on Artificial Intelligence and Assistive Medicine, AI-AM/NetMed 2014 (pp. 16-20). Prague, Czech Republic: Ceur Workshop Proceedings.

Serrano, J. A. (2010). Fostering User Participation in Ambient Assisted Living Projects. [IJEHMC]. *International Journal of E-Health and Medical Communications*, *1*(2), 36–50. doi:10.4018/jehmc.2010040104

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Sun, H., De Florio, V., Gui, N., & Blondia, C. (2009). Promises and challenges of ambient assisted living systems. In *Information technology: New generations, 2009. ITGN '09. 6th International Conference on* (pp. 1201-1207). Las Vegas, NV, USA: IEEE. 10.1109/ITNG.2009.169

The Ambient Assisted Living (AAL) Joint Programme. (2004). Retrieved October 22, 2013, from http://www.aal-europe.eu/

Vacher, M., Portet, F., Fleury, A., & Noury, N. (2011). Development of Audio Sensing Technology for Ambient Assisted Living: Applications and Challenges. [IJEHMC]. *International Journal of E-Health and Medical Communications*, *2*(1), 35–54. doi:10.4018/jehmc.2011010103

Wolf, P., Schmidt, A., Otte, J. P., Klein, M., Rollwage, S., König-Ries, B., & Gabdulkhakova, A. (2010). Openaal – the open source middleware for ambient-assisted living (aal). In AALIANCE conference (pp. 1-5). Malaga, ES.

Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: Theory and practice. *The Knowledge Engineering Review*, *10*(2), 115–152. doi:10.1017/S0269888900008122

World Health Organization. (2011). Global health and ageing. World Health Organization, Geneve, CH.

Yu, E., & Giorgini, P. Maiden, N, & Mylopoulos, J. (2011). Social Modeling for Requirements Engineering. The MIT Press.

ENDNOTES

1. The ratio between the number of people over 65 years and those of working age.
2. It maintains an accurate, complete and timely list of agents, allowing them to publish the description of provided services (Bellifemine et al., 2007).
3. For example the VCA can send a request to turn the light on in the hallway because the Ambient Agent controlling the PIR sensor detects the presence of the assisted person.
4. The example has to be intended to purely show the syntax of the rules expressed in the AgentSpeak language.

This research was previously published in the International Journal of E-Health and Medical Communications (IJEHMC), 6(4); edited by Joel J.P.C. Rodrigues; pages 62-76; copyright year 2015 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 77
An Imagination of Organizations in the Future: Rethinking McKinsey's 7S Model

Oya Zincir
Istanbul University, Turkey

Ayşegül Özbebek Tunç
Istanbul University, Turkey

ABSTRACT

Organizations try to survive in a hypercompetitive, changing, and unpredictable environment. The form of this survival continuously changes and requires different tools, solutions, dynamics and drivers according to the actual time. Today organizations face with a big paradigm shift, the industrialization of information age. Organizations should find a new form on the basis of the new age requirements. Some authors have suggested some strategic prescriptions harmonized with the actual term such as McKinsey's 7S Model including strategy, structure, systems, skills, staff, style and shared goals. It is considered that there is a gap related to updating and upgrading these formulas by activating some new concepts such as morphing, organizational ambidexterity and so on. In this study, it is aimed to understand the organizations in the future with internal elements and outer changes which would affect them at the edge of the new era by using this model and offer many propositions.

INTRODUCTION

I don't read history. That's in the past. I'm thinking of the future. Henry Ford (Lander, 2014)

Managerial and organizational studies' aspects have changed through the years and so have organizations. According to the well-known study of Alvin Toffler (1980), three types of societies -based on the concept of waves; each wave pushes the older ones aside- have evolved through first, second and third waves as agricultural, industrial and information ages. First wave is the society after Agrarian Revolu-

DOI: 10.4018/978-1-7998-0951-7.ch077

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

tion which characterized by the enclosure of common land and the introduction of some technological innovations such as the seed drill. The concepts of this age led to the development of job specialization, complex political structures, non-portable possessions, architecture, and the rise of industry and commerce. Second wave is the society during the Industrial Revolution which characterized by mass production, distribution, consumption, education, recreation and destruction that combines those things with standardization, centralization, concentration, and synchronization. The structure of organization was bureaucracy. Third wave is the Post-Industrial Society characterized by actionable knowledge as a primary resource, demassification, diversity, knowledge based productions and non-linear change. The structure of organization is adhocracy. Based on these developments, organizations have evolved into various types through the ages.

With the 21st century, it has been recognized that new social and technical aspects of work systems, organizational structures, managerial practices, competencies and skills are needed. Some authors (Anderson, 2012; Rifkin, 2011, 2014) define this new age as "Third Industrial Revolution" and explain it as "the industrialization of Information Age". It can been said that it is the collaboration age; the modern enterprise is already a cyberfactory with the usage of robotics, internet of things and high-automation (Czarniawska, 2012) and high collaboration is very important for organizations and people who work for them.

Rifkin adresses renewable energy by mentioning its sustainable use and he stresses that this energy will be one of the main power element to change the understanding of world (Rifkin, 2011). Further, Rifkin (online, para.36-37) also explains organizations and workforce of the future by mentioning some important aspects: 'The new high tech workforce of the Third Industrial Revolution will need to be skilled in renewable energy technologies, green construction, IT and embedded computing, nanotechnology, sustainable chemistry, fuel-cell development, digital power grid management, hybrid electric and hydrogen-powered transport and hundreds of other technical fields. Entrepreneurs and managers will need to be educated to take advantage of cutting edge business models, including distributed and collaborative research and development strategies, open source and networked commerce, performance contracting, shared savings agreements, and sustainable low-carbon logistics and supply chain management". Organizations shift from centralized global companies to distributed small and medium size enterprise networks. Klaus Schwab (2016), using the term of Fourth Industrial Revolution, explained future concepts on the topics of mobile supercomputing, self-driving cars, geneting editing and many others by defining them with some megatrends (physical, digital and biological). These developments will affect economies, organizations, governments, societies and individuals. Schwab (2016) also emphasizes that the speed, breadth and depth and the holistic effect of the Fourth Industrial Revolution is a way more different than the other industrial revolutions. It is non-liner; will change many paradigms and recreate them by gathering some of them together and it needs to be considered holistically with countries, companies, sectors and more. He mentions that the new era is the "collaboration era", so that these institutions should collaborate together to do the necessary changes. The developments in advanced robotics, 3D printing, circular economy, internet of things (IoT), sharing economy, artificial intelligence, innovation making, humanistic approaches and many more, will reshape the world. In this manner, we need to understand new organizational concepts and economics systems and they should be renewed according to these developments. Machines will be used even more in doing business and so, not "race against the machine" but "race with the machine" understanding should be complementary. In short, people and machines should be team mates. Data driven developments help goods to be better produced, and this leads to increase asset productivity. Organizations collaborate with different partners

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

and learn from them even more and so forth. Thus, with all these, it can be said that the skill levels and managerial styles might be quite different from now; more agile, sustainable, collaborative, humanistic, flexible, technologic, creative, innovative, etc.

In this study, it is aimed to prescribe organizations of the future by rethinking McKinsey's 7S Model. It is a management model which developed by Robert Waterman and Tom Peters, most often used as an organizational analysis tool to assess and monitor changes in the internal situation of an organization. By using this model, it is aimed to understand organizations in the future with internal elements and outer changes which would affect them at the edge of the new era. On this framework, based on future thinking and new concepts, 'organizations in the future' is prescribed on some topics and furthermore, propositions are presented about some concepts. Some of them are liquid organization model, morphing, organizational ambidexterity, roles (staff who work for not jobs but for multiple roles), cyborganization, holacracy.

LITERATURE REVIEW

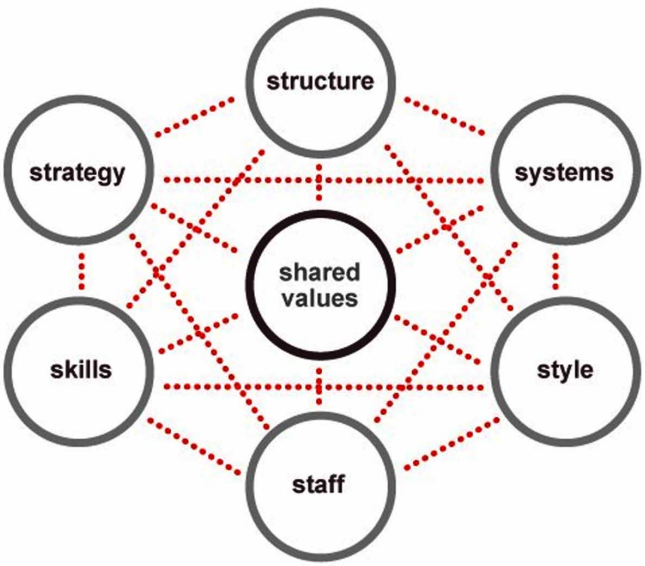
Introduced in the June 1980 issue of *Business Horizons*, in an article titled "Structure is not Organization" by Thomas J. Peters, Robert H. Waterman and Julien Phillips, McKinsey's 7S Model maps a constellation of interrelated factors that influence an organization's ability to change on seven elements. It suggests that significant progress in one element of the organization will be difficult without working on the others. The model has been used as an important tool to understand the complexity of organizations and provides a deep understanding of that structure alone doesn't refer to the whole organization.

The 7S are (which can be seen in Figure 1); hard elements that of structure, strategy, systems and soft elements that of skills, style, staff and shared values. As the authors explained (Waterman, Peters & Phillips, 1980), the elements are interconnected, they have effects on each other and do not follow a hierarchy or an importance degree. Strategy is the plan devised to maintain and build competitive advantage over the competition; structure is the way an organization is structured and who reports to whom, and systems is the daily activities and procedures that staff members engage in to 'get the work done'. These are hard elements; easy to define or identify and management can directly influence them. Shared values -called "superordinate goals" when the model was first developed- are the core values of the company that are evidenced in the corporate culture and the general work ethic; style is the style of leadership adopted; staff is the employees and their general capabilities, and skills are the actual skills and competencies of the employees who work for the company. These are soft elements; more difficult to describe, less tangible and more influenced by culture (Waterman, Peters & Phillips, 1980).

Peters (2011, para.15) also explains that now, in the new era, "hard is soft, soft is hard. The plans and the numbers are often 'soft' and the people (staff) and shared values (corporate culture) and skills (core competencies) which are truly 'hard' that is, the bedrock upon which the adaptive and enduring enterprise is built". It illustrates the paradoxical characteristics of the Hard Ss and Soft Ss elements. This perspective grounds on the "people" dimension. People's competences differentiate and change constantly (Dahlgaard-Park, 2013). Hamm (2015, para.3) explains this as that "the traditional hard skills are readily available and they cover a small percentage of what it takes to be 'excellent' as a company. On the other hand, the soft skills are really difficult to master. Finding people with the skills that are highest in demand is difficult". Currently, workforce is changing. Digitalization and other developments brought some serious effects on employees and customers. Employees expect a harmonized work-life balance. Customer expectations are

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 1. McKinsey's 7S Model



more specialized today. They are the forcing dimension of business (one example is user innovation). In this demand economy, people also choose which organizations that they want to work with and there are human resources intermediaries who gather people and organizations. Organizations will have the chance to use "human clouds" to choose its future employees. All these grounds on "people dimension" of the model. It is the main dimension which is effected by the changes in workplace behaviors and working conditions. When considering the model, it can be said that, social aspects of works, leadership/management style and staff' skills and competencies are the pillars of the model now by considering all the elements.

RETHINKING MCKINSEY'S 7S MODEL FOR ORGANIZATIONS IN THE FUTURE

The methodology of this study bases on rethinking McKinsey's 7S Model for the future understanding of organizations. As explained above, the model includes seven elements which interact each other. Mentioning that 'hard is soft, soft is hard' in this new era, in this study, every element are explained on some topics. For example, strategy is defined as a soft element and conceptualized by tactics and maneuvers, emergent strategies, strategic flexibility, transient advantage strategy etc. Staff, another example, is conceptualized by t-shaped skills, roles in organizations, technophilia, etc.

Soft Elements (Soft Ss)

As Peters (2011) explained lately, strategy, structure and systems are the soft elements of the model. In this study, it is considered to explain these elements under this category. When it is first written, the authors explains these elements as hard elements. By the time, business world has changed and so have organizations. First, each element is identified, and than some new concepts are explained and lastly propositions are introduced for each concept.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Strategy

In the model, strategy is defined as the plan devised to maintain and build competitive advantage over the competition. The focus is on 'building sustainable competitive advantage over the competition'. Strategy has been seen as a hard element and business world was (and still is) focused on building sustainable competitive advantage by creating a strong position among rivals and defending it for extended periods of time (McGrath, 2013). Lippitt (2007) underlined the fact that 'execution, not strategy, offers an exclusive competitive advantage'. Strategy execution which aims a competitive advantage is still very important for companies, but companies need to understand new rules of this new era. In today's world, competitors and customers are very unpredictable and amorphous (McGrath, 2013), and high-velocity competitive contexts are fast becoming the norm (Leavy, 2014).

In hyper-changing industries that effected by globalization and the digital revolution, competitive advantage often evaporates in a year. So, if companies want to stay ahead, they need to constantly start new strategic initiatives, building and exploiting many transient competitive advantage at once (McGrath, 2013). For this manner, first, it is essential to define 'transient advantage'. It is a strategy that prioritizes fast innovation to build a pipeline of competitive advantages where business environment requires the capacity to surf through waves of short-lived opportunities.

Proposition 1: In turbulent industries, organizations which follows transient advantage strategy can cope with emergent strategies more easily than the others.

Where being fast and changing fast are important aspects, and strategic flexibility is a vital capability, organizations need to give importance to tactics and maneuvers more. The reason is that tactics are actions that take advantage of opportunities offered by the gaps within a given strategic system. Maneuvers are actions that improve or maintain strategic situation of the company in a competitive environment or avoid a worse situation. 'Strategy is undertaken before the battle; tactics and maneuvers are implemented during the battle'. In this manner, where execution of the strategy is very important, organizations need to prioritize maneuvers and tactics to execute strategies. Strategic flexibility, also another important concept, is defined as the ability to precipitate intentional changes and adopt to environmental changes through the continuous rethinking of current strategies, asset deployment and investment strategies (Roberts & Stockport, 2014).

Proposition 2: In turbulent industries, organizations might prioritize maneuvers and tactics to execute strategies faster.

Proposition 3: Organizations that internalized strategic flexibility can adopt to environmental changes faster and easier.

Waterman et. al (1980), explained the strategy as the way a company aims to improve its position vis-a-vis competition. The focus is now on transient advantage strategies and wining the competition for a short period of time by, for example, creating new strategic initiatives, building organizational dynamics on strategic flexibility, etc. From this perspective:

1. Making innovation an everyday proficiency;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2. Practicing healthy disengagement as a normal, regular activity and continuously reconfiguring resources; and
3. Doing activities to achieve a dynamic balance between stability and agility (Leavy, 2014) are essential for organizations who want to cope with new environment changes and be competitive with the changing dynamics.

Proposition 4: Transient advantage strategy prioritizes fast innovation practices for organizations to compete among rivals.

Proposition 5: Transient advantage strategy generates the development of new strategic advantages for the organizations.

Porter (1980) asserted that companies should concentrate for one direction for their strategies to gain competitive advantage. Therefore he recommended to choose either differentiation-oriented or low cost-oriented as strategic positioning. He argues that there are insuperable trade-offs between these strategies, so that organizations have to make explicit choices (Porter, 1996). This assumption was a cornerstone for strategy field and also for companies who practice the findings of related academic studies in 1990s. Porter was not only a management thinker to support one point of view that means making choices that favor one activity over the other. According to this perspective, companies that pursue mixed strategies risk becoming "stuck in the middle," which may lead to poor performance (Raisch & Birkinshaw, 2008). This metaphor obviously presents how the duality is perceived as risky.

This duality is called 'ambidextrous approach' in terms of not only strategy but also structure. Ambidexterity concept is first used by Duncan (1976) who argued that firms need to innovate in order to insure long-term success and should consider a dual structure, one to initiate and one to execute, as a means to achieve these ends. This concept covers the idea that organizations should put in place dual structures, with one unit focusing on exploration and the other focusing on exploitation (Birkinshaw & Gupta, 2013). Exploration includes such things captured by terms such as search, variation, risk taking, experimentation, play, flexibility, discovery, innovation. Exploitation includes such things as refinement, choice, production, efficiency, selection, implementation, execution (March, 1991). It is argued that too much emphasis on exploitation to the exclusion of exploration leads to competency traps, inertia, and ultimately obsolescence. By contrast, too much focus on exploration to the exclusion of exploitation leads to a failure trap in which organizations gains no returns from its knowledge (Simsek et al., 2009). For this reason, some authors support the balance between these contrary ways. Actually today's organizations focus on strategies, in other words, how to formulate the best strategy to survive. On the other hand, future organizations should concentrate on dynamic capabilities, expertise activities, competencies and try to reveal the best mixture to be successful. This mixture provides them flexibility, diversity, harmony capability and holistic paradigm for organizations.

Proposition 6: Ambidexterity strategies might provide flexibility and diversity for organizations.

Proposition 7: Ambidexterity strategies might push the development of dynamic capabilities for organizations.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Structure

Structure is the way an organization is structured and who reports to whom. It divides tasks and then provides coordination, trades off specialization and integration (Waterman, Peters & Phillips, 1980). There are various structure types, such as functional, matrix, adhocratic, etc. A traditional hierarchy is not always the most effective structure; different organizations have different types of structure and these different structures effect the way they perform. Mintzberg (1980) identifies five main successful organizational structure as the entrepreneurial organization, the machine organization (bureaucracy), the professional organization, the divisional (diversified) organization and the innovative organization (adhocracy). Structure effects the organizational practices such as standardization of work processes, unit size, centralization/decentralization etc. It is important to mention that there's no "right" organizational structure, thus, to understand how structure relates to the variety of attributes in an organization is essential.

There is also a new concept which is called 'holacracy'. It is defined as a system of organizational governance in which authority and decision-making are distributed throughout a holarchy of self-organizing teams rather than being vested in a management hierarchy (Rud, 2009). As Robertson (2007) explains, holacracy aligns the explicit structure of an organization with its more organic natural form, replacing artificial hierarchy with a fractal "holarchy" of self-organizing teams ("circles"). A "circle" in holacracy is a self-organizing team. Each circle has an aim (purpose), and the authority to define and assign its own roles and accountabilities. Each circle is a holon -a whole self-organizing entity in its own right-, and a part of a larger circle. Each one connects to other of its sub-circles via a double-link, where a member of each circle is appointed to sit on the other, creating a bidirectional flow of information and rapid feedback loops. Each circle governs itself by uncovering the roles needed to reach the aim of the circle, and assigning circle members to fill them (Robertson, 2007).

Proposition 8: In holacratic organizations, staff might perform innovative work behavior more often.
Proposition 9: In holacratic organizations, staff might be more focused to team and organizational goals.
Proposition 10: In holacratic organizations, staff might get more rapid feedback and information from other members of the organization.

There is another concept that should be mentioned. It is 'morphing' which means continuous ongoing transformation. Morphing is used especially for change management practices. Generally, it mentions ongoing changes of the whole organization; for example, the changes of organization's strategies, structure and processes etc. Morphing embraces fluidity and continuous change rather than stability and certainty. It explains the rapid, seamless and more or less total change of the organizations which are mostly in high-velocity industries seek to regenerate their competitive advantage under hyper-competition conditions (Marshak, 2004). As explained above, transient advantage strategy helps organization to win the competition for a short period of time. This means that morphing could be an internal structural change process[1] to fit this strategical movements and flexibility.

Proposition 11: Morphing might help organizations to adapt strategic flexibility more easily.
Proposition 12: Morphing might help organizations to adapt short-term strategic changes more easily.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Liquid organization model is another concept that could be mentioned as a structural practice. Liquefying an organization means disrupting the industrial-age driven assumptions on which rigid structures are designed and move on to make it adaptive, dynamic and anti-fragile. Based on lean management and open collaboration principles, the liquid organization model is flat, meritocratic and value-driven, enabling stigmergic behavior and "organic" effectiveness (Verzera, 2014). Its meritocratic feature defines it as a performance-measured system. Stigmergic behavior means efficient collaboration between members and parts of the organization. As a whole, collaborative working, transparency, lean perspective, concentrated decisional power, agility, etc. are the basic features of the liquid organization. New organizations which could be called as "game changers" ground on these concepts. These entities lead to structural changes and new organizational structures; and new practices that change business world's dynamics.

Proposition 13: Liquid organization model might enhances stigmergic behaviors of staff by enabling collaboration.

Proposition 14: Liquid organization model might helps to improve organic effectiveness of the organization.

Czarniawska (2011) has developed an analytical metaphor of 'cyberfactory' that refers to a news agency where software has a central importance in the control of news production through the professional journalist. This terms brings the affirmation of the interconnectedness of people and machines in perceiving the world and brings another term that is 'cyborganization'. It is a term which describes the concept of a future form of organization, shaped by technology-mediated coordination, social computing and organizational algorithms. It basically describes a hybrid of a human and computer shaped form of organization (Kaminski, 2014). With new arising technologies and collaborative networks of resources and knowledge, one can assume that daily life and work will change dramatically within the next decades which have never seen before. Organizations can basically become structural artifacts that reflects as cyborganizations of the *Second Machine Age* which has been defined as a new era that makes people and software-driven machines substitutes, rather than complements. It is argued that it will bring technology-driven opportunities and challenges which will transform economy in the new era and will require human-beings to learn 'race with the machines' (Brynjolfsson & McAfee, 2014).

Proposition 15: In cyborganizations, staff can gain technological skills and be more innovative.

Proposition 16: Cyborganizational structure can act as a catalyst between staff and technological assets of an organization.

In terms of organizational paradigm, chaos theory refers to a subset of chaos incorporates principle of quantum mechanics and presents them in a complex system environment. On the basis of this theory, organizations face with the layout in chaos as an inevitable part of the environment (Ozbebek Tunc & Turan, 2013). In such an environment possibly in the future, fractal organization structure can help companies to survive. Because this type of organization offers open, self-organized, reconfigurable, nonlinear, self-contained and acquired structure. It discusses on the meaning of authority and hierarchy. From this aspect, it shows similarity with holacratic organization. This structure is self-similar, too irregular, goal-oriented, dynamically restructured and suitable for a turbulent environment (Ryu & Jung, 2004).

Proposition 17: Fractal organizations can help to link with internal regularity and external irregularity.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Systems

Systems are formal and informal daily activities and procedures that staff members engage in to 'get the work done'. Financial systems, promotion and reward systems, recruitment systems, and information systems are some examples of internal systems. Through these organizational systems, all the processes, information flows and key activities are carried out (Dahlgaard-Park, 2013). The way systems mirror the state of an organization is very important to enhance organizational effectiveness.

Getting Things Done (GTD) approach focuses on workflow and time management. It rests on the idea of moving planned tasks and projects out of the mind by recording them externally and then breaking them into actionable work items which allows one to focus attention on taking action on tasks (Allen, 2001). GTD approach can help staff and organizations simplify tasks, arrange them and engage in. This can increase efficiency and reduce time waste. These are really important for all types of organizations.

Proposition 18: GTD approach can help organizations to increase efficiency.
Proposition 19: GTD approach can help organizations to be more effective when doing business-related activities.

Integrative Decision Making (IDM) is defined as a decision-making process in which an individual balances tensions between opposing variables. It follows a four-step process. The first step (salience) seeks to define the relevant aspects of a problem. The second step (causality) seeks to determine the relationships between related and seemingly unrelated parts of the problem. The third step (architecture) involves the creation of a model that outlines the relationships defined in the previous two steps. The fourth step (resolution) outlines the decision and how it was reached (Business Dictionary).

Proposition 20: IDM can help staff to gain effectiveness in decision making process.
Proposition 21: IDM can help organizations to be more faster in strategy implementation process by making decision making process more effective.

All groups of individuals or organizations get things done via a 'crowd of wisdom'. Collective intelligence (CI) is defined as groups of individuals doing things collectively that seem intelligent. Wikipedia is an example of collective intelligence. CI refers to anyone in a large group who chooses to do an activity, without being assigned by someone in a position of authority (Malone et. al., 2009). Collaboration platforms such as Wikipedia allows organizations to connect and engage people and information, anywhere, anytime, and on any device. More data means more potential information that can be gathered about anything. This can lead to better decision making, being able to filter the information and make sense of it is crucial (Morgan, 2014). Thus, it can be said that knowledge and information gathering is essential for people and organizations so much more, especially for making new innovations, and CI helps to gather people for these purposes.

Proposition 22: CI can help staff to harness knowledge from other members of an organization, interest people and other organizations.
Proposition 23: CI can help staff to collaborate of other members of the organization, interest people and other organizations by displaying collective behavior.
Proposition 24: CI can help staff to learn new ways of doing business.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

A meme is an idea, behavior, style or usage that spreads from person to person within a culture. It acts as a unit for carrying cultural ideas, symbols, or practices that can be transmitted from one mind to another through writing, speech, gestures, rituals, or other imitable phenomena with a mimicked theme (Dawkins, 1990). Techno-memes (temes) are memes which live in technological artifacts instead of the human mind. Communication systems spread memes over longer distances, and broadcasting increases the number of copies sent out; sound and image recording systems raise longevity by storing previously ephemeral information; and new technologies raise fidelity enormously by making the shift from analogue to digital recording. The systems that do these are increasingly copying, storing and propagating information without human help (Blackmore, 2009). The Internet of Things is helping to create the vast amounts of big data. Devices will be able to "talk" to each other and to people on their own without human intervention or activation. This has the potential to make lives and workplaces easier and more streamlined as well as help people to better understand themselves about how they work, and how they live (Morgan, 2014). The current developments for the Internet of Things (IoT) can be a way of exchange temes between machines and this can reshape most industries by changing the way of doing the formal and informal daily activities and procedures.

Proposition 25: Temes can reflect as a factor that increase efficiency of an organization.
Proposition 26: Temes can reflect as a factor that make work-related activities more easily.

Artificial intelligence (AI) is a field of science and engineering concerned with the computational understanding of what is called 'intelligent behavior' and with the creation of artifacts that exhibit such behavior (Shapiro, 1992). It is defined as the study and design of "intelligent agents (machines)" which refers to the system that perceives its environment and takes actions that maximize its chances of success. AI promises to replace some human jobs, freeing people for other pursuits by automating manufacturing and transportation, self-modifying, self-writing, and learning software which relieves programmers of the burdensome task of specifying the whole of a program's functionality. AI can see relationships in enormous or diverse bodies of data that a human could not (Rajani, 2011).

Proposition 27: AI can help to increase staff's innovative work behavior by freeing them on thinking more about work-related issues.

Another concept that is interested in this chapter is biomimicry. It was first used by Benyus (1997) and defined as 'an approach to innovation that seeks sustainable solutions to human challenges by emulating nature's time-tested patterns and strategies' by Biomimicry Insitute (online). In short, it is about inspring from the nature. Companies inspire from the nature related to innovation, structure, strategies, business models, etc. It can be assessed that this inspiration is a systemic challenge because of its holistic effect to the whole. If a company design all activities, roles, competencies, goods and services on the basis of biomimicry view, it can be said that the organization is systematically biomimetic and so adapt to the environment as an eco-system. They can come up with organizational problems by considering nature's solution systems and ways. For example, they can model a kind of bird on producing new airplane or an ant subsystem on organizing team work.

Proposition 28: Organizations adopt biomimicric thinking can change their systems to the nature and so adapt easily to the future environment.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Hard Elements (Hard Ss)

As Peters (2011) explained lately, skills, staff, styles and shared goals are the hard elements of the model. Related new concepts are categorized and explained theoritically under every element. After that they are linked with future organization abilities and some propositions have been offered.

Skills

Skills are the actual skills and competencies of the employees who work for the company. They are the distinctive capabilities possessed by individuals, groups, and an organization as a whole. The skills variable can be referred to as the core competencies of an organization. It is a strong component of competitive advantages of organizations (Dahlgaard-Park, 2013).

Intrapreneurial behavior (IB) refers to the voluntary employee behavior and is a "bottom-top" process related to the spontaneous behaviors of individual employees which is aimed at the perception of opportunities, generation of ideas, creation of new products, as well as the development of new business lines. IB is an extra-role behavior which is self-initiated and reflects in individuals' risk-taking, innovation, and proactiveness (Antoncic & Zorn 2004; De Jong et al. 2011).

Proposition 29: By the time, the mission of being change agent has shifted from corporations to individuals.

Shirky (2010) argues that people are learning how to use free time more constructively for creative acts rather than consumptive ones, particularly with the advent of online tools that allow new forms of collaboration. People use online tools such as social media platforms and tools and create something collectively. Cognitive surplus (CS) is defined as the free time of the world's educated citizenry as an aggregate (Shirky, 2010). Cognitive surplus frames the abundance of free time as a collective resource that society may be able to use more effectively (Tomlinson & Silberman, 2012).

Proposition 30: Future employees can simultaneously and collaboratively create, contribute and use information via online tools.

Proposition 31: Future employees can spend their free time to learn and create something.

T-shaped skills as Hansen (2010) explains, is a metaphor used in job recruitment to describe the abilities of people in the workforce. It represents as a method for assessing résumés to build interdisciplinary work teams for creative processes. T-shaped people have two kinds of characteristics: The vertical stroke of the "T" is a depth of skill that allows them to contribute to the creative process that can be from any number of different fields such as an industrial designing, architecture, social science, etc. The horizontal stroke of the "T" is the disposition for collaboration across disciplines which is composed of two things:

1. Empathy that allows people to imagine the problem from another perspective; and
2. Enthusiasm about other people's disciplines, to the point that these people may actually start to practice the disciplines (Hansen, 2010).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Proposition 32: T-shaped skills can make workforce more collaborative.
Proposition 33: T-shaped people can have deeper understanding about theirselves and the others.

Staff

Staff is the employees of an organiation and their general capabilities. Staff is the crucial asset of an organization, it is a "sine qua non", a must. An organization should consider its staff and their general capabilities. The workforce is changing. New employee behaviors entering organizations are challenging the conventional idea of how employees work and what they expect from an organization. A new generational workplace means new behaviors, approaches, attitudes, and expectations about work and the workplace. Mobility allows employees to stay connected and working even when they are on the go. This helps make employee location independent. Organizations, large and small, can be comprised of international teams. Organizations can also develop a presence in any part of the world and work without boundaries (Morgan, 2014).

First, Y Generation entered the business life; in the near future, Millennials will enter. And they are quite different from other generations; they are shaped by technology (can be called as techno-born). They value community and creativity in their work. They have invested in human capital more than previous generations. They will contend with the effects of starting their careers during a historic downturn for years to come (Millennials Report, 2014). In short, the future employee wants to work for an organization that believes and invests in sustainability and community development, corporate social responsibility, health and wellness, employee happiness, ethical choices, and in making the world a better place (Morgan, 2014).

'Role' in an organization means one specific activity that the organization is counting on. For example, creating and maintaining a project release plan is a Project Manager's role. Whenever a new role is defined, it is then assigned to a member of the organization to fill and execute. This formally gives that member control to do what is needed to enact the accountabilities of the role, within any defined limits. It also gives others in the organization the reasonable expectation that they can ask the individual to account for any of the role's explicit accountabilities. In holacracy (explained above), the title of a role becomes secondary, the role is the list of explicit accountabilities. With the detailed accountabilities used in holacracy, any given role may have dozens of accountabilities, and any given individual may fill multiple roles (Robertson, 2007).

Proposition 34: Role-shaped organizations can help staff to gain new talents and competences.

Technology decisions are now in the hands of anyone and everyone within an organization, not just IT department or management. Cloud-based technologies can decrease the time to deploy, make upgrades and deployments easier and faster, and allow organizations to put together their own "stacks" made up of components from various vendors (Morgan, 2014). Technophilia is a term that refers to the enthusiasm toward technology with its rewarded and knowledgeable adoption (Ronit, 2011). It can reflect as techno-creativity of staff and, with the staff-driven activities, organizations can become techno-creative and change some industries' parameters and dynamics.

Proposition 35: Staff that reflects technophilic behaviors can help organizations to become techno-creative organizations.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Green-collar workers can be defined as workers who have interest, expertise, and knowledge in environmental issues and practices within organizations and are employed primarily for these reasons. A green-collar worker is an environmentalist, someone whose functions at the workplace include protecting the environment, encouraging sustainable workplace practices, and raising environmental awareness among co-workers. They advocate for the environment within their organizations and are cognizant of whether the organization acts in a socially responsible manner with regard to environmental issues (Harvey et. al., 2010).

Proposition 36: Green-collar workers can help an organization to give more importance to sustainability-related activities.

Proposition 37: Green-collar workers can help co-workers to be aware of environmental issues and social responsibilities of an organization.

Style

Style is the style of leadership adopted. It refers to the issues of how key managers behave in achieving organizational goals. It is also considered to encompass the cultural style of an organization. All organizations have their own specific culture and management style. The style/culture includes the dominant values, beliefs, norms, and traditions that are relatively enduring features that permeate organizational life (Dahlgaard-Park, 2013). Morgan (2014:91) explains that the future managers needs to "be a leader, follow from the front, understand technology. They need to lead by example, embrace vulnerability, believe in sharing and collective intelligence, be a fire starter, practice real-time recognition and feedback, be conscious of personal boundaries and adapt to the future employee". To be a fire-starter means to challenge conventional ideas about management and work not just take things at face value. A fire-starter must be curious ditch the template and the management formula. Lead by example means to be a role-model for others in every actions.

Proposition 38: When management/leadership style become more people-centric, staff can be more committed to the organization.

Proposition 39: When manager/leader is a 'fire starter', staff can be more innovative.

Some organizations can be 'managerless'. They are managed by employee-appointed committees. But how could they be managed? They can be guided by coaches and mentors. They can be called as 'employee-managed' organizations. Transparency, freedom with accountability, mentorship and support, communication and long-term approach are very important for these organizations. They must reflect values, hire the right people and understand that employees have their own lives (Morgan, 2014). 'Self-management' concept can reflect for these organizations.

Proposition 40: Self-management can help staff to be harmonized by collaborated and cooperated activities.

Proposition 41: Self-management can help staff to make decisions faster.

Proposition 42: In employee-managed organizations, staff can be more committed to the organization.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Shared Goals

Shared Values -called "superordinate goals" when the model was first developed- are the core values of a company that are evidenced in the corporate culture and the general work ethic. This element is considered to be the interconnecting center of all other variables and refers to the guiding concepts and meaning or the purpose of organizations' existence that are shared among all organizational members (Dahlgaard-Park, 2013). Morgan (2014) explains that future organizations will focus on creating a place of "want" instead of a place of "need", adapt to change faster, innovate anywhere and all the time, shift from profit to prosperity, and be more flatter. It can be said that future organizations will be more people-centric, humanistic, lean, innovative, technologic and flexible.

Proposition 43: When organizations become more people-centric, staff can have more flexibility with their roles, functions and work styles.
Proposition 44: When organizations become more flexible, staff can adapt to change faster.
Proposition 45: When organizations focus on creating prosperity, staff can be more people-centric and sustainable.

DISCUSSION AND CONCLUDING REMARKS

In this chapter, it is tried to relate between business concepts and futuristic trends for organizations. It is mentioned about the paradigm shift over years and described the dynamics of today. On the basis of today, it is assumed that some propositions for the future. The major purpose of this study is to prescribe the organizations of the future by rethinking McKinsey's 7S Model by mentioning some concepts such as morphing, holacracy, biomimcry etc. All these concepts are newly discussed theoretically and practically. Even though they are new concepts for today, they will be pillars and corner stones of future. They do shape not only companies but also whole business environment. Market, companies and even employees will have to change theirselves to this aura.

Furthermore, in this study, some propositions are offered to understand the organizations in the future with internal elements and outer changes which would affect them at the edge of the new era which is called as Second Machine Age or Third/Fourth Industrial Revolution by considering different dynamics and concepts. This revolution forces companies to turn traditional management roles upside down and inside out via digital mind on the basis of digitalization. This big change create a paradigm shift about business manners. No doubt that today's organizations should not pass over the symptoms of the new era to be a good example of 'future organization'. For this, they will have to be highly sensitive to all variables of environment and fast-changing as well as the world. This chapter aimed to spotlight on the question 'what is the future?' by offering some futuristic predictions which will be probably realized.

On the other part of this study, McKinsey's 7S Model maps a constellation of interrelated factors that influence an organization's ability to change on seven elements which have been called as hard and soft elements. Hard elements have been conceived as strategy, structure and systems. Soft elements have been conceived as staff, skills, shared values and style. Being hard means a 'hardly changeable, easy to define' structure, and being soft lies upon an 'uneasy to define' concept according to the model. But conversely, in this new era, which grounds on chaotic, hyper-changing and competitive environment, it can be said that 'hard is soft, soft is hard'. This reverse view let companies think all variables and

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

parameters as deeply embedded parts of the system and people as its center. It is because that the model basically grounds on the 'people' dimension now. Human-beings are developing more and more each day and gain new skills and competences. Furthermore, the growth rate of technologic developments is incrementally rising and human-machine collaboration occurs more often. In the near future, it is expected to see many new technologies and developments that grounds on it. This may lead new line of businesses and challenges in the current ones. Organizations may face with a big paradigm shift. The radical changes may push companies to be more innovative but unstructured, holistic but niche-based activities. Organization may look for new ways to build enduring but flexible structures and evolve from self-managed to unmanaged. They may have indulgence to democracy of ideas coming from employees, partners, customers, and even society. In other words, organizations may have to find a new form on the basis of the new age.

All in all, this chapter let anyone to rethink actual strategic models and imagine a future organization. It provides to find an answer to the basic question 'How may be a future organization?'. Future organizations may be structured with no hierarchy, holacratic, liquid, open to morphing, cyborganization, role-shaped, people-centric and fractal formation. They may adopt the systems including getting things done approach, integrative decision making process and temes. They may let to improve collective intelligence and artificial intelligence as a whole. They may prefer to formulate emergent tactics and maneuvers rather than long-term strategies, so they have internalized strategic flexibility. They also may plan to attain transient advantage which is contrary to today's dynamics. They may focus on contrary competencies as being ambidextrous and having duality. They may let their employees to be more entrepreneur, collaborative, innovative, t-shaped and learn to learn at their free time. They may prefer to work with green-collar workers and encourage them to be self-managed people. In conclusion, companies which do the best practice for these concepts in a sustainable manner may be able to survive with destructive competition. No doubt this survival story of companies may start with the exploration of their managerial DNA and then mapping out a route. For companies, the future requires to define uncertainties and make them certainties as soon as their rivals.

REFERENCES

Anderson, C. (2012). *Makers: the new industrial revolution*. New York: Crown Business.

Antoncic, B., & Zorn, O. (2004). The mediating role of corporate entrepreneurship in the organizational support performance relationship: An empirical examination. *Managing Global Transitions*, *2*(1), 5–14.

Benyus, J. M. (n.d.). *Biomimicry: Innovation Inspired by Nature*. Retrieved from https://books.google.com.tr/books?id=mDHKVQyJ94gC&printsec=frontcover&dq=biomimicry&hl=tr&sa=X&ei=WOSvVKW7JYX3UKOFg6gH&redir_esc=y#v=snippet&q=3.8&f=false

Biomimicry Enstitute Web Site. (n.d.). Retrieved 15 October 2015 from http://biomimicry.org/what-is-biomimicry/

Birkinshaw, J., & Gupta, K. (2013). Clarifying The Distincitve Contribution of Ambidexterity to the Field of Organization Studies. *The Academy of Management Perspectives*, *27*(4), 287–298. doi:10.5465/amp.2012.0167

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Blackmore, S. (2009). Dangerous Memes; or, What the Pandorans let loose. In S. Dick & M. Lupisella (Eds.), *Cosmos and Culture: Cultural Evolution in a Cosmic Context* (pp. 297–318). NASA.

Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age: Work, Progress and Prosperity in a Time of Brilliant Technologies*. New York: W.W. Norton & Company.

Dictionary, B. (2015), *Integrative Decision Making*. Retrieved from http://www.businessdictionary.com/definition/integrative-thinking.html

Czarniawska, B. (2011). *Cyberfactories: How News Agencies Produce News*. Edgar Publishing. doi:10.4337/9780857939135

Dahlgaard-Park, S. M. (2013). Seven-S Framework. In E. H. Kessler (Ed.), *Encyclopedia of Management Theory*. Thousand Oaks, CA: Sage Pub.,Inc. doi:10.4135/9781452276090.n217

Dawkins, R. (1990). *The Selfish Gene*. New York: Oxford University Press, Inc.

De Jong, J. P. J., Parker, S. K., Wennekers, S., & Wu, C. (2011). Corporate entrepreneurship at the individual level: measurement and determinants. In *EIM research reports*. Zoetermeer: EIM.

Duncan, R. (1976). The Ambidextrous Organization: Designing Dual Structures for Innovation. In R. H. Killman, L. R. Pondy, & D. Sleven (Eds.), *The Management of Organization, C.1* (pp. 167–188). New York: North Holland.

Hamm, D. (2015). *Investing in soft skills pays off for workers, employers: create a work environment that values communication, decision making*. Retrieved 1 October 2015 from http://www.insightonmfg.com/ArticleText.aspx?articleId=3854&magzId=2

Hansen, M. T. (2010). *T-Shaped Stars: The Backbone of IDEO's Collaborative Culture: An Interview with IDEO CEO Tim Brown*. Retrieved 5 October from http://chiefexecutive.net/ideo-ceo-tim-brown-t-shaped-stars-the-backbone-of-ideoae%E2%84%A2s-collaborative-culture/

Harvey, D. M., Bosco, S. M., & Emanuele, G. (2010). *The impact of "green-collar workers" on Integrative Decision Making*. Retrieved 5 October 2015 from http://www.businessdictionary.com/ definition/integrative-thinking.html #ixzz3si4xw9uY

Kaminski, J. (2014). *Cyborganization*. Retrieved 15 October 2015 from http://jermainkaminski.com/cyborganization/

Lander, S. (2014). *Conversations with History*. Hay House, Inc.

Leavy, B. (2014). Strategy, organization and leadership in a new transient-advantage world. *Strategy and Leadership*, *42*(4), 3–13. doi:10.1108/SL-05-2014-0038

Lippitt, M. (2007). Fix the disconnect between strategy and execution. *T&D, 61*(8), 54-57.

Malone, T. W., Laubacher, R., & Chrysanthos, D. (2010). The collective ıntelligence genome. *MIT Sloan Management Review*, *51*(3), 21–22.

March, J. G. (1991). Exploration and Exploitation in Organizational Learning. *Organization Science*, *2*(1), 71–87. doi:10.1287/orsc.2.1.71

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Marshak, R. J. (2004). Morphing: The leading edge of organizational change in the 21st century. *Organization Development Journal*, *22*(3), 8–21.

McGrath, R. M. (2013, June). Transient advantage. *Harvard Business Review*, 64–70.

Mintzberg, H. (1980). Structure in 5's: A synthesis of the research on organizational design. *Management Science*, *26*(3), 322–341. doi:10.1287/mnsc.26.3.322

Morgan, J. (2014). *The Future of Work: Attract New Talent, Build Better Leaders and Create a Competitive Organization*. Wiley.

Peters, T. (2011). *A Brief History of 7S (McKinsey's 7S) Model.* Retrieved 15 August 2015 from http://tompeters.com/docs/7SHistory.pdf

Porter, M. E. (1980). *Competitive strategy: techniques for analyzing industries and competitors*. New York: Free Press.

Porter, M. E. (1996). What is strategy? *Harvard Business Review*, *74*(6), 61–81. PMID:10158474

Raisch, S., & Birkinshaw, J. (2008). Organizational Ambidexterity: Antecedents, Outcomes, and Moderators. *Journal of Management*, *34*(3), 375–409. doi:10.1177/0149206308316058

Rajani, S. (2011). Artificial intelligence: Man or machıne? *International Journal of Information Technology and Knowledge Management*, *4*(1), 173–176.

Report, M. (2014). *15 economic facts about millennials*. Executive Office of the President of the United States: The Council of Economic Advisers. Retrieved from 10 October 2015 https://www.whitehouse.gov/sites/default/files/docs/millennials_report.pdf

Rifkin, J. (2011). *Third industrial revolution: how lateral power is transforming energy, the economy, and the world*. London: Palgrave Macmillan.

Rifkin, J. (2014). *The zero marginal cost society: the internet of things, the collaborative commons, and the eclipse of capitalism*. London: Palgrave Macmillan.

Rifkin, J. (n.d.). *Book Summary: Third industrial revolution: how lateral power is transforming energy, the economy, and the world*. Retrieved from http://www.wermutham.com/pdf/The%20Third%20Industrial %20Revolution.pdf

Roberts, N., & Stockport, G. J. (2014). Defining Strategic Flexibility. In The Flexible Enterprise, Flexible Systems Management (vol. 1, pp. 37-45). Springer Publishing. doi:10.1007/978-81-322-1560-8_3

Robertson, B. J. (2007). *Organization at the Leading Edge: Introducing Holacracy*. Retrieved 15 October 2015 from http://www.integralesleben.org/fileadmin/user_upload/images/DIA/Flyer/ Organization_at_the_Leading_Edge_2007-06_01.pdf

Ronit, R. (2011). Technophilia: a new model for technology adoption. *UK Academy for Information Systems Conference Proceedings*. Retrieved from http://aisel.aisnet.org/cgi/ukais2011/41

Rud, O. P. (2009). *Business intelligence success factors: tools for aligning your business in the global economy*. John Wiley & Sons, Inc.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ryu, K., & Jung, M. (2004). *Fractal Approach to Managing Intelligent Enterprises. In Creating Knowledge Based Organizations*. Idea Group. Retrieved from http://books.google.com.tr/books?id=7aHM3hjWsjUC&lpg=PA319&dq=fractal%20organization &hl=tr&pg=PA348#v=onepage&q=fractal%20organization&f=false

Schwab, K. (2016). *The fourth industrial revolution*. World Economic Forum.

Shapiro, S. C. (1992). Artificial Intelligence. In S. C. Shapiro (Ed.), *Encyclopedia of Artificial Intelligence* (2nd ed.; pp. 54–57). New York: John Wiley & Sons, Inc.

Shirky, C. (2010). *Cognitive surplus: how technology makes consumers into collaborators*. New York: Penguin Press.

Simsek, Z., Heavey, C., Veiga, J. F., & Souder, D. (2009). A Typology for Aligning Organizational Ambidexterity's Conceptualizations, Antecedents, and Outcomes. *Journal of Management Studies*, *46*(5), 864–894. doi:10.1111/j.1467-6486.2009.00841.x

Toffler, A. (1980). *Third wave*. New York: William Morrow and Co.,Inc.

Tomlinson, B., & Silberman, M. S. (2012). The cognitive surplus is made of fossil fuels. *First Monday*, *17*(11). doi:10.5210/fm.v17i11.4120

Tunc, O., & Aysegul, A. T. (2013). Businesses as Fractal Organisms in Chaotic Aura. *International Journal of Business and Management Studies*, 10-20.

Verzera, S. (2014). *Liquid organizations: building the next evolutionary stage of anti-fragility*. Academic Press.

Waterman, R. Jr, Peters, T., & Phillips, J. R. (1980). Structure is not Organization. *Business Horizons*, *23*(3), 14–26. doi:10.1016/0007-6813(80)90027-0

KEY TERMS AND DEFINITIONS

Artificial Intelligence: Artificial Intelligence refers to computational understanding of intelligent behavior which works as a system that perceives its environment and takes actions that maximize its chances of success.

Cyborganization: Cyborganization refers to a hybrid of a human and computer shaped form of organization that shaped by technology-mediated coordination, social computing and organizational algorithms.

Holacracy: Holacracy refers to an organizational form which grounds on a peer-to-peer workplace that of flexible structure, self-organized holons, unique decision making process, etc.

Morphing: Morphing means continuous ongoing transformation of an organization that of structural change, strategical change etc.

Second Machine Age: The Second Machine Age refers to the new age that grounds on software-driven technological developments and human-machine collaboration.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Third Industrial Age: The Third Industrial Age refers to the industrialization of information age that grounds on digitalization and renewable, distributable and smart energy for every continent of the world.

Transient Competitive Advantage: Transient Competitive Advantage is a strategy that grounds on short-lived opportunities and focuses to gain short-term strategic advantage of competition among rivals.

ENDNOTE

[1] It is not only a 'structural' practice; in this section, it is used for explaining possible future practices.

This research was previously published in Strategic Imperatives and Core Competencies in the Era of Robotics and Artificial Intelligence edited by Roman Batko and Anna Szopa; pages 101-125; copyright year 2017 by Business Science Reference (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Chapter 78
Events Automatic Extraction from Arabic Texts

Emna Hkiri
University of Monastir, Tunisia

Souheyl Mallat
University of Monastir, Tunisia

Mounir Zrigui
University of Monastir, Tunisia

ABSTRACT

The event extraction task consists in determining and classifying events within an open-domain text. It is very new for the Arabic language, whereas it attained its maturity for some languages such as English and French. Events extraction was also proved to help Natural Language Processing tasks such as Information Retrieval and Question Answering, text mining, machine translation etc… to obtain a higher performance. In this article, we present an ongoing effort to build a system for event extraction from Arabic texts using Gate platform and other tools.

1. INTRODUCTION

With the increase amount of textual data, there is a need of methodologies for analyzing, and representing these data in the best form. One of the challenges of the textual data analysis is information extraction (IE). IE as defined in the message understanding conferences (MUC), aims to analyze natural language texts and extract useful information from a particular domain or application. IE is used in processing text for Arabic systems such as search engines, clustering, classification, text mining systems Question and answering, information retrieval (Atta-Allah al, 2006), text summary, indexation, and classification(Cimiano et al, 2004), (Cohen et al, 2009).

DOI: 10.4018/978-1-7998-0951-7.ch078

Copyright © 2020, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

If this search domain is already extensively covered, the event extraction remains a vital and complex task. Event extraction from Arabic text could be beneficial for IE systems in various ways. For example, being able to determine events could improve the performance of news systems (Borsje et al, 2010), monitoring systems (kamijo et al, 2000) risk analysis applications (Capet et al, 2008) and decision making tools (wei et al, 2004).Extracted events are also extensively applied within the medical domain (Cohen et al, 2009), (Yakuji et al, 2001)

In this paper, the input of our information extraction system is a set of unclassified news websites articles written in Arabic language, and the output is a set of events with their attributes. In this context, we adopt the definition of the ACE (Automatic Content Extraction) (George, 2004) model that defines the event as an action, a process in which participants are connected. We describe in this paper the different steps followed to develop our event extraction system from Arabic texts.

The rest of the paper is organized as follows: Section (2) presents Arabic language and details its unique features. Section (3) deals with the definition of the event and present some related works. In section (3), we present our approach for the automatic extraction of events. The implementation of our approach in the Gate platform is described in section (4). In section (5), we evaluate the system to demonstrate its capabilities. Finally, we conclude our work with some perspectives.

2. ARABIC LANGUAGE

In Arabic processing domain, the research started in the 1970, even before editing texts problems are fully solved. Early studies focused mainly on lexicons. For ten years, the Web internationalization and the proliferation of media in Arabic demonstrated the usefulness of a large number of potential applications of the Arabic NLP. Therefore, researches have begun to address issues more varied as syntax, automatic translation, automatic indexing of documents, information retrieval, etc. (Farber; 2008).

In what follows, we present orthographic and morphological systems of the Arabic language and some problems of its automatic processing as the lack of free resources, lack of vowels and the agglutination of words. Those are the main issues that characterize Arabic and strongly contribute to the delay of its automatic processing. For morphological analysis, the absence of vowels adds additional ambiguity of Arabic words. As for the agglutination, it makes it more difficult to identify the segments that make up these words.

2.1. Features of Arabic Language

Arabic language is considered difficult to control in the NLP. It is a semitic language: it is written from right to left and its alphabet is an Abjad. It is composed of consonants and long vowels "ي", "ا" and "و".

In addition, Arabic writing is unicameral: capital letters and lowercase letters do not exist. It is a semi cursive language that most of its letters are attached to each other, their spellings differ depending on whether they are preceded or / and followed by other letters or are isolated. Only the letters "ا" and "و, د, ر" are never attach to the next letter.

The Arabic lexicon is composed of three main classes: nouns, verbs and particles. The class of particles is extended to include grammatical morphemes. This extension leads to a reorganization into four classes: verbs, nouns, pronouns and tools words and (Khoja et al, 2001):

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- **Verbs:** Most of them are formed on radical of 3 consonants such as the verb (“بتآ” kataba - write) and possibly 4 consonants such as the verb “wrدحج” dahrağa - drag, to drag). These roots may give rise to several patterns as a result of one or more morphological changes (eg, repetition of a consonant, a vowel lengthening, the addition of a morpheme, etc.);
- **Nouns:** The morphological system of Arabic nouns is composed of:
  - **Primitive Nouns:** Cannot be attached to a verbal root. Example: سأر (RAAS - head), (مد dam - blood), (مف fam - mouth), (بأ AAB - father) (خأ ah - brother) etc.;
  - **Deverbal Nouns:** Nouns that can be derived from a verbal root;
- The particle class is composed of numbers, pronouns and tool words:
  - **Numbers:** This category consists of names of simple numerals representing units: from (“رفص” sifr-zero, 0) to (“ةعست” tis’at - nine, 9); tens (“ةرشع” ašarat- dix, 10) (“نورشع” išruwn - twenty, 20) ... and (“نوعست”tis’uwn - ninety, 90); hundreds, etc. and composed numerals such as (“رشع ةعست” tis’at’ Asara - nineteen 19);
  - The pronoun is formed of:
    - Demonstrative pronouns (“ةراشإ ءامسأ” ásmaā’ íšaārat): (“اذه” hadaā – this);
    - Relative pronouns (“ةلوصوم ءامسأ”Asmaa mawsuwlat): they are related to the noun or pronoun that precedes (“الذي” al-ladiy - one), (“اللتيتن”) al-latayni - those, feminine, duel “اللذيذن” (al-ladayni - those, masculine, plural), etc.;
    - Personal pronouns (“ةلصفنم ريمض” damaāýir munfasilat): they identify the grammatical persons;
  - Tool words are used to locate facts or objects relative to time or place. Tool words are divided into quantifiers and adverbs, prepositions, etc.;
- **Prepositions:** eg (“في” fiy - in) or (“ىلع” Alaa - on);
- **Conjunctions of Coordination:** eg (“مث” tumma - then, then);
- **Adverbs:** eg (“ادبا” Abada - ever) or (“بشكلياع” bišaklĩ ‘aādiyyĩ - normally normal way);
- **Quantifiers:** eg (“لك” kulla - all) or (“ضعب” ba’d% a - a little).

2.2. Morpho-Syntactic Problems

Arabic is considered as difficult language to apprehend (written and spoken) in the domain of NLP (Atwell et al, 2004), (Debili and al, 2002), it presents specific morphological, syntactic, phonetic and phonological features; many researches have been conducted to study the specific linguistic phenomena in Arabic such as non-vocalization and agglutination. The difficulty occurred in a complex implementation of a theoretical model or prototype into a real system usable in large scale applications.

2.2.1. Lack of Vowels (or Diacritics)

The automatic analysis of the Arab faces the problem of the lack of vowels (or diacritics) in the texts. This problem causes a number of lexical ambiguity for example the word non-vowelized (“بتك” ktb) have 17 different voyellations.

A statistical study made by F. Debili (Debili, 2001) shows that only 19% of words in the corpus are unambiguous and the average is six voyellations by word. Thus, given the large proportion of ambiguous words, we can conclude that this phenomenon is very frequent for Arabic processing and requires serious studies. Vowels play a crucial role for Arabic; every letter of every word should have its short

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

vowel which is generally not the case in the texts. Short vowels remove the ambiguity, and they give the grammatical label of a word regardless of its position in the sentence.

2.2.2. Agglutination of Arabic Words

A major problem comes from the Arabic agglutination as a word can mean a whole sentence by itself. For example, the form agglutinated ("لياستكملها" is translated into lystkmlhA) to make it complete. The structure of the Arabic word is divided into five components: proclitic, prefix, base, suffix and enclitic (Larkey et al, 2002), (Tuerlinckx, 2004).

The proclitics are particles (conjunctions and / or prepositions and / or articles) .

Prefixes and suffixes express grammatical features and functions to indicate the nouns and the mode of verbs (active / passive) and the (number, gender, person, ...) .

The enclitics are personal pronouns.

The base is a combination of a root (The set of two, three or four consonants that represent a structure of language or verbal behavior of speakers).

To detect the root of a word, we must know the schema from which it was derived and remove inflectional elements that have been added:

- The last example: لياستكملها / lystkmlhA: to make it complete;
- Proclitic (ل / l): preposition = > (to);
- Prefix (ي / y): verbal prefix of unfulfilled time;
- Base: (إستكمل / istkml) derived from the root كمل by scheme (إستفعل / IstfEl) C1 C2 C3 = > (to be completed);
- Enclitic (ها / hA) suffix pronoun = > (it).

In any automatic processing perspective, the problem is to decompose the word into parts. This decomposition requires knowledge of higher level in case of ambiguities (if the word accepts multiple segmentations).

2.2.3. Segmentation of Texts

To analyze an Arabic text, we must add a segmentation phase because a paragraph may not contain any punctuation other than a point at the end. In addition, the Arabic sentence is long and complex; the recognition of the end is difficult because the punctuation is not systematic and sometimes particles can define another role than separating sentences. To solve this problem a number of studies were carried out as work Ouersighni and the results are quite modest (Ouersighni, 1998), (Ouersighni, 2000), (Ouersighni, 2002).

2.2.4. Word Order in the Sentence

The order of words in Arabic sentence is relatively flexible. The sentence can be ordered as: verb +subject + complement, Subject + verb + complement, complement +verb +subject So, we conclude that the relatively free words order accumulate the morpho-syntactic problems specific to Arabic. Indeed, taking into account all possible combinations of rules of sentence constituents cause, probably, additional syntactic ambiguities.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

2.3. Other Problems

The lack of strict lexical and grammatical structures in natural language poses many challenges for automatic analysis especially for the Arabic language, which by its morphological and syntactic properties is considered a difficult language in the domain of automatic processing of language (Larckey et al, 2002)(Mallat et al, 2013).

3. STATE OF ART

Named entity recognition acquired in recent years considerable attention. Several works has been done in NE recognition in Arabic. The majority of Arabic NER system focused mainly on traditional classes like (person, organization and location). Some systems introduced additional classes and categories related to a specific domain. Here we present a brief survey.

Several works applied statistical learning algorithms to Arabic NER. Abdul-Hamid, (2010) applied Support Vector Machines algorithm. Nezda (2006) as well as Benajiba (2007) use Maximum Entropy Models. Farber (2008) and Mohit (2012) use Perceptron. A range of works on Arabic named entity recognition includes labeled datasets, creation of gazetteers, rule-based and also statistical systems. RENAR (Rule-based Arabic named entity recognition system) is a recent rule-based approach (Zaghouani, 2012). It is based on gazetteers and grammar recognition rules to build the lexicon entities. The second system is proposed by Oudah and Shaalan (2012), it is a hybrid approach and combines rule-based approach with various statistical classifiers for extracting out of a set of NE classes.

Wikipedia has been the test-bed for few studies and systems of NER, they take advantages of its multilingual links and disambiguation and redirect pages to identify NE equivalences in target language. Behrang (2012) extended classes of named entities and developed a NER system using Arabic Wikipedia. Wentland et al. (2008) leveraged Wikipedia links and pages to produce NE for German and Swahili language. Other works combined Wikipidea with other resources to identify names and variants; Stern & Sagot (2010) exploited French wikipedia and GeoNamess to create a French Database. In addition, Toral et al. (2008) joint Wordnet to Wikipedia. Their resource contains 310,000 entities and variants found in Wikipedia

But while the problem of detecting and extracting named entities is relatively well understood, the event extraction remains a problem far from being solved tasks because an event mention can be expressed by several linguistic expressions and several sentences (Béatrice, 2012)(Aymen, 2007). The same definition of event varies depending on the domain of application: software engineering, history, philosophy and linguistics ...etc.

In Table 1 we present an overview of event extraction researches and their approaches.

Extracting event attracted a number of works, mostly restricted to English (Gabriel, 2006) or French (Gabriel, 2008), (Ludovic, 2011). This domain of research is widespread, but it is still very little studied for the Arabic language because it has several unique features that do not exist in other languages (Al-Jumaily, 2012) (Soraya, 2013). For other languages different approaches have proposed different models in the definition, which can be summarized in two main flows: The approach or TimeML standard and the ACE Model.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 1. Overview of event extraction researches and their approaches

Event Extraction Approach	Research	Method	Event	Year
KNOWLEDGE BASED	(Abuleil, 2007)	Pattern matching in Arabic news document	DIFFRENT DOMAIN	2007
	(Xu et al, 2006)	Using seeds and pattern learning by system MINIPAR	News events of Nobel reward domain	2006
	(Vargas – veraand Celjuska, 2004)	Semantic-syntactic patternss	KMI news event	2004
	(Aoneand Ramos-Santacruz, 2000)	Lexicon-syntactic patternss	Different domain	2000
DATA ORIENTED	(Naughton, 2008)	Support vector machine and linguistic modeling	ACE	2008
	(Ahn, 2006)	Nearest neighbor and maximum entropy classifiers	ACE	2006
HYBRID	(Sangeetha et al, 2010a)	Trigger detection using bayesien theory and argument detection using Ontosem system and text meaning representation	ACE	2010
	(Sangeetha et al, 2010b)	Using lexical chain to trigger detection and patter, matching to argument detection	Wikipedia document local event	2010
	(Piskorski et al, 2007)	Clustering pattern matching	Security related in news document	2007

The ACE model (Automatic Content Extraction) defines the event as an action, a trial at which the participants are connected. Events are represented according to their attributes and their participants. ACE event can have a number of participants, and each participant is characterized by a role (agent, object, source, target).The event attributes are its types (destruction, creation, transfer, movement, interaction) and the modality of an event is (real and not real). Several systems and projects were founded on this model such as the Aone and Ramos-Santacruz approach, which is interested in events and relations between them; it is based on patterns for the labeling. The idea is to search find all informations about the event and to fill the grid corresponding to it for example the event "bying" their REES tool do extract from the text informations about the time, place, the seller and the buyer (Aone and Ramos-Santacruz, 2000).

The approach of Ahn is founded on machine learning. It is composed of 4 modules, each one is generated by TiMBL tool these modules are trigger identification; events anchors; attributes identification and the last one corresponds to the coreference of the events (Ahn, 2006).

The same principle of modules is used in Beatrice approach for French events extraction, the module of anchors identification is based on the prepositions or determinants and the events could be nouns, verbs, adjectives or adverbs. The identification of these anchors is based on classification of words in the texts. The basic indices of the classification are morpho-syntactic categories of words, lexical features (lemma, form, depth in the dependency tree) issued from WordNet, words surrounding the event, etc...

The approach or TimeML standard (Roser, 2006) is based on three fundamental concepts: time, event and relations. The TimeML project was created with the aim of improving Q-A systems to deal with questions of temporal order on the entities and events. TimeML adopts a large conception of events. The <event> tag classifies an annotated event by the TimeML ontology; the latter includes different classes of events (aspectual, intentional action, intentional state, perception, and state case).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Parent used TimeML guide to manually annotate the English texts. Their model is based on patterns and syntactic analysis. The author is interested in adverbs, verbs, adjectives and nouns events. In their approach, they are based on words of action verbs and syntactic rules to annotate nouns events which depend on temporal prepositions. These prepositions are initially issued from TimeML guide (Parent et al., 2008).

The approach in (Saurí et al., 2005) focused on the disambiguation of nouns that may have event interpretation using SemCor corpus and TimeBank1.2. He extracted 25 sub-tree of the semantic network, which contains essentially nouns referring to events. This approach is also based on a statistical module and WordNet for event annotation. If the process of events search in WordNet is worthless they pass to use their Bayesian classifier. This last is founded on a set of rules learned from the SemCor corpus. The evaluation of their results is done by comparison with the tags of the TimeBank1.2.

In our work we are interested in the EVITA system (Sauri, 2005). This system is part of the project TARSQI (Verhagen, 2005) respecting the TimeML specification, EVITA is a tool for event detection for the English language. In this system, we consider verbs, nouns and adjectives as events. These textual elements are considered the most meaningful. In this system for each type of "trigger", an extraction method is associated particularly considering the grammatical and textual features.

EVITA system tests first if the category of a word is a verb, noun or adjective and then classifies it among the types of predefined events. These processes are complemented by the identification of the various arguments whose objective is the identification of event participants. The identification is made by detecting pairs between the trigger and the other entities of the sentence.

Detection of pairs is realized based on indices such as the morpho-syntactic category of the two elements, the type of event, the type of entity or the dependency relations in the pair.

4. PROPOSED METHOD

In our work, as already mentioned, we were inspired by the EVITA system. EVITA considers verbs, nouns and adjectives as the most meaningful events. For us, we judged that only verbs and nouns as events while adjectives as less significant.

Our event extraction system was built using GATE General Architecture for Text Engineering). Our system identifies predefined named entities (names of "people", "places" "organization", and "dates"), and the relations between the entities and the defined events. The extraction of an event consists of the discovery of links between the "trigger" of the event and its arguments/attributes. The extraction of the link is established based on a syntactic analysis "dependency analysis» and of extraction rules exploiting this analysis. To implement this task we used JAPE transducer (JAVA Annotation Pattern Engine) provided by Gate Toolkit.

Our event detection system is composed of five main phases summarized below.

The first phase is processing that aims to represent data in a form that can be efficiently analyzed and to improve the data quality by reducing the quantity of trivial noise (the non-Arabic words). We carry out text transliteration and segmentation. The next step is segmentation and tokenization of each sentence in the arabic text. We applied a morphosyntactic analysis and stop-word elimination and we finished by the normalization.

The second phase is lemmatization; for each word in the Arabic text, we assign its lemma. We obtained verbal and nominal lemma depending of the grammatical category of the term.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

The third phase is the identification of triggers; it is done by the use of gazetteers. These lists are composed of verbal lemmas for verbal triggers and nominal lemmas for nominal triggers. This phase is realized as the following steps:

- We compare each lemma to the list of triggers (list of gazetteers);
- If matching, we annotate the corresponding lemmatized word as an event trigger.

In the fourth phase, as in the EVITA system we associated the trigger to the class that it represents. At this stage of our work we do not deal with polysemous triggers, we just developed gazetteers for monosemic triggers.

The last phase includes the identification of participants and their semantic roles. For identification, the transducer is used to extract the main dependencies "subject", "object", "preposition", "agent » by executing the developed rules. The purpose of the semantic analysis is the assignment of semantic roles to participants extracted in the previous phase.

The algorithm corresponding to the proposed method is presented as shown in Algorithm 1.

5. SYSTEM IMPLEMENTATION

5.1. WORK Environment

5.1.1. GATE

Most of existing tools for text engineering, if not all, were not originally developed for the treatment of the Arabic language. Most tools were built for English, French or other languages as ACABIT, LEXTER, FASTER, ANA, EXIT (Roch, 2004). Some tools can support or be modified to treat Arabic as Nooj and GATE. We chose the latter for the implementation of our system. GATE[1] (General Architecture for Text Engineering) is a Java open source platform dedicated to textual engineering; it appeared to us well suited for the development of our system. It is a toolkit for natural language processing and very useful for information extraction.

By a system of plugins GATE provides to users a variety of modules dedicated to textual analysis. The most commonly used are tokenizers (segmenters), Part Of Speech Taggers (morpho-syntactic taggers), Gazetteers (lexicons) and transducers (JAPE).

Named entities extracted by GATE correspond to person names, organization, location, dates, etc... To realize new annotation Gate permits to load new resources and plugins. Also, it permits to combine and parameter them within the same treatment chain.

Figure 1 shows the interface of the Gate platform and its basic components for extraction of different types of named entities mentioned above:

- "Language Resources" is composed of linguistic data such as documents, corpora or ontologies;
- "Processing resources" for different resources used for annotation, such as analyzers, segmenters, taggers etc.;

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Algorithm 1. Event annotation algorithm

Input of (Preprocessing)
Corpus C, Document D
(1) Preprocessing of the documents
For each document D in Corpus C do
Remove(digits and punctuation marks) (=,+,$,*.,:,/,, [,] (,)_,&,|,-,^,, {, }, !,§,-)
Filter(the non Arabic words)
Transliteration
Segmentation of text into sentences
Segmentation of sentences
Morphosyntactic analysis
Extraction of simple and composed terms
Remove (Arabic stop words by the anti-dictionary)
Normalization
End for
Output: Preprocessed Document D

Input of (Lemmatization**):**
$S = \{S_i; \bigcup_n^1 S_i = D\}$ the set of segments forming each document D in the corpus.
$T = \{T_i; \bigcup_n^1 T_i = S_i\}$ with T the set of terms forming Si
(2) Lemmatization of T
Output of (1):
$L = \{L_i; \bigcup_n^1 L_i = S_i\}$ with L the set of lemmas forming Si

Input of (Triggers identification & Class Attribution**):**
$L = \{L_i; \bigcup_n^1 L_i = S_i\}$ L set of lemmas forming S_i from**(1).**
$EI = \{EI_i; \bigcup_n^1 EI_i ; C_i\}$EI a set of indicators relating to Events,
Ci corresponds to the event Classes
(3,4) Triggers identification & Class Attribution
For each L_i in S_i do
If ($\exists L_i \in EI$) then
Annotate (T_i) with T_i term which corresponds to lemma L_i
Return (T_i, C_i)
End if
End for
Output of(3,4):
{T, C} the set of terms with their corresponding event_class

Input of (Identification of participants/ arguments in the event**):**
{T, C} the output of **(3,4):**
$N = \{\bigcup_n^1 N_i, RN_i; \in S_i\}$ N is the set of named entities annotated by gate inS_i.
RNi={date, location, person, organization }
$Ev = \bigcup_n^1 J_i , J_i\{I_i\}$ Represents Jape event annotation rule having Ii as trigger/ indicator.
(5) Identification of participants/ arguments in the event (Syntactic and semantic analysis)
For each J_i in Ev do
Extract (Ni of I_i) with Ni is the set of participants
End for
Output of (5):
{T,C,N} the set of triggers, class and the participants of each trigger

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- On the right part, different named entities are colorized and annotated by GATE, which are the default_token (Date, Location, Organization, Person) we do mention that GATE does not the annotation of events.

5.1.2. External Resources

For the Arabic language, Gate does not have a satisfactory number of instances in his predefined Gazetteers of named entities. For example, the Gazetteers_Personne is composed only of two types (female_names.lst and male_names.lst) whereas for English almost 20000 entries, classified in nine different types (person_ambig.lst, person_female_cap.lst, person_female.lst,person_full.lst, person_male_cap.lst,person_male.lst, person_relig.lst,person_sci.lst, person_spur.lst).

To overcome, this problem of deficiency in the Gazetteers, we have used external resources to enrich Gate predefined Gazetteers, which consist of five different types; all built manually using our corpus and web resources (see also Table 2):

- **Person gazetteer:** A list of 3257 complete names of people found in Wikipedia and other websites. These names are normalized and formalized to end with a list of 3000 names (first and final names);
- **Organization gazetteer:** We collected 2400 names of organizations;

Figure 1. Interface of GATE platform



EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

- **Location gazetteer:** We enrich it by Arabic Wikipedia, taking the page labeled "countries of the world in Arabic" to retrieve location names. It consists of 2500 of continents, cities, countries… etc.;
- **Date Gazetteers:** We did enrich this predefined gazetteer by 910 new entities;
- **Event Gazetteers:** Gate as mentioned above do not do the annotation of event and have no predefined Gazetteer for it, that's why we did create a new one, composed of verbal and nominal list of triggers. These lists are collected from our domain corpus. To enrich it we added to their extracted lemmas their synonyms using Arabic WordNet[2] and as well as the argument prepositions structures (for verbs) by the Arabic dictionary. This final Gazetteer contains 700 entries.

5.2. System Description

The implementation in GATE platform has necessitated the use of additional Gazetteers and tools; therefore the installation of new plugins.

For the lemmatization phase, we used morphological analyzer (GATE Morphological Analyser). This analyzer has allowed us to obtain, for each word in the Arabic text, the associated lemma. These lemmas are then used in the next phase.

The second phase of identifying triggers is performed using resources Flexible Gazetteer, which allowed us to compare the tokens of the Arabic text to lemmas in both created Gazetteers (one for verbal lemmas and the other for nominal lemmas).

In the third phase, we classified the verb and noun lemmas into subclasses of the "Event" class (Attack, Military_Operation, Crash, Shooting, Damage, Bombing, Death, Kidnapping, War, Injure).

In the fourth phase, for the identification of participants, sentences were cut with the Noun Phrase Chunker and VP Chunker resources. For syntactic analysis, Stanford Parser was used and configured to process the Arabic language.

In the last phase, we pass to develop our linguistic rules with JAPE transducer (Java Annotation Patterns Engine**).** The role of the transducer is to identify named entities (Person, Organization, Location, and Date). So in our case, JAPE [3]executes the developed rules to extract the various arguments of the event. Writing these rules is followed by a test phase which aims to detect annotation errors and therefore the correction of these rules.

The Figure 2 shows an example of Arabic text where events are annotated by our tool: they belong to some predefined event classes such as:

Table 2. Enrichment of Gate Gazetteers

Gate	Predefined Entries	Enriched Entries	Overall
Person	1700	3000	4700
Organization	96	2400	2496
Location	485	2500	2985
Date	84	910	994
Event	0	700	700

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Figure 2. Example of annotated events from an Arabic text



- **Death:** مقتل
- **Attack:** الهجوم
- **InjureEvent:** إصابة
- **Military_ Status:** تعزيزات عسكرية
- **War:** الحرب العالمية الثانية / الحرب الأهلية

6. EXPERIMENTAL EVALUATION

6.1. Experimental Setup

6.1.1. Data

We used MILTcorp to test our system. MILTcorp is annotated especially for the Event extraction task. This corpus is related to our domain research; the military domain. Its articles are collected from news websites and news wire like (aljaziraa, al_arabia, al_manar, France24 (Arabic) also from electronic journals (al-Quds).

6.1.2. Annotation

The evaluation of the annotation is performed as follows:

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 3. Data collection

	Test Corpus
Number of sentences	1650
Number of words	45350
Number of words/sentences	24

- Three human annotators manually annotated texts of test corpus. These three human annotations are gathered in one annotation used as "the reference annotation;"
- In a second time, we annotated automatically these texts with our proposed system.
- We have established a comparison between these two annotations with Corpus Quality Assurance. This Gate tool makes the comparison between two annotations on a corpus;
- The result of the comparison is the generation of precision, recall and f-measure values.

6.2. Evaluation Metrics

To measure the effectiveness of our proposed system, we used a set of well-known evaluation metrics: precision, recall, and F- measure. Precision measurement is defined by the percentage of entities found by the system and which are correct, and the recall is defined as the ratio between the numbers of found correct entities by the number of entities extracted from the reference articles. The overall evaluation of entities extracted by our system compared to reference articles is based on the use of F-measure. This measure combines the precision and the recall, it is defined as following (we have used the $F_{ß}$=1; the precision and the recall are weighted equally):

$$\text{Precision} = \frac{\text{number of found correct entities}}{\text{number of entities found}}$$

$$\text{Recall} = \frac{\text{number of found correct entities}}{\text{number of entities in the corpus}}$$

$$\text{F} - \text{measure} = \frac{\left(ß^2 + 1\right)*\text{Precision}*\text{Recalll}}{ß^2*\text{Precision} + \text{Recall}}$$

6.3. Experimental Results

To evaluate the overall system, we established comparison with Gate baseline system. As we mentioned before Gate do not annotate event entity in texts. Table 4 shows results for different entities annotated by Gate (using only the basic predefined gazetteers without the enriched gazetteers). The results are automatically generated by the Corpus Quality Assurance.

Table 5 presents the results obtained by using the enriched gazetteers (person, organization, location, date) and our method of event extraction.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Table 4. Gate baseline results

NE	Precision	Recall	F-Measure
Date	51%	29%	36, 97%
Person	33%	15%	20, 62%
Organization	52%	33%	40,73%
Location	59%	44%	50,40%
Event	0%	0%	0%
Overall	**39%**	**24,2%**	**29,67%**

Table 5. Annotation results with event, date, person, organization, location gazetteers

Named Entities	Precision	Recall	F-Measure
Date	84.03%	77.19%	80,64%
Person	83.04%	79.34%	81,14%
Organization	78.6%	62.11%	69,38%
Location	86.14%	77.42%	81,54%
Event	70%	35%	46,66%
Overall	**80,36%**	**66,21%**	**71,87%**

Table 4 shows the results obtained by using the enriched gazetteers (person, organization, location, date) and our method of event extraction.

From Table 5, we see the effect of the annotation of events on the annotation of named entities extracted by Gate. We obtain an improvement of accuracy of 39% to 80,36% same with the recall from 24,2% to 66,21%.

6.4. Discussion

The strength of our Annotation system is in recognizing Event, Location and Date entities, which is attributed to the high named entity coverage of our NE base containing the enriched gazetteers.

It is important to observe that our system shows good recall for Person Names which were more abundant in the training corpus. Besides this, the corpus had a heterogeneous mix of Person proper names not only from Arabic countries but also from the continents of Africa, Asia and America. A good recall percentage for Person entity names is encouraging as the named entities related to south Asia and America have no phonetic similarity with the native person names entities available in Arabic countries. This gives good credence to our annotation process presented above.

A detailed look to the Event Recall reveals that our NE recognition system performs poorly for Event entities that appear in the corpus, this is due to the specificities of Arabic sentences such as free word order that affects the identification of participants and their semantic roles. In addition, we notice that our system cannot efficiently handle the acronyms, short names or the abbreviations of organization entities.

These previous results are satisfying according to limitations due to the complexity of Arabic sentences and the lack of adequate tools for the treatment of the Arabic language.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

In the next future, we plan to complete the event Gazetteers (verbal and nominal) and the disambiguation of polysemous triggers also to increase the size of our corpus to obtain a higher performance of the system.

7. CONCLUSION

In this paper we have presented our system to detect and extract Event named entity in Arabic text using Gate tool. The event extraction was performed through several stages; collection of data, preprocessing, Triggers identification, Class Attribution, Identification of participants/ arguments in the event. The evaluation results show that our method is very promising, This system can be generalized for the purposes of NLP enrichment which can be implemented in many fields such as question answering, text mining, summarization, etc... There are many directions for future work. One of the basic directions is to identify and annotate different types of relations between the named entities.

REFERENCES

Abdul Halin, A., Rajeswari, M., & Abbasnejad, M. (2013). Soccer Event Detection via Collaborative Multimodal Feature Analysis and Candidate IAJIT First Online Publication Ranking, International Arab. *Journal of Information Technology*, *10*(5).

Abdul-Hamid, A., & Darwish, K. (2010). Simplified feature set for Arabic named entity recognition. In: *Proceedings of the 2010 named entities workshop*. Association for Computational Linguistics, Uppsala, pp 110–115

Abuleil, S. (2007) Using NLP Techniques For Tagging Events in Arabic Text, in *Proceedings of the 19th IEEE International Conference on Tools with Artificial Intelligence*, vol. 2, pp. 440-443 10.1109/ ICTAI.2007.179

Ahn, D. (2006). The stages of event extraction. In ARTE '06: Proceedings of the Workshop on Annotating and Reasoning about Time and Events, pages 1_8, Morristown, NJ, USA, jul 2006. Association for Computational Linguistics. 49 10.3115/1629235.1629236

Al-Jumaily, H. T., Martínez, P., Martínez-Fernández, J. L., & Van der Goot, E. (2012). A real time Named Entity Recognition system for Arabic text mining. *Language Resources and Evaluation*, *46*(4), 543–563. doi:10.100710579-011-9146-z

Antonio, T., Muñoz, R., & Monachini, M. (2008). Named Entity WordNet. In *Proceedings of the 6th Conference on Language Resources and Evaluation*. Marrakesh (Morocco)

Aone, C., & Ramos-Santacruz, M. (2000). REES: a Large-Scale Relation and Event Extraction System. In Proceedings of the sixth conference on Applied natural language processing, ANLC '00, pages 76-83, Stroudsburg, PA, USA. Association for Computational Linguistics. 4, 42, 47, 50, 76 10.3115/974147.974158

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Aone, C., & Ramos-Santacruz, M. (2000)"REES: A Large-Scale Relation and Event Extraction System," in *Proceedings of the sixth conference on Applied natural language processing*, pp. 76- 83 10.3115/974147.974158

Ataa-Allah, F., Boulaknadel, S., El Qadi, A., & Aboutajdine, D. (2006).Arabic Information Retrieval System Based on Noun Phrases., ICTTA'06, Information and Communication Technologies, nº 24-28, vol. 1, Avril, Damas, Syrie, pp: 1720-1725.

Atwell E, L. Al-Sulaiti, S. Al Osaimi, Abu Shawar B.(2004) "Un Examen d'Outils pour l'Analyse de Corpus Arabes", JEP TALN, Session on Arabic Language Processing, Fès, 2004

Benajiba, Y., & Rosso, P. BenedíRuiz JM. 2007. ANERsys: an Arabic named entity recognition System based on maximum entropy. In: Gelbukh A (ed) Proceedings of CICLing, Mexico City. Springer, pp 143–153

Bethard, S., & Martin, J. H. (2006). Identification of event mentions and their semantic class. In *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing (EMNLP 2006)*, p. 146–154, Sydney. 10.3115/1610075.1610098

Borsje, J., Hogenboom, F., & Frasincar, F. (2010). Semi-Automatic Financial Events Discovery Based on Lexico-Semantic Patterns. *International Journal of Web Engineering and Technology*, *6*(2), 115–140. doi:10.1504/IJWET.2010.038242

Capet, P., Delavallade, T., Nakamura, T., Sandor, A., Tarsitano, C., & Voyatzi, S. (2008).: Intelligent Information Processing IV, IFIP International Federation for Information Processing, vol. 288, chap. A Risk Assessment System with Automatic Extraction of Event Types, pp. 220-229. Springer Boston (2008)

Cimiano, P., Staab, S.(2004): Learning by Googling. SIGKDD Explorations Newsletter6(2), 24-33 (2004)

Cohen, K. B., Verspoor, K., Johnson, H. L., Roeder, C., Ogren, P. V., Baumgartner, W. A., Jr., . . . Hunter, L. (2009).: High-Precision Biological Event Extraction with a Concept Recognizer. In: Workshop on BioNLP: Shared Task collocated with the NAACL-HLT 2009 Meeting. pp. 50-58. Association for Computational Linguistics (2009) 10.3115/1572340.1572348

Daille, B., Fourour, N., & Morin, E. (2000). Catégorisation des noms propres: Une étude en corpus. *Cahiers de Grammaire*, *25*, 115–129.

Debili, F(2001) traitement automatique de l'arabe voyéllé ou non IRMC tunis 2001

Debili, F., Achour, H., & Souici, E. (2002). "La langue arabe et l'ordinateur: de l'étiquetage grammatical à la voyellation automatique", Correspondances de l'IRMC, N° 71, pp. 10-28, 2002.

Doddington, G., Mitchell, A., & Przybocki, M. (2004). The Automatic Content Extraction (ACE) Program Tasks, Data, and Evaluation. *Proceedings of LREC*, *2004*, 837–840.

Elkhlifi, A., & Faiz, R. (2007).: Machine Learning Approach for the Automatic Annotation of the Events. FLAIRS Conference 2007: 362-367.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Farber, B., Freitag, D., Habash, N., & Rambow, O. (2008). Improving NER in Arabic using a morphological tagger. In *Calzolari N, Choukri K, Maegaard B, Mariani J, Odjik J*. Piperidis.

Frasincar, F., Borsje, J., & Levering, L. (2009). A Semantic Web-Based Approach for Building Personalized News Services. *International Journal of E-Business Research*, *5*(3), 35–53.

Gabriel, P., Gagnon, M., & Muller, P. (2008). Annotation d'expressions temporelles et d'événements en français. In Frédéric Béchet, editor, Traitement Automatique des Langues Naturelles (TALN), Avignon, 09/06/08-13/06/08, page (support électronique), http://www.atala.org/, 2008. Association pour le Traitement Automatique des Langues (ATALA). 49, 51, 100, 150, 152, 153.

Hassoun, 2002. Mallat S, Zouaghi Z, Hkiri E, Zrigui M, (2013). Automatic translation of Arabic queries based on lexical disambiguation method for Bilingual information retrieval Fourth International Conference on Information and Communication Technology and Accessibility. TUNISIA, Hammamet, October, 24-26, conference Program

Kamijo, S., Matsushita, Y., Ikeuchi, K., & Sakauchi, M. (2000). Tra_c monitoring and accident detection at intersections. *IEEE Transactions on Intelligent Transportation Systems*, *1*(2), 108–118. doi:10.1109/6979.880968

Khoja, S. (2001) "APT: Arabic part-of speech tagger". Proceeding of student workshop at the 2nd meeting of the NAACL, (NAACL'01), Carnegie Mellon University, Pennsylvania, 2001

Larkey, L. S., Ballesteros, L., & Connell, M. (2002) "Improving Stemming for Arabic Information Retrieval: Light Stemming and Cooccurrence Analysis", In *Proceedings of the 25th Annual International Conference on Research and Development in Information Retrieval (SIGIR 2002)*, Tampere, Finland, pp. 275-282, 2002. 10.1145/564376.564425

Ludovic Jean-Louis. (2011). Approches supervisées et faiblement supervisées pour l'extraction d'événements et le peuplement de bases de connaissances. Thèse, Université Paris Sud - Paris XI, Decembre 2011.

Mohit, B. N. S., Bhowmick, R., Oflazer, K., & Smith, N. A. (2012). Recall-Oriented Learning of Named Entities in Arabic Wikipedia. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, Avignon, France.

Naughton, M., Stokes, N., & Carthy, J. Investigating Techniques for Sentence-Level Event Classification, (2008) in *Proceedings of the 22nd International Conference on Computational Linguistics*, pp. 617-62.

Nezda, L., Hickl, A., Lehmann, J., & Fayyaz, S. (2006). What in the world is a Shahab? Wide coverage named entity recognition for Arabic. In: Proccedings of LREC, Genoa, pp 41–46

Oudah, M., & Shaalan, K. (2012). A pipeline Arabic named entity recognition using a hybrid approach. In: Proceedings of COLING 2012, Mumbai. The COLING 2012 Organizing Committee, pp 2159–2176

Ouersighni, R. (1998) "An approach for the conception of an arabic parser based on affix grammars over finite lattice", ICEMCO 98, Cambridge, 1998.

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Ouersighni R.(2000) "A major offshoot of the DIINAR-MBC project:AraParse, a morphosyntactic analyzer for unvowelled Arabic texts", ENSSIB, 2000.

Ouersighni, R. (2002) "Analyse syntaxique robuste de la langue arabe ",Université Lyon 2, thèse sous la direction de M.

Parent, G., Gagnon, M., & Muller, P. (2008). Annotation d'expressions temporelles et d'événements en français TALN 2008, Avignon, 9–13 juin 2008.

Piskorski, J., Tanev, H., & Wennerberg, P. (2007). *Extracting Violent Events From On Line News for Ontology Population* (pp. 287–300). SpringerVerlag Berlin Heidelberg. doi:10.1007/978-3-540-72035-5_22

Roche, M., Heitz, T., Matte-Tailliez, O.,Kodratoff, Y., EXIT(2004). Un système itératif pour l'extraction de la terminologie du domaine à partir de corpus spécialisés, 2004.

Rosa, S., & Sagot, B. (2010). Resources for named entity recognition and resolution in news wires. *Proceedings of the LREC 2010 Workshop on Resources and Evaluation for Identity Matching, Entity Resolution and Entity Management*, Malta.

Sangeetha S., Takur R., and Arock M (2010a) Domain Independent Event Extraction System Using Text Meaning Representation Adopted for Semantic Web, Publisher International Journal of Computer Information Systems and Industrial Management Applications(IJCISI.M), vol.2, pp.252-261,.

Sangeetha, S., Takur, R., & Arock, M. (2010b). *Event Detection Using Lexical Chain* (pp. 314–316). Springer-Verlag Berlin Heidelberg.

Saurí, R., Knippen, R., Verhagen, M., & Pustejovsky, J. (2005). Evita: A Robust Event Recognizer for QA Systems. In *Proceedings of HLT/EMNLP 2005*, p. 700–707. 10.3115/1220575.1220663

Roser Sauri, Jessica Littman, Bob Knippen, Robert Gaizauskas, Andrea Setzer, and James Pustejovsky(2006).TimeML Annotation Guidelines Version 1.2. January 31, 2006

Soraya Zaidi–Ayad (2013). Une plateforme pour la construction d'ontologie en arabe: Extraction des termes et des relations à partir de textes (Application sur le Saint Coran) These, Université Badij Mokhtar, Annaba 2013.

Tuerlinckx, L. (2004). "La lemmatisation de l'arabe non classique ", JADT 2004: 7es Journées internationales d'Analyse statistique des Données Textuelles, 2004.

Vargas-Vera, M., & Celjuska, D. (2004), "Event Recognition on News Stories and SemiAutomatic Population of an Ontology," in *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 615-618 10.1109/WI.2004.10148

Verhagen, M., Mani, I., Saurí, R., Knippen, R., Littman, J., & Pustejovsky, J. (2005). Automating Temporal Annotation with TARSQI. In *Proceedings of the ACL 2005* 10.3115/1225753.1225774

Wei, C. P., & Lee, Y. H. (2004). Event detection from Online News Documents for Supporting Environmental Scanning. *Decision Support Systems*, *36*(4), 385–401. doi:10.1016/S0167-9236(03)00028-9

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Wolodja, W., Knopp, J., Silberer, C., & Hartung, M. (2008). Building a Multilingual Lexical Resource for Named Entity Disambiguation, Translation and Transliteration. *Proceedings of LREC 2006*, Genoa, Italy

Xu, F., Uszkoreit, H., & Li, H. (2006)., "Automatic Event and Relation Detection with Seeds of Varying Complexity," in *Proceedings of the AAAI 2006 Workshop Event Extraction and Synthesis*

Yakushiji, A., Tateisi, Y., & Miyao, Y. (2001): Event Extraction from Biomedical Papers using a Full Parser. In: 6th Paci_c Symposium on Biocomputing (PSB 2001). pp. 408-419 (2001)

Zaghouani, W. (2012). RENAR: A rule-based Arabic named entity recognition system. [TALIP]. *ACM Transactions on Asian Language Information Processing*, *11*(1), 1–13. doi:10.1145/2090176.2090178

ENDNOTES

1 http://gate.ac.uk/
2 Arabic Wordnet http://sourceforge.net/projects/awnbrowser/
3 http://gate.ac.uk/sale/tao/splitch8.html#chap:jape

This research was previously published in the International Journal of Information Retrieval Research (IJIRR), 6(1); edited by Zhongyu (Joan) Lu; pages 36-51; copyright year 2016 by IGI Publishing (an imprint of IGI Global).

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Index

3D Virtual Worlds (3DVW) 916, 1325, 1408-1411, 1419, 1434
3LNI 746, 758-761, 763-765, 769

A

AAL 1652-1656, 1663
Abstractive Summarization 637
Actor-Network Theory 1286-1287
Aerospace 112-115, 128, 137
Affective Computing 282, 1115, 1130, 1199-1202, 1205, 1207, 1215
affective technology 1199
agent-based computing 84-86
agent-based simulation 1058, 1060, 1084
Agents Based on Goals 1539, 1543, 1562
Agreement Technologies 464-467, 477, 484
AI Confinement Problem 1235, 1260, 1266
AI-Box 1250, 1255
AIML 912, 918-919, 925-926, 933, 1410-1411, 1420-1421, 1429, 1434
Ambient Assisted Living 1652-1653, 1655
Ambient Intelligence 1655
AmI 1655, 1663
Analyzing Risk 537, 1027, 1057, 1379-1380, 1395, 1402, 1407
Animation 496, 563, 745, 925, 1060, 1080-1081, 1084, 1204, 1272, 1284, 1317, 1324, 1331-1332, 1459-1461, 1463, 1466-1467, 1469-1470, 1472-1479
anomaly detection 112, 114-115, 120, 123-124, 132, 138
Ant Colony 154, 160, 163, 178, 335, 724-725
Antirheumatic Drugs 1536, 1563
apartheid 1214, 1223
Arabic NLP 1687
archetypal representations 343, 345, 348, 354-355, 359-360, 369-371
archetype 355, 359, 361, 370
Artificial General Intelligence 486-487, 1299, 1603
Artificial Intelligence 1, 3, 5, 9, 23-25, 29, 50-53, 65-66, 68, 84, 112, 114-115, 119-120, 123-124, 138-139, 141-142, 146-147, 155-156, 163, 168, 184-185, 197, 207-209, 222, 234, 266-267, 269, 294, 322-325, 328-329, 337, 351, 353, 362-363, 370-371, 375, 397, 401, 405-406, 410, 440, 457, 459-460, 464, 468, 473-474, 477, 497, 509-512, 517, 530-532, 537, 539-541, 544-545, 549-550, 555-556, 561-562, 580, 584-586, 592, 600, 609, 611-613, 658, 660, 669, 671, 708-709, 715, 722, 724, 744, 771, 793, 796, 803, 810, 817, 860-861, 864-866, 872, 916-920, 933, 957-958, 960, 976-978, 982-985, 987, 1010, 1026-1028, 1051-1052, 1057-1059, 1149, 1166, 1183-1185, 1201, 1208, 1215-1216, 1220, 1225, 1232, 1235-1236, 1265, 1274, 1280-1282, 1284, 1286-1287, 1290, 1298-1299, 1301, 1304, 1314-1315, 1329-1330, 1349, 1355, 1357, 1402-1403, 1407, 1410-1411, 1420-1421, 1434, 1468, 1470, 1479-1482, 1495-1496, 1505, 1507-1510, 1517, 1519-1521, 1525, 1529, 1533, 1538, 1540, 1543, 1557, 1567, 1582-1584, 1587, 1601-1603, 1629, 1652-1654, 1656, 1663, 1668, 1676, 1681, 1684
Artificial Neural Networks 26, 29, 53, 59, 68, 147-148, 150, 161, 440-441, 472, 548, 697, 1149, 1204, 1498-1499, 1505, 1527, 1627, 1629
aspect based sentiment analysis 31, 33, 35, 38, 40-41, 45
aspect model 843-847, 850
Assamese 1139-1141, 1145-1147
Assisted Hybrid Structured 3D Virtual Environment (3DVW) 1434
Authoring Tool 292-293, 562, 579
automatic emotion classification 1116, 1121, 1134
Automatic evaluation 625, 772, 787
automatic text summarization 619-620, 623, 625
autonomous planning and scheduling 248-249
Avatars 351, 370-371, 487, 566-567, 578, 918, 1180, 1259, 1412, 1459-1466, 1470, 1479-1480, 1483

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

B

Back Propagation Algorithm 1017, 1020, 1022, 1158
Basle accord 49, 51, 55-56
Bayes Classifier 47, 957-960, 970, 972, 974
Bayesian network system 28
Bayesian Networks (BN) 28, 47, 291, 372, 374-378, 380-382, 385, 389-390, 393, 545, 691-692, 1184-1185, 1187, 1192, 1195, 1623, 1627
BDI 4-5, 11, 87, 101, 103-104, 1652, 1654-1658, 1660, 1663
behavior generation 562
behavioral biometrics 112, 115, 125, 136
Belief-Desire-Intention 87, 101, 1652, 1654
Big Brother 1292
Big Data 77, 285, 458, 619, 698, 939-941, 1490, 1495-1496, 1501, 1564-1568, 1570-1571, 1574, 1579-1584, 1588, 1590-1592, 1594, 1597-1606, 1676
Biological therapy 1536, 1563
Biometrics 112, 115, 125, 136, 138
biomimetics 1507-1508
BIS tagset 1139
Bostrom 1217, 1225, 1232-1236, 1249, 1251, 1253, 1257, 1260, 1266, 1298, 1311, 1368-1370, 1375, 1571, 1587
Bot 457, 486, 507, 710, 925, 1422
Broad Agency Announcement (BAA) 1585, 1618

C

Call For Paper 394, 397
capability control 1238-1239, 1241, 1265
Cellular Automaton 1618
Chaos Theory 325, 1674
Chaotic Crossover 335, 342
chaotic maps 325, 337, 342
Chaotic Mutation 336, 342
Chaotic pseudo-random number generators (CPRNG) 325, 328, 342
chatterbots 918, 923, 925, 938
Chunk 491, 507, 775, 794, 824, 999
Classification Theory 959, 982
clinical semiology 1533, 1537, 1540, 1552, 1557, 1563
Cloud Computing 143, 158, 322-323, 329, 1580
Cloud Storage 1573, 1618
Clustering 126, 128, 138, 148, 641, 700, 728, 818, 822, 855, 896, 1098, 1127, 1524-1526, 1623, 1631-1632, 1636, 1686
Co-Creation 748, 759, 765, 769, 916, 1165
codesign 872
Cognitive Architecture 485-489, 491-492, 495, 499-500, 507, 746, 791, 793, 796-798, 800, 803, 1233
cognitive informatics 985, 987
Cognitive Intelligence 725, 739, 1203
cognitive modeling 119
cognitive research 344, 348, 350, 354, 356, 361
Cognitive Systems 248, 746, 793, 798, 802-803
Collaborative Filtering 638, 641, 643, 688-690, 695-702, 707, 1622-1624, 1626-1628, 1630
Collective Intelligence 861, 864-865, 1586, 1675, 1679, 1681
Collective Narrative Generation 748
collective personae 344-345, 347, 354, 358-359, 362, 370
compartmental model 939, 946
complex decision problems 958, 962, 977-978
Complexity Theory 1618
Computational Creativity 765, 769
Computational Linguistics 31-32, 43, 817, 1010, 1085
computational models 798, 1081-1082
Computer Networks 913-914, 916, 920, 922, 925-926, 931, 933, 938, 1257, 1603
Computer Science 23-24, 84, 120, 138, 356, 361, 370-371, 397, 399, 465, 509-511, 528, 530, 584-585, 587, 592, 597, 602, 659, 723, 817, 1009-1010, 1027-1028, 1051-1052, 1085, 1096, 1149, 1380, 1557, 1575, 1656
computer-interpretable clinical guidelines 659
Conceptual Graphs 770-773, 776-778, 780
Conceptual Model 3, 56, 682, 1471
Conceptualization 488, 570, 800, 864-865, 878-881, 887-889, 895-897, 1409, 1493
Conditional Random Field (CRF) 998, 1009
Configuration Management 112-115, 117, 127-128, 137-138
Consciousness 161, 344, 349, 355, 357-358, 360, 363, 370, 791-792, 796-797, 800-801, 803, 976, 1249, 1299, 1308-1310, 1507-1514, 1517-1521, 1525, 1567, 1582-1585, 1587, 1592-1593, 1596-1597, 1601-1602
constraint propagation 172, 250
content filtering 1622, 1624-1626
content generation 1350, 1352, 1355-1356, 1360
Content-Based Filtering 689, 695-696, 707, 1624-1625
context-aware computing 914
Contour Modeling 538, 594, 602, 607, 1041, 1407
Conversational Interaction 918, 938, 1420, 1425-1426, 1434
Convolutional Neural Network 446, 457, 1149-1151, 1154-1155, 1158-1159
copyright 539, 541-542, 546-547
CPRNG Tweak Function 342
Creative Works 1058, 1084
credit rating migration 65-66, 68, 71, 77

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

crime linkage 375, 379, 381, 389, 393
crime mining 389-390
Crime Pattern 393
Crime Tactical Analysis 393
criminal profiling 379-380, 389, 393
cultural applications 1279, 1352, 1360
cultural heritage 1314-1316, 1333, 1335-1337, 1348-1350, 1352-1357, 1359
curriculum sequencing 722-725, 727, 731, 738-739
curve extrapolation 1403
Curve Interpolation 509-512, 517, 522, 527-531, 538, 585-587, 589, 591-593, 597, 599-600, 607, 609-611, 618, 1026-1028, 1030, 1032, 1040-1041, 1049-1051, 1057, 1379-1380, 1403, 1407
cyborganization 1669, 1674, 1681, 1684
Cytopathology 268

D

Data Extrapolation 510, 528-529, 538, 618, 1032, 1057, 1379-1381, 1395, 1407
data management 306
Data Mining 31-32, 67, 70, 112, 115, 120-122, 124, 138, 285, 291, 379, 401, 405, 689-690, 707, 724, 1216-1217, 1356, 1490, 1631, 1633
Data Modeling 512, 531-532, 584, 586, 604, 606, 612, 1027, 1036, 1379
Davos Forum 1304
decision maker 958, 960-963, 967, 972, 977-978, 982
Decision Making 2, 5-6, 11, 13, 17, 23-25, 54-55, 147, 149-150, 192, 222, 240, 243, 372-373, 375, 389, 465, 499-500, 509-510, 512, 527-529, 532, 538, 542-544, 546, 551, 556, 584, 586, 589, 597, 602, 607-608, 612, 658, 682, 695, 710, 809, 829, 861-862, 864, 889, 941, 949, 957-958, 960, 978, 982, 1026-1028, 1044, 1050-1052, 1057, 1217, 1220, 1294, 1307, 1379-1381, 1395, 1400-1402, 1407, 1483, 1509-1510, 1520, 1525, 1528, 1586, 1597-1598, 1622, 1629, 1631, 1640, 1675, 1681, 1684, 1687
Decision Problem 958, 960, 963, 972, 982, 1540
Decision Support 372, 375, 379, 386, 546, 659-661, 671, 682, 861, 977, 1207, 1411
decision tree classifiers 47
Declarative Knowledge 495, 507, 794, 1518
Deduction 5, 17, 187, 207, 1273, 1570, 1575, 1618
Deep Learning 216, 440-442, 445-446, 449-450, 452-453, 455-457, 459, 461, 1023, 1121, 1149, 1158, 1304-1305, 1309, 1523, 1525, 1527, 1529
digital age 1272
Digital Animation 1284
Digital Game 215, 217, 563-564
Disease Modifying Anti-Rheumatic Drugs (DMARDs) 1536, 1563
Distributed Artificial Intelligence 1, 3, 84, 234
Distributed Generation 159
Domain Knowledge 10, 304, 456, 470, 727-728, 883, 1181, 1621-1622, 1634, 1636-1637, 1640-1641
drawing operator 1070, 1072-1073, 1084
Dyadic Binary Formula 1578, 1618
Dynamic Bayesian Networks (DBNs) 393
dynamic temporal reasoning 248-249, 252

E

E-Business 878-880, 883, 888-897, 911, 1622
educational game design 216, 224
educational informatics 983-984, 986-987
ELAI 912-913, 920, 922-924, 927-928, 930-933, 938
Electronic Institutions 1408, 1411-1412, 1418
Emotional Agents 4, 1199, 1205
emotional intelligence 1199-1208
Emotions 87, 162, 282, 348, 350, 359, 564, 574, 641, 744, 808-815, 942-943, 953, 1077, 1115-1117, 1119-1122, 1129-1130, 1200-1208, 1469, 1478, 1507, 1515, 1517, 1583
Empathy 349, 1203, 1253, 1478-1479, 1517, 1529
endometrium 267, 269
enterprise search engine 302-303, 314, 317
entry relationship 838-839, 846-848, 851-852, 855
Ethics 52, 281, 284, 346, 350, 1214-1226, 1246, 1253, 1259, 1263, 1308, 1508, 1586, 1598, 1601
Evaluation Metrics 780, 782, 1003, 1698
event annotation 1692
exception treatment 666, 682
Extractive Summarization 621, 637

F

Face Detection 1152
facial animation 1204, 1459, 1469-1470, 1473
fair use 539-542, 547-549, 552, 555
fog computing 1564-1567, 1570, 1579-1580, 1584, 1591, 1599-1600, 1602-1603, 1605-1606
Formal Concept Analysis 724, 1436-1438, 1454, 1636
Formulate Goal 1563
Frey and Osborne 1306
fun space 1367-1372, 1374-1376
Fuzzy Inference 27, 818
Fuzzy Logic 27, 149-150, 197, 818, 1629, 1656
fuzzy systems 17, 24, 27, 29

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

G

Game Engine 224, 491, 708, 712-713, 715-716, 1315, 1331, 1351
game logs 1179, 1183, 1189-1190, 1194, 1196
Game-Based Learning 214-217, 220, 224, 226, 228-229, 1335
Gamification 862, 986-987, 1335-1336, 1349, 1352, 1410, 1428-1429
Gate platform 1686-1687, 1693, 1695-1696
generated scenarios 1178-1179, 1189-1190, 1195-1196
generating comics 1070
Genetic Algorithms 24, 152-153, 163, 333, 470, 477, 546, 724, 1216, 1627, 1629
gentrification 1275-1277, 1280, 1282, 1284
geometric method 250-251
goal-based agents 1557

H

hard elements 1669-1670, 1677, 1680
Health Insurance Portability and Accountability Act (HIPAA) 992, 1009
Healthcare 284, 322, 459-460, 658-659, 683, 1002, 1104, 1215-1216, 1223-1224, 1286, 1289, 1463-1468, 1481-1482, 1484, 1540, 1655
Hidden Markov Model 517, 996, 1001, 1013, 1139-1141, 1143-1144, 1478
holacracy 1669, 1673, 1678, 1680, 1684
holography 354, 357-358, 370
homo-di-fict 1272-1275, 1280, 1282, 1284
Human Computer Interaction 32, 1155, 1200, 1206, 1226, 1359
Hurwitz-Radon Matrices 523, 529, 587, 589, 1027, 1030, 1051, 1379-1381, 1388, 1403
hybrid approaches 702
hybrid collaborative filtering 696, 698, 707
Hybrid Structured 3D Virtual Environment (3DVW) 1434

I

Image Analysis 267-268, 275, 1149
IMPLEMENTATION STRATEGIES 282, 288
Industrial Revolution 139, 460, 1299-1304, 1307, 1668, 1680
inflammatory 1533-1534, 1544-1546, 1549, 1563
Information And Communication Technologies 862, 913, 983, 1162
Information Extraction 32, 150, 394-395, 398, 401, 407, 772, 840, 996, 1686-1687, 1693
Information Network 148, 159, 1287
Information Retrieval 31, 38, 304, 413, 620, 651, 693, 695, 749, 818-820, 824, 826, 839-840, 942, 1088-1089, 1096, 1110, 1140, 1438, 1446, 1622, 1686-1687
information society 743
Information Systems 11, 149, 879, 881, 1085, 1355, 1675
Inteligent Agent 938
intellectology 1367-1368, 1376
intelligence gathering 1591
Intelligent Agent 2, 4, 11, 18, 102, 175, 912-914, 917, 920-924, 926-928, 932-933, 938, 1204, 1234, 1290, 1533, 1539-1540, 1543, 1563
Intelligent Tutoring System 228-229, 281, 287, 722-723, 725, 727, 739
Intelligent Tutors 280-282, 284-286, 294, 794
Interaction Design 222, 486, 565-566, 573, 763, 1544
Interactivity 11, 347, 357, 361, 363, 370-371, 500, 562, 927, 1472
Internet Of Things 323, 1214-1215, 1217-1218, 1286-1288, 1291-1292, 1564, 1566, 1668, 1676

K

k-NN 1204, 1624, 1630
Knowledge Base 7, 9-10, 169, 197, 498, 669, 692, 694-695, 775, 785, 912-914, 918, 923, 925-927, 931, 1116, 1481, 1544-1545, 1548, 1636, 1653-1654, 1656
knowledge modelling 1350, 1352-1353, 1355-1356, 1360
Knowledge Representation 9, 100, 234, 396, 662, 690, 771, 773, 1013, 1381, 1410, 1634

L

Latent Semantic Analysis 415, 418, 772, 829, 839, 878-879, 881, 884, 895, 904, 908, 1089-1090, 1162, 1166, 1174, 1492
learning analytics 1179
Learning Outcomes 215-216, 227, 280-281, 283, 286, 294, 717
learning style 722, 724-725, 727-729, 735-736, 738
Learning Vector Quantizer 266-267
legal practice 51, 540, 556

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Level Of Expertise 283, 912-913, 920-921, 928, 930, 938
Lexical 31, 33, 37, 40, 42, 45, 620-622, 646, 812-813, 818-820, 822-825, 834, 994-998, 1001, 1006, 1087, 1116, 1120-1121, 1126, 1130, 1491, 1688, 1690-1691
longevity 1676
Luddite fallacy 1301-1302, 1305

M

Machine Learning 10, 17, 25, 37, 44-45, 51, 55, 65, 123, 146, 148, 207, 267, 379, 381, 441, 443, 446, 455, 458-459, 461, 509, 549-551, 555-556, 584, 621-622, 690, 693, 696-697, 703, 808-810, 817, 959, 976, 991, 993-994, 997-998, 1006, 1009, 1026, 1118, 1149-1150, 1159, 1208, 1216-1217, 1221, 1304, 1306, 1499, 1508, 1525-1529, 1625, 1627, 1631-1632, 1636, 1641, 1656, 1691
Machine Translation 410-414, 419, 452, 818, 820, 1140, 1686
MADM 960-964, 967, 972, 976-978, 982
management education 280-282, 284-288, 290, 293-295
MAS 2-3, 11, 16-17, 83, 86-87, 89, 95, 97, 105, 234-236, 240, 466, 473-475, 477-478, 484, 1434, 1652-1656, 1662-1663
Maximum Entropy Classifier 47
MCDM 957-960, 969, 974, 982, 1631
measurement of features 186
media dream 343, 345, 347, 349-350, 354, 356-359, 361, 363, 371
Medical Identification System 1539-1540, 1543, 1552
Medical Natural Language Processing 993, 1010
Medical Signs 1563
Medical Symptoms 1547, 1563
memory-based approaches 697
metaheuristic algorithms 139
Metaverse 343-345, 347-348, 350-351, 353-354, 356, 359, 361, 363, 369-371, 916, 924-925
Metering 149, 159
MHR Method 512-514, 516, 518, 522, 525, 529-532, 538, 586-587, 589, 591, 597, 611-612, 618, 1027-1028, 1030-1032, 1045, 1048, 1050-1051, 1057, 1379-1381, 1383, 1388, 1390-1398, 1400-1403, 1407
microgeneration 143-145
mind space 1261, 1375
Minsky 1, 24, 52, 68, 160-163, 178, 443, 495
Model-Based Approaches 697, 1623, 1631
Modus Operandi 373, 381, 393
morphing 1473, 1667, 1669, 1673, 1680-1681, 1684
morphometry 267
Motion Capture 457, 1467, 1471-1472, 1474-1475, 1477
Multi-Agent System (MAS) 1-3, 6-7, 11, 16, 83-84, 86-89, 99, 101, 105, 160-161, 163-164, 172-173, 178, 235-236, 484, 499, 700, 1330, 1409, 1411-1412, 1429, 1434, 1652
Multi-Criteria Decision Making 546, 1622, 1631, 1640
multi-layer perceptron 49, 51, 54, 66, 68-70, 544-545, 550, 552, 554-555
museum education 215

N

Naive Bayes Classifier 47
Named Entity Recognition (NER) 992-993, 996, 1010, 1140, 1690
narrative discourse 744, 749-751, 760, 769
narrative expression 755, 759-760, 769
Narrative Generation System 743-745, 763, 769
narrative structure 759-760, 765, 995
Natural Language Processing (NLP) 31-32, 38, 44, 410, 412, 440-441, 450, 489, 746, 770, 772, 774, 808-810, 817, 819-820, 842, 939-940, 942, 946, 949, 953, 993, 995, 1010, 1012-1014, 1149, 1151, 1156, 1162-1163, 1165-1166, 1305, 1350, 1490-1491, 1505, 1628, 1686, 1693
Negotiation 1-2, 6, 11, 15, 98, 464-473, 475-477, 484, 545-546, 1226, 1413, 1419, 1633, 1655
NetMedicine 1652-1653, 1663
Neural Network (NN) 17-18, 23-29, 47, 49, 52-55, 57, 59, 61, 65-70, 74-75, 77, 137, 147-148, 150, 161, 207, 267, 285, 350, 440-442, 445-446, 448-450, 452, 457, 460-461, 472, 545, 547-551, 697, 724, 774, 984, 1006, 1011-1012, 1014, 1016, 1019, 1022-1023, 1126, 1149-1151, 1153-1156, 1158-1159, 1204, 1216, 1498-1500, 1505, 1510, 1519-1521, 1527, 1625, 1627, 1629
Neural-Gaming 1505
news propagation 939-941, 943, 948
Non-Playing Character 494, 507, 914, 917, 923, 938
Nonsteroidal Anti-Inflammatory Drugs (NSAIDs) 1536, 1563
NRC Emolex 1133-1134

O

OHR Operator 514-515, 529, 538, 589, 618, 1051, 1057, 1389, 1396, 1407
online Internet encyclopedia 838

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Open Multi-Agent System 484, 700
Opinion Mining 32, 34, 44, 638, 641, 646, 809-810, 817, 1130
optical output 185
Oracle AI 1234, 1265-1266
Organization Centered Multi-Agent System (OCMAS) 1434
organizations in the future 1667, 1669-1670, 1680
original picture 1080, 1084
Oxford thesaurus 1124

P

Part of Speech Tagger 1139
participatory democracy 862, 871
participatory prototyping 860, 865, 871
Particle Swarm Optimization 150, 235, 332
pattern classification 186, 443
Pattern Recognition 17, 114, 120, 138, 184-185, 207, 210, 269, 398, 509-511, 517, 530-531, 585-586, 592, 599, 609, 611, 1009, 1026, 1028, 1051, 1151, 1158-1159, 1204, 1306, 1574, 1578
pedagogical agents 217, 220
Personalised Virtual Museums 1348, 1350, 1352, 1356
photodiode 185, 198, 202, 204-205, 207-208
platoon control 232-235, 240-243
Point Interpolation 529, 1027, 1051, 1402
Pos Tag 39, 417, 772, 997, 1140, 1144
pre-knowledge level 727-729, 731, 733, 735, 739
Pre-Processing 307, 398, 414, 441, 550, 552, 554, 620, 622-623, 625-626, 631, 813, 884, 909, 998-999, 1125-1126, 1625, 1634, 1636
probabilistic reasoning 372-374, 385, 389, 393
probability formula 958, 960, 963-964, 966
Problem Solving 3, 7, 86, 99, 160-162, 164, 166, 171-173, 178-179, 285, 294, 747, 771, 840, 1201, 1335, 1349, 1598
Product Reviews 31, 41, 815
Project Management 288, 879, 883, 897, 984
project reports 878-880, 883, 896, 908
protected health information (PHI) 991-992
Pseudo-Random Number Generators (PRNGs) 324, 342
psychecology 345, 349, 356, 358, 361, 371
Psychecology Video Games 371
Psychological Functions 348, 371
Psychology 287, 343-344, 356-357, 361-363, 381, 456, 545, 723, 792, 794, 808-809, 984, 1085, 1201, 1220, 1225, 1249, 1252, 1255, 1259, 1538, 1593, 1601
public opinion 688, 695, 862, 939, 949

Q

Quantum Computing 1571-1572, 1602, 1618

R

racism 1215-1216, 1218, 1310
random number generation 323-325, 342
ranking methods 961, 978, 982
realistic simulations 1314, 1316, 1329, 1337, 1349
Recommendation System 307, 641, 651, 655, 696, 1623, 1625, 1627, 1631, 1635, 1640
Recommender Systems 639, 643, 688-690, 693-698, 702, 707, 1410, 1428-1429, 1434, 1621-1622, 1624-1633, 1635, 1640-1641
Recursion 355, 360, 371, 800-801, 1498, 1523, 1527, 1575, 1577, 1603, 1619
Reductionism (Cartesianism) 1619
Reinforcement Learning 10, 149, 176, 489-491, 499, 1241, 1243, 1355, 1525
relational data 690-692, 700-702
rheumatic diseases 1533-1534, 1536-1537, 1540-1543, 1545, 1547, 1552, 1554-1556, 1563
Rings of Power 1293
RITA 1461-1463, 1467-1468, 1479-1481
Robotics 202, 351-352, 362-363, 370-371, 441, 459, 794, 1207, 1220, 1286, 1299, 1304-1306, 1471, 1507-1509, 1519-1521, 1538, 1544, 1668
Rough sets 1436, 1438-1440, 1442-1443
Rule-Based 47, 197, 292, 474, 488-490, 492, 545, 728, 993-994, 997-998, 1012, 1120, 1238, 1241, 1244, 1623, 1626-1627, 1690
Rule-Based Classifiers 47
Rules of Behavior 1547

S

Safety Critical Systems 112, 115-116, 138
scanning technology 187
Scheduling 139, 141, 144-145, 152-156, 248-252, 255-256, 258, 260-262, 470, 510, 528, 564, 1027, 1157, 1402-1403, 1540
Science Fiction 371, 441, 1251, 1286-1287, 1292, 1294, 1299, 1480-1481, 1483, 1602
Search Space 450, 471-472, 1563
Second Life 489, 574-575, 577-580, 916, 1255, 1315, 1411, 1503, 1583, 1591
Second Machine Age 1490, 1674, 1680, 1684
Second Mind 561-562, 572, 574-580
Second Order Cybernetics 1603-1604, 1619
self-sustaining 791-792, 799-803

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

Semantic Analysis 45, 306, 317, 410, 413, 415, 418, 638-639, 642, 644, 646, 772, 829, 838-840, 842-843, 847-848, 855, 878-879, 881, 884, 895, 904, 908, 1011-1013, 1089-1090, 1162, 1166, 1174-1175, 1492, 1693
semantic analysis model 842, 855
Semantic Knowledge 745, 1107, 1633, 1640
semantic networks 565, 771, 1085, 1087, 1098, 1106, 1109-1110
Semantic Relatedness 1085-1087, 1089-1093, 1096-1099, 1102, 1104-1107, 1109-1111, 1115-1116, 1127, 1133-1134, 1443
Semantic Search 305, 318, 1436-1440, 1442-1443, 1445, 1454
Semantic Similarity 579, 622, 818-824, 834, 1086-1089, 1115-1116, 1121, 1125, 1127, 1130, 1133, 1443, 1446
Semantic Web 95, 304, 640, 1437, 1622, 1634-1635, 1637, 1640
semantic web usage mining 1622, 1635, 1637, 1640
semantics of answers 773
Sentence Features 637
sentence scoring 619-620, 622-623, 625-626
Sentiment Analysis 31-35, 37-38, 40-41, 43-45, 644, 808-809, 939-940, 942, 946, 949, 953, 1115, 1117, 1120, 1156
Sentiment Score 31, 33, 646, 648
Sentiwordnet 31, 33, 40, 42, 45, 646
Serial Crime 393
serial offender 382-383, 386, 389, 393
serious game architecture 1178-1179, 1183-1184
Serious Games 215, 357, 1178-1184, 1195-1196, 1315, 1326, 1336, 1349, 1352, 1355, 1357
service providers 71, 148, 159, 466
Shape Representation 509-511, 529-531, 585-587, 589, 592, 597, 599, 601, 609, 611, 618, 1026-1028, 1051-1052
signal processing 1323, 1383
SimCoach 1461-1462
similarity reasoning 1436-1437, 1439-1440, 1442, 1454
SimSensei 1466, 1468, 1481
Sina Weibo microblogging 939-941, 946, 949
Singularity 352, 1233, 1235, 1254-1255, 1263, 1266, 1299, 1305, 1584, 1603-1604
Slow-Acting Antirheumatic Drugs (SAARDs) 1536, 1563
smart city 1164, 1223-1224, 1291, 1574
smart grid devices 152, 159
social application 743
social brain 1596, 1600-1603, 1605
Social Network Analysis 16, 642, 646, 689, 1094, 1096, 1110
Social Network Mining (SNM) 689, 707
Social Networks 362, 655, 688-689, 693, 700-701, 1090, 1096, 1115, 1437, 1454
social recommendation 642
social robotics 1507, 1509
society of agents 160-161, 164, 166, 168, 178, 466
Society of Mind 160-162, 165, 178-179
socio-technical systems 1217
soft elements 1669-1670, 1680
soft law 49-50, 52, 55, 59, 61
Software Agent 84-85, 87, 162, 1533-1534, 1538, 1558, 1635
Software Agents 84-85, 87, 89, 162, 169, 371, 464-465, 794, 1181, 1204, 1409, 1412, 1414, 1417, 1434, 1543-1544, 1547-1549, 1557-1558, 1563, 1653
South Park 1272-1278, 1280-1282
Sparsity 643, 688, 697-698, 702-703, 1624-1626, 1629
speech generation 1460, 1467, 1477-1478
Speech Recognition 375, 452, 992, 1001, 1140, 1149, 1155, 1467, 1490, 1495-1497, 1656
statistical approach 1089
Stemming 45, 317, 620, 626, 631, 637, 813, 842, 1492, 1494-1496
stochastic information 962, 964, 976, 982
stopword 620, 637, 909
Story Vector 1065, 1084
story world 755, 769
storyboard 1058, 1080-1081, 1084
Strong AI 488, 1298-1299, 1603
subjective answers 770-772
Summary Generation 620, 623, 625-626, 631, 1011, 1492
superintelligence 1232-1234, 1252, 1258-1259, 1262, 1265, 1298, 1603
Support Vector Machine (SVM) 47, 185, 207, 395, 822, 996, 1010, 1080, 1204
support vector machine regression 185
surface computing tables 220
suspect modelling 372, 374, 385
syneoid 1299, 1308, 1310

T

technological unemployment 1301-1302, 1306, 1309
text analysis 32, 304, 307, 309-311
Text Classification 31, 38, 699, 817
Text Mining 375, 394-395, 398, 401, 407, 772, 884, 1010, 1491-1492, 1631, 1686, 1700

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use

text summarization 619-623, 625, 637
Third Industrial Age 1685
transhumanism 1602
Transient Competitive Advantage 1671, 1685
True-Random Number Generators (TRNGs) 342
Tukey Criteria 132, 134

U

Ubiquitous Computing 914-915, 920, 931, 938, 1288, 1292
ubiquitous learning 915
uploading 1368, 1640
user-awareness 303-305, 317
user-ratings 696-698, 700, 703

V

Value Anticipation 510, 512, 527-529, 532, 586, 589, 597, 602, 607, 612, 618, 1028, 1050-1051, 1057, 1379-1380, 1407
Vectorization 1090, 1127
Verbal Apraxia 1505
Verbal Dyspraxia 1505
Victimization Prediction 393
Video Games 350-351, 354, 357-359, 361-362, 364, 370-371, 488, 561, 745, 930, 986-987, 1178, 1180-1182, 1501
Virtual Agent 485-489, 491-493, 496, 498-500, 507, 1330, 1466
Virtual Carer 1652-1654, 1656-1663
Virtual Environments 220, 485-486, 491-492, 496, 499, 912-914, 916, 918-921, 933, 1314, 1316, 1329-1331, 1337, 1348, 1350, 1352-1353, 1355-1356, 1359-1360, 1410-1411, 1421-1424, 1462
virtual museums 1314-1316, 1330, 1335, 1348-1356, 1359-1360
Virtual Organization 484
Virtual Reality 284, 343, 350-351, 354, 358, 363, 492, 1259, 1314, 1332, 1348-1350, 1352-1353, 1356, 1583
Virtual Training 350, 1408
Virtual Worlds 353, 371, 487, 561-562, 913-914, 916, 918-920, 932, 938, 1233, 1329, 1336, 1352, 1357-1358, 1408-1412, 1419-1420, 1583
visual interface 760-761, 763, 765
visual realism 1316, 1324, 1472
Visualization Technique 1619
Voice Recognition 440-441, 445, 452, 457, 1481, 1505

W

Web Search 651, 819-820, 824, 1436-1437
web usage mining 1622, 1626, 1631-1637, 1640
Wikipedia 825, 838-840, 842, 846, 848-852, 855, 1090, 1126, 1288-1290, 1292, 1436-1437, 1440, 1442-1443, 1454, 1566, 1570-1571, 1573-1574, 1578, 1582, 1590, 1593, 1597-1598, 1603-1604, 1606, 1675, 1690
Wikipedia article 840, 851
Word Sense Disambiguation 38, 44, 412, 415, 418-419, 818, 820, 1015, 1022
Wordnet 37, 410, 412, 415-416, 418, 622, 646, 772, 812, 819-820, 822-823, 825, 1086-1087, 1090, 1098, 1109, 1120, 1122-1124, 1126, 1437, 1690-1692

EBSCOhost - printed on 2/9/2023 6:08 AM via . All use subject to https://www.ebsco.com/terms-of-use