Rossitza Setchi
Ivan Jordanov
Robert J. Howlett
Lakhmi C. Jain (Eds.)

LNAI 6279

Knowledge-Based and Intelligent Information and Engineering Systems

14th International Conference, KES 2010
Cardiff, UK, September 2010
Proceedings, Part IV

Part IV

Springer

Lecture Notes in Artificial Intelligence 6279

Edited by R. Goebel, J. Siekmann, and W. Wahlster

Subseries of Lecture Notes in Computer Science

Rossitza Setchi   Ivan Jordanov
Robert J. Howlett   Lakhmi C. Jain (Eds.)

Knowledge-Based and Intelligent Information and Engineering Systems

14th International Conference, KES 2010
Cardiff, UK, September 8-10, 2010
Proceedings, Part IV

Springer

Series Editors

Randy Goebel, University of Alberta, Edmonton, Canada
Jörg Siekmann, University of Saarland, Saarbrücken, Germany
Wolfgang Wahlster, DFKI and University of Saarland, Saarbrücken, Germany

Volume Editors

Rossitza Setchi
Cardiff University, School of Engineering
The Parade, Cardiff CF24 3AA, UK
E-mail: Setchi@cf.ac.uk

Ivan Jordanov
University of Portsmouth, Dept. of Computer Science and Software Engineering
Buckingham Building, Lion Terrace, Portsmouth, PO1 3HE, UK
E-mail: Ivan.Jordanov@port.ac.uk

Robert J. Howlett
KES International
145-157 St. John Street, London EC1V 4PY, UK
E-mail: rjhowlett@kesinternational.org

Lakhmi C. Jain
University of South Australia, School of Electrical and Information Engineering
Adelaide, Mawson Lakes Campus, SA 5095, Australia
E-mail: Lakhmi.Jain@unisa.edu.au

Library of Congress Control Number: 2010932879

CR Subject Classification (1998): I.2, H.4, H.3, I.4, H.5, I.5

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN 0302-9743
ISBN-10 3-642-15383-6 Springer Berlin Heidelberg New York
ISBN-13 978-3-642-15383-9 Springer Berlin Heidelberg New York

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper 06/3180

Preface

The 14th International Conference on Knowledge-Based and Intelligent Information and Engineering Systems was held during September 8–10, 2010 in Cardiff, UK. The conference was organized by the School of Engineering at Cardiff University, UK and KES International.

KES2010 provided an international scientific forum for the presentation of the results of high-quality research on a broad range of intelligent systems topics. The conference attracted over 360 submissions from 42 countries and 6 continents: Argentina, Australia, Belgium, Brazil, Bulgaria, Canada, Chile, China, Croatia, Czech Republic, Denmark, Finland, France, Germany, Greece, Hong Kong ROC, Hungary, India, Iran, Ireland, Israel, Italy, Japan, Korea, Malaysia, Mexico, The Netherlands, New Zealand, Pakistan, Poland, Romania, Singapore, Slovenia, Spain, Sweden, Syria, Taiwan, Tunisia, Turkey, UK, USA and Vietnam.

The conference consisted of 6 keynote talks, 11 general tracks and 29 invited sessions and workshops, on the applications and theory of intelligent systems and related areas. The distinguished keynote speakers were Christopher Bishop, UK, Nikola Kasabov, New Zealand, Saeid Nahavandi, Australia, Tetsuo Sawaragi, Japan, Yuzuru Tanaka, Japan and Roger Whitaker, UK.

Over 240 oral and poster presentations provided excellent opportunities for the presentation of interesting new research results and discussion about them, leading to knowledge transfer and generation of new ideas.

Extended versions of selected papers were considered for publication in the *International Journal of Knowledge-Based and Intelligent Engineering Systems, Engineering Applications of Artificial Intelligence, Journal of Intelligent Manufacturing*, and *Neural Computing and Applications*.

We would like to acknowledge the contribution of the Track Chairs, Invited Sessions Chairs, all members of the Program Committee and external reviewers for coordinating and monitoring the review process. We are grateful to the editorial team of Springer led by Alfred Hofmann. Our sincere gratitude goes to all participants and the authors of the submitted papers.

September 2010

Rossitza Setchi
Ivan Jordanov
Robert J. Howlett
Lakhmi C. Jain

Organization

KES 2010 was hosted and organized by the School of Engineering at Cardiff University, UK and KES International. The conference was held at the Mercure Holland House Hotel, September 8–10, 2010.

Conference Committee

General Chair	Rossi Setchi, Cardiff University, UK
Conference Co-chair	Lakhmi C. Jain, University of South Australia, Australia
Executive Chair	Robert J. Howlett, University of Brighton, UK
Chair of the Organizing Committee	Y. Hicks, Cardiff University, UK
Program Chair	I. Jordanov, University of Portsmouth, UK

Organizing Committee

KES Operations Manager	Peter Cushion, KES International
Publicity Chairs	D. Todorov, Cardiff University, UK
	Yu-kun Lai, Cardiff University, UK
KES Systems Support	Shaun Lee, KES International
Members	Engku Fadzli, Cardiff University, UK
	Lei Shi, Cardiff University, UK
	Nedyalko Petrov, Portsmouth University, UK
	Panagiotis Loukakos, Cardiff University, UK

Track Chairs

Bruno Apolloni	University of Milan, Italy
Bojana Dalbelo Basic	University of Zagreb, Croatia
Floriana Esposito	University of Bari, Italy
Anne Hakansson	Stockholm University, Sweden
Ron Hartung	Franklyn University, USA
Honghai Liu	University of Portsmouth, UK
Ngoc Thanh Nguyen	Wroclaw University of Technology, Poland
Andreas Nuernberger	University of Magdeburg, Germany
Bernd Reusch	University of Dortmund, Germany
Tuan Pham	University of New South Wales, Australia
Toyohide Watanabe	Nagoya University, Japan

Invited Sessions Chairs

3D Visualization of Natural Language	Minhua Eunice Ma, University of Derby, UK Nikolaos Antonopoulos, University of Derby, UK Bob Coyne, Columbia University, USA
Intelligent Data Processing in Process Systems and Plants	Kazuhiro Takeda, Shizuoka University, Japan Takashi Hamaguchi, Nagoya Institute of Technology, Japan
A Meta-Heuristic Approach to Management Engineering	Junzo Watada, Waseda University, Japan Taki Kanda, Bunri University of Hospitality, Japan Huey-Ming Lee, Chinese Culture University, Taiwan Lily Lin, China University of Technology, Taiwan
Knowledge Engineering and Smart Systems	Cesar Sanin, University of Newcastle, Australia
Skill Acquisition and Ubiquitous Human–Computer Interaction	Hirokazu Taki, Wakayama University, Japan Masato Soga, Wakayama University, Japan
Application of Knowledge Models in Healthcare	István Vassányi, University of Pannonia, Hungary György Surján, National Institute for Strategic Health Research, Hungary
Knowledge Environment for Supporting Creative Learning	Toyohide Watanabe, Nagoya University, Japan Tomoko Kojiri, Nagoya University, Japan
ICT in Innovation and Creativity	Toyohide Watanabe, Nagoya University, Japan Taketoshi Ushiama, Kyushu University, Japan
Intelligent Support for Designing Social Information Infrastructure	Toyohide Watanabe, Nagoya University, Japan Naoto Mukai, Tokyo University of Science, Japan
Intelligent Systems in Ambient-Assisted Living Environments	Antonio F. Gómez-Skarmeta, Universidad de Murcia, Spain Juan A. Botía, Universidad de Murcia, Spain
Knowledge-Based Systems for e-Business	Kazuhiko Tsuda, University of Tsukuba, Japan Nobuo Suzuki, KDDI Corporation, Japan
Quality Assurance and Intelligent Web-Based Information Technology	Anastasia N. Kastania, Athens University of Economics and Business, Greece Stelios Zimeras, University of the Aegean, Greece
Knowledge-Based Interface Systems	Yuji Iwahori, Chubu University, Japan Naohiro Ishii, Aichi Institute of Technology, Japan Yoshinori Adachi, Chubu University, Japan Nobuhiro Inuzuka, Nagoya Institute of Technology, Japan
Reasoning-Based Intelligent Systems	Kazumi Nakamatsu, University of Hyogo, Japan Jair Minoro Abe, University of Sao Paulo, Brazil
Data Mining and Service Science for Innovation	Katsutoshi Yada, Kansai University, Japan Takahira Yamaguchi, Keio University, Japan Maria Alessandra Torsello, University of Bari, Italy

Web 2.0: Opportunities and Challenges for Social Recommender Systems	Jose J. Pazos-Arias, University of Vigo, Spain Ana Fernandez-Vilas, University of Vigo, Spain
Innovations in Chance Discovery	Akinori Abe, University of Tokyo, Japan
Personalization of Web Contents and Services	In-Young Ko, Korea Advanced Institute of Science and Technology (KAIST), Korea Juan D. Velásquez, University of Chile, Chile
Advanced Knowledge-Based Systems	Alfredo Cuzzocrea, ICAR-CNR and University of Calabria, Italy
Knowledge-Based Creativity Support Systems	Susumu Kunifuji, Jaist, Japan Kazuo Misue, University of Tsukuba, Japan Hidehiko Hayashi, Naruto University of Education, Japan Motoki Miura, Kyushu Institute of Technology, Japan Toyohisa Nakada, Niigata University of International and Information Studies, Japan Tessai Hayama, JAIST, Japan
Intelligent Network and Service	Jun Munemori, Wakayama University, Japan Takaya Yuizono, JAIST, Japan
Real-World Data Mining and Digital Intelligence	Rashid Mehmood, Swansea School of Engineering, UK Omer F. Rana, Cardiff University, UK Ziad Salem, Aleppo University, Syria
Advanced Design Techniques for Adaptive Systems	Sorin Hintea, Technical University of Cluj-Napoca, Romania Hernando Fernández-Canque, Glasgow Caledonian University, UK Gabriel Oltean, Technical University of Cluj-Napoca, Romania
Soft Computing Techniques and Their Intelligent Utilizations Toward Gaming, Robotics, Stock Markets etc.	Norio Baba, Osaka Kyoiku University, Japan
Methods and Techniques of Artificial and Computational Intelligence in Engineering Design	Argyris Dentsoras, University of Patras, Greece Nikos Aspragathos, University of Patras, Greece Vassilis Moulianitis, University of the Aegean, Greece
Philosophical and Methodological Aspects of Reasoning and Decision Making	Vesa A. Niskanen, University of Helsinki, Finland

Semantic Technologies for Knowledge Workers	Andreas Dengel, German Research Center for Arificial Intelligence (DFKI), Germany Ansgar Bernardi, German Research Center for Arificial Intelligence (DFKI), Germany
Tools and Techniques for Effective Creation and Exploitation of Biodiversity Knowledge	Andrew C. Jones, Cardiff University, UK Richard J. White, Cardiff University, UK Gabriele Gianini, Università degli Studi di Milan, Italy Antonia Azzini, Università degli Studi di Milan, Italy Stefania Marrara, Università degli Studi di Milan, Italy
Immunity-Based Systems	Yoshiteru Ishida, Toyohashi University of Technology, Japan Takeshi Okamoto, Kanagawa Institute of Technology, Japan Yuji Watanabe, Nagoya City University, Japan Koji Harada, Toyohashi University of Technology, Japan

Program Committee

Abe, Akinori	IREIIMS University, Japan
Alexandre de Matos Araujo	Rui, University of Coimbra, Portugal
Angelov, Plamen	Lancaster University, UK
Anwer, Nabil	LURPA - ENS CACHAN, France
Aoki, Shingo	Osaka Prefecture University, Japan
Apolloni, Bruno	University of Milan, Italy
Aspragathos, Nikos A.	University of Patras, Greece
Bannore, Vivek	University of South Australia, Australia
Barb, Adrian S.	Penn State University, USA
Becker-Asano	Christian, Intelligent Robotics and Communication Labs, Japan
Bianchini, Monica	University of Siena, Italy
Bichindaritz, Isabelle	University of Washington, USA
Boeva, Veselka	Technical University of Sofia, Bulgaria
Boutalis, Yiannis	Democritus University of Thrace, Greece
Brna, Paul	University of Glasgow, UK
Buckingham, Christopher	Aston University, UK
Camastra, Francesco	University of Naples Parthenope, Italy
Cao, Cungen	Chinese Academy of Sciences, China
Ceccarelli, Michele	University of Sannio, Italy
Chalup, Stephan	The University of Newcastle, Australia
Chang, Bao Rong	National University of Kaohsiung, Taiwan
Chen, Lihui	Nanyang Technological University, Singapore
Chen, Toly	Feng Chia University, Taiwan
Cheng, Kai	Brunel University, UK
Cheung, Benny	Honk Kong Polytechnic University, Hong Kong
Cobos Pérez, Ruth	Universidad Autónoma de Madrid, Spain
Crippa, Paolo	Università Politecnica delle Marche, Italy

Cuzzocrea, Alfredo	University of Calabria, Italy
Damiana, Maria Luisa,	University of Milan, Italy
Dasiopoulou, Stamatia	Informatics and Telematics Institute, Greece
De Cock, Martine	University of Washington, USA
De Wilde, Philippe	Heriot-Watt University, UK
Dengel, Andreas	German Research Center for Arificial Intelligence (DFKI), Germany
Duro, Richard J.	Universidade da Coruña, Spain
Dustdar, Schahram	Vienna University of Technology, Austria
Elomaa, Tapio	Tampere University of Technology, Finland
Fernandez-Canque, Hernando	Glasgow Caledonian University, UK
Georgieva, Petia	University of Aveiro, Portugal
Godoy, Daniela	UNICEN University, Argentina
Grabot, Bernard	LGP-ENIT, France
Graña Romay, Manuel	Universidad del Pais Vasco, Spain
Grecos, Christos	University of West Scotland, UK
Hara, Takahiro	Osaka University, Japan
Hintea, Sorin	Cluj-Napoca University, Romania
Honda, Katsuhiro	Osaka Prefecture University, Japan
Hong, Tzung-Pei	National University of Kaohsiung, Taiwan
Hu, Chenyi	University of Central Arkansas, USA
Hurtado Larrain, Carlos	University of Chile, Chile
Ichalkaranje, Nikhil	University of South Australia, Australia
Ishibuchi, Hisao	Osaka Prefecture University, Japan
Ishida, Yoshiteru	Toyohashi University of Technology, Japan
Ito, Takayuki	Massachusetts Institute of Technology, USA
Ivancevic, Tijana	University of South Australia, Australia
Janicki, Ryszard	McMaster University, Canada
Jastroch, Norbert	MET Communications GmbH, Germany
Jensen, Richard	Aberystwyth University, UK
Jones, Andrew	Cardiff University, UK
Jordanov, Ivan	University of Portsmouth, UK
Jung, Jason J.	Yeungnam University, Korea
Juric, Matjaz B.	University of Maribor, Slovenia
Katagiri, Hideki	Hiroshima University, Japan
Ko, In-Young	KAIST, Korea
Kodogiannis, Vassilis S.	University of Westminster, UK
Koenig, Andreas	Technische Universitaet Kaiserslautern, Germany
Kojadinovic, Ivan	University of Auckland, New Zealand
Kompatsiaris, Yiannis	Informatics and Telematics Institute, Greece
Konar, Amit	Jadavpur University, India
Koshizen, Takamasa	Honda R&D Co., Ltd., Japan
Koychev, Ivan	University of Sofia, Bulgaria
Kwong, C.K.	The Hong Kong Polytechnic University, Hong Kong
Lee, Dah-Jye	Brigham Young University, USA
Lee, W.B.	Hong Kong Polytechnic University, Hong Kong
Likas, Aristidis	University of Ioannina, Greece

Lim, C.P.	Universiti Sains Malaysia, Malaysia
Liu, Lei	Beijing University of Technology, China
Maglogiannis, Ilias	University of Central Greece, Greece
Maier, Patrick	The University of Edinburgh, UK
Marinov, Milko T.	University of Ruse, Bulgaria
McCauley Bush, Pamela	University of Central Florida, USA
Montani, Stefania	Università del Piemonte Orientale, Italy
Moreno Jimenez, Ramón	Universidad del Pais Vasco, Spain
Nguyen, Ngoc Thanh	Wroclaw University of Technology, Poland
Nishida, Toyoaki	Kyoto University, Japan
Niskanen, Vesa A.	University of Helsinki, Finland
Ohkura, Kazuhiro	Hiroshima University, Japan
Palade, Vasile	Oxford University, UK
Pallares, Alvaro	Plastiasite S.A., Spain
Paranjape, Raman	University of Regina, Canada
Pasek, Zbigniew J.	University of Windsor, Canada
Pasi, Gabriella	University of Milan, Italy
Passerini, Andrea	Università degli Studi di Trento, Italy
Pazos-Arias, Jose	University of Vigo, Spain
Petrosino, Alfredo	Università di Napoli Parthenope, Italy
Prada, Rui	IST-UTL and INESC-ID, Portugal
Pratihar, Dilip Kumar	Osaka Prefecture University, Japan
Putnik, Goran D.	University of Minho, Portugal
Reidsema, Carl	University of New South Wales, Australia
Resconi, Germano	Catholic University in Brescia, Italy
Rovetta, Stefano	University of Genoa, Italy
Sansone, Carlo	Università di Napoli Federico II, Italy
Sarangapani, Jagannathan	Missouri University of Science and Technology, USA
Sato-Ilic, Mika	University of Tsukuba, Japan
Schockaert, Steven	Ghent University, Belgium
Seiffert, Udo	Fraunhofer-Institute IFF Magdeburg, Germany
Simperl, Elena	University of Innsbruck, Austria
Smrz, Pavel	Brno University of Technology, Czech Republic
Soroka, Anthony	Cardiff University, UK
Szczerbicki, Edward	The University of Newcastle, Australia
Tanaka, Takushi	Fukuoka Institute of Technology, Japan
Teng, Wei-Chung	National Taiwan University of Science and Technology, Taiwan
Tichy, Pavel	Rockwell Automation Research Centre, Czech Republic
Tino, Peter	The University of Birmingham, UK
Tolk, Andreas	Old Dominion University, USA
Toro, Carlos	VICOMTech, Spain
Torra, Vicenc	IIIA-CSIC, Spain
Tsihrintzis, George	University of Piraeus, Greece
Tsiporkova, Elena	Sirris, Belgium
Turchetti, Claudio	Università Politecnica delle Marche, Italy

Uchino, Eiji	Yamaguchi University, Japan
Urlings, Pierre	DSTO, Department of Defence, Australia
Vadera, Sunil	University of Salford, UK
Valdéz Vela, Mercedes	Universidad de Murcia, Spain
Vellido, Alfredo	Universitat Politècnica de Catalunya, Spain
Virvou, Maria	University of Piraeus, Greece
Wang, Zidong	Brunel University, UK
Watts, Mike	TBA, New Zealand
White, Richard J.	Cardiff University, UK
Williams, M. Howard	Heriot-Watt University, UK
Yang, Zijiang	York University, Canada
Yoshida, Hiroyuki	Harvard Medical School, USA
Zanni-Merk, Cecilia	LGeCo - INSA de Strasbourg, France
Zheng, Li-Rong	Royal Institute of Technology (KTH), Sweden

Reviewers

Adam Nowak
Adam Slowik
Adrian S. Barb
Akinori Abe
Akira Hattori
Alan Paton
Alessandra Micheletti
Alfredo Cuzzocrea
Ammar Aljer
Amparo Vila
Ana Fernandez-Vilas
Anastasia Kastania
Anastasius Moumtzoglou
Andrea Visconti
Andreas Abecker
Andreas Dengel
Andreas Oikonomou
Andrew Jones
Annalisa Appice
Anne Håkansson
Ansgar Bernardi
Anthony Soroka
Antonio Gomez-Skarmeta
Antonio Zippo
Aristidis Likas
Armando Buzzanca
Artur Silic
Athina Lazakidou
Azizul Azhar Ramli
Balázs Gaál
Bao Rong Chang
Benjamin Adrian
Bernard Grabot
Bernd Reusch
Bettina Waldvogel
Björn Forcher
Bob Coyne
Bojan Basrak
Bojana Dalbelo Basic
Bozidar Ivankovic
Branko Zitko
Bruno Apolloni
Calin Ciufudean
Carlo Sansone
Carlos Ocampo
Carlos Pedrinaci
Carlos Toro
Cecilia Zanni-Merk
Cesar Sanin
Chang-Tien Lu
Christian Becker-Asano
Christine Mumford
Chunbo Chu
Costantino Lucisano
C.P. Lim
Cristos Orovas
Daniela Godoy
Danijel Radosevic
Danilo Dell'Agnello
David Martens
David Vallejo
Davor Skrlec
Dickson Lukose
Dilip Pratihar
Doctor Jair Abe
Don Jeng
Donggang Yu
Doris Csipkes
Eduardo Cerqueira
Eduardo Merlo
Edward Szczerbicki
Eiji Uchino
Elena Pagani
Elena Simperl
Esmail Bonakdarian
Esmiralda Moradian
Francesco Camastra
Frane Saric
Fujiki Morii
Fumihiko Anma
Fumitaka Uchio
Gabbar Hossam
Gabor Csipkes
Gabriel Oltean
Gabriella Pasi
George Mitchell
George Tsihrintzis
Gergely Héja
Gianluca Sforza
Giovanna Castellano

Giovanni Gomez Zuluaga
Gunnar Grimnes
Gyorgy Surjan
Haoxi Dorje Zhang
Haruhiko H. Nishimura
Haruhiko Haruhiko Nishimura
Haruki Kawanaka
Hector Alvarez
Hernando Fernandez-Canque
Hideaki Ito
Hidehiko Hayashi
Hideo Funaoi
Hideyuki Matsumoto
Hirokazu Miura
Hisayoshi Kunimune
Hrvoje Markovic
Huey-Ming Lee
Ilias Maglogiannis
Ing. Angelo Ciccazzo
Ivan Koychev
Ivan Stajduhar
J. Mattila
Jair Abe
Jari Kortelainen
Jayanthi Ranjan
Jerome Darmont
Jessie Kennedy
Jesualdo Tomás Fernández-Breis
Jiangtao Cao
Jim Sheng
Johnson Fader
Jose Manuel Molina
Juan Botia
Juan Manuel Corchado
Juan Pavon
Julia Hirschberg
Jun Munemori
Jun Sawamoto
Junzo Watada
Jure Mijic
Katalina Grigorova
Katsuhiro Honda
Katsumi Yamashita
Kazuhiko Tsuda
Kazuhiro Ohkura
Kazuhiro Takeda
Kazuhisa Seta
Kazumi Nakamatsu
Kazunori Nishino
Kazuo Misue
Keiichiro Mitani
Kenji Matsuura
Koji Harada
Kouji Yoshida
Lars Hildebrand
Laura Caponetti
Lei Liu
Lelia Festila
Leonardo Mancilla Amaya
Lily Lin
Ljiljana Stojanovic
Lorenzo Magnani
Lorenzo Valerio
Ludger van Elst
Manuel Grana
Marek Malski
Maria Torsello
Mario Koeppen
Marko Banek
Martin Lopez-Nores
Martine Cock
Masakazu Takahashi
Masaru Noda
Masato Soga
Masayoshi Aritsugi
Mayumi Ueda
Melita Hajdinjak
Michelangelo Ceci
Michele Missikoff
Miguel Delgado
Milko Marinov
Minhua Ma
Minoru Minoru Fukumi
Monica Bianchini
Motoi Iwashita
Motoki Miura
Nahla Barakat
Naohiro Ishii
Naoto Mukai
Naoyuki Naoyuki Kubota
Narayanan Kulathuramaiyer
Nikica Hlupi
Nikola Ljubesic
Nikos Tsourveloudis
Nobuhiro Inuzuka
Nobuo Suzuki
Norbert Jastroch
Norio Baba
Noriyuki Matsuda
Omar Rana
Orleo Marinaro
Paolo Crippa
Pasquale Di Meo
Pavel Tichy
Philippe Wilde
Rafael Batres
Raffaele Cannone
Ramón Jimenez
Rashid Mehmood
Richard Pyle
Richard White
Robert Howlett
Roberto Cordone
Ronald Hartung
Roumen Kountchev
Rozália Lakner
Ruediger Oehlmann
Ruth Cobos
Ryohei Nakano
Ryuuki Sakamoto
Sachio Hirokawa
Satoru Fujii
Sebastian Rios
Sebastian Weber
Seiji Isotani
Seiki Akama
Setsuya Kurahashi
Shamshul Bahar Yaakob
Shinji Fukui
Shusaku Tsumoto
Shyue-Liang Wang
Simone Bassis
Sophia Kossida
Stamatia Dasiopoulou
Stefan Zinsmeister
Stefania Marrara

Stefania Montani
Stephan Chalup
Steven Schockaert
Sunil Vadera
susumu hashizume
Susumu Kunifuji
Takanobu Umetsu
Takashi Hamaguchi
Takashi Mitsuishi
Takashi Yukawa
Takaya Yuizono
Takeshi Okamoto
Taketoshi Kurooka
Taketoshi Ushiama
Takushi Tanaka
Tapio Elomaa
Tatiana Tambouratzis
Tessai Hayama
Thomas Roth-Berghofer
Tomislav Hrkac
Tomoko Kojiri
Toru Fukumoto
Toshihiro Hayashi
Toshio Mochizuki
Toumoto
Toyohide Watanabe
Toyohis Nakada
Tsuyoshi Nakamura
Tuan Pham
Valerio Arnaboldi
Vassilis Kodogiannis
Vassilis Moulianitis
Vesa Niskanen
Veselka Boeva
Vivek Bannore
Wataru Sunayama
Wei-Chung Teng
William Hochstettler
Winston Jain
Wolfgang Stock
Xiaofei Ji
Yi Xiao
Yiannis Boutalis
Yoshifumi Tsuge
Yoshihiro Okada
Yoshihiro Takuya
Yoshinori Adachi
Yoshiyuki Yamashita
Youji Ochi
Young Ko
Yuichiro Tateiwa
Yuji Iwahori
Yuji Wada
Yuji Watanabe
Yuki Hayashi
Yukio Ohsawa
Yumiko Nara
Yurie Iribe
Zdenek Zdrahal
Ziad Salem
Zijiang Yang
Zlatko Drmac
Zuwairie Ibrahim

Table of Contents – Part IV

Knowledge Based and Expert Systems

Knowledge Environment for Supporting Creative Learning

Information Communication Technology in Innovation and Creativity

Intelligent Support for Designing Social Information Infrastructure

Intelligent Systems in Ambient Assisted Living Environments

3D Visualisation of Natural Language

Knowledge-Based Creativity Support Systems

Tools and Techniques for Effective Creation and Exploitation of Biodiversity Knowledge

Real World Data Mining and Digital Intelligence

Advanced Design Techniques for Adaptive Systems

Emotion Judgment Method from an Utterance Sentence

Seiji Tsuchiya, Eriko Yoshimura, and Hirokazu Watabe

Dept. of Intelligent Information Engineering and Sciences, Doshisha University
Kyo-Tanabe, Kyoto, 610-0394, Japan
{stsuchiy,eyoshimu,hwatabe}@mail.doshisha.ac.jp

Abstract. Authors focus on the emotion of such common sense and attempt to establish a method to judge the user's emotions based on utterances. A speaker's utterance sentence includes a linguistic proposition and a linguistic modality. The linguistic modality is an important factor to represent emotions for an utterance sentence. Therefore, in this paper, a method is proposed which judges speaker's emotions from the linguistic proposition and modality included in the utterance sentence. The proposed method uses knowledge base and an Association Mechanism. As a result, the accuracy of the proposed emotion judgment method processing the linguistic modality is improved approximately 30% compared with an existing method.

Keywords: Emotion, Common Sense, Concept Base, Degree of Association.

1 Introduction

Authors are conducting research aiming to develop new interfaces that follow the mechanism of human communication, focusing on human common sense. Humans, in such communication, are able to appropriately interpret ambiguous information that they receive and carry on a smooth conversation. Common sense is knowledge (ability) that only man has. The person can express, and act feeling neither sense of incompatibility nor unnatural by using common sense. Moreover, when the sense of incompatibility and unnatural are felt, the person can appropriately interpret them.

Especially, authors focus on the emotion of such common sense and attempt to establish a method to judge the user's emotion based on the utterance. For example, using this system can output not an utterance cultivating sadness as "You are no good at anything" when a user say "Because I failed in work, it has been scolded by the superior" but a consoling utterance as "The odds are in your favor". Thus, using this system can select an appropriate expression if the content that the system tries to provide the user contains expressions that are unpleasant or remind the user of unhappy events.

Such system and method have already been developed. The developed method[1,2] judges user's emotions, categorized into 10 types, from a sentence

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 1–10, 2010.

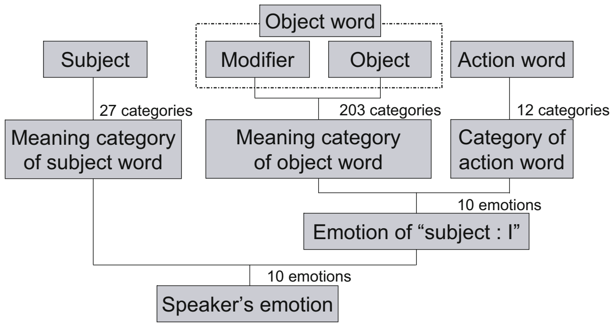


Fig. 1. Outline of the existing Emotion Judgment System

of the user utterance, based on the four components of the sentence: "subject", "modifier", "object word", and "action word".

However, the user utterance sentence is not always consisting of the four components. For example, the subject is often omitted in the conversation. Moreover, the intention and emotions change depending on ending of words of a sentence in Japanese. Therefore, in this paper, a method is proposed which judges speaker's emotions from the linguistic proposition and modality included in the utterance sentence. The proposed method uses knowledge base and an Association Mechanism.

2 The Existing Emotion Judgment System

The components of the utterance sentences to be used to judge speaker's emotion were limited to four ("subjects", "modifiers", "objects" and "action words")[1,2]. Figure 1 shows outline of the existing Emotion Judgment method.

A "subject" was categorized into 3 attributes: liking (likes and dislikes), familiarity (closeness) and sociality (good and evil). These 3 attributes have 3 values. In short, "subject" is categorized into 27 categories. The categorization is processed using the knowledge base based on thesaurus[3].

A "modifier" was an adjective or "adjectival verb" that modifier the "object" which follows the modifier. "Modifiers" may be omitted, as they were not always necessary in textual expression. The direct modification and dependent modification types were further divided into different groups having similar meaning according to the adjectives describing the modifiers, and they were registered in the knowledge base for emotion judgment.

An "object" was a noun that denotes the object of the subject's action, behavior, or state. Objects were also classified according to their meanings using the 203 sense words that the Sense Judgment System[4,5] can judge. These 203 sense words share the common meaning categories with the modifiers discussed earlier. In addition, "modifiers" and "objects" collectively were referred to as "object words". In short, the 203 sense words were used to categorize the meanings of the object words.

An “action word” was a verb, adjective, or “adjectival verb” that describes the subject’s action, behavior, or state. An action word converted the feature related to the sense and perception that associated with an object word. Features expressed in terms of senses and perceptions could be roughly divided into positive and negative expressions. Likewise, emotions could also be categorized into two groups, positive and negative. Therefore, four types of effect could be found in the action words.

A sentential actor’s (“subject”) emotion was judged based on the “object words”, and “action words”. With respect to the emotions that were generated, those associated with a total of 406 pairs of the meaning categories of object words (203 categories) and action words (2 categories of “succession” and “opposite”) were manually defined and registered in the system’s Emotion Judgment Knowledge Base. Therefore, speaker’s emotion is judged by combination of attributes values (27 categories) of sentential actor (“subject”) and judged emotions of sentential actor. 270 rules which are combination of 27 categories (attributes values) of sentential actor and 10 emotions are registered in the knowledge base.

Some knowledge related to the “generation of emotion”, the “action words”, and the “modifiers” of the “object words” were registered in the Emotion Judgment Knowledge Base. Based on this, the system associated words and expanded its knowledge within the range of common sense, making it possible to handle many expressions. The word association was realized by using the huge Concept Base[6,7] that was automatically built from multiple digital dictionaries, and a method to calculate the Degree of Association[8] that evaluates the relationship between words. Hereafter, this Concept Base and the calculation method are called the “Association Mechanism”.

3 Elemental Technique

3.1 Concept Base

The Concept Base is a large-scale database that is constructed both manually and automatically using words from multiple electronic dictionaries as concepts and independent words in the explanations under the entry words as concept attributes. In the present research, a Concept Base containing approximately 90,000 concepts was used, in which auto-refining processing was carried out after the base had been manually constructed. In this processing, attributes considered inappropriate from the standpoint of human sensibility were deleted and necessary attributes were added.

In the Concept Base, Concept A is expressed by Attributes a_i indicating the features and meaning of the concept in relation to a Weight w_i denoting how important an Attribute a_i is in expressing the meaning of Concept A. Assuming that the number of attributes of Concept A is N, Concept A is expressed as indicated below. Here, the Attributes a_i are called Primary Attributes.

$$A = \{(a_1, w_1), (a_2, w_2), \cdots, (a_N, w_N)\}$$

	train, 0.36	locomotive, 0.21	railroad, 0.10	...	a_i, w_i	Primary Attributes
	train, 0.36	locomotive, 0.21	railroad, 0.10	...	a_{i1}, w_{i1}	Secondary Attributes
train	locomotive, 0.21	streetcar, 0.23	subway, 0.25	...	a_{i2}, w_{i2}	
	:	:	:	:	:	
Concept	a_{1j}, w_{1j}	a_{2j}, w_{2j}	a_{3j}, w_{3j}	...	a_{ij}, w_{ij}	

Fig. 2. Example of the Concept "train" expanded as far as Secondary Attributes

Because Primary Attributes a_i of Concept A are taken as the concepts defined in the Concept Base, attributes can be similarly elucidated from a_i. The Attributes a_{ij} of a_i are called Secondary Attributes of Concept A. Figure 2 shows the elements of the Concept "train" expanded as far as Secondary Attributes.

3.2 Degree of Association Algorithm

Each concept is defined as a set of attributes and each attribute is also a concept as described in section 3.1. Each concept is defined by an infinite chain of attributes (concepts). In this paper, a method to derive the Degree of Association between concepts using up to Secondary Attributes is used.

The Degree of Match between Two Sets of Primary Attributes: DM The relation is higher if two concepts have more of the same attributes since a concept has related concepts as the attributes. When two concepts are defined as the following equations using Primary Attributes a_i, b_j and their weights u_i, v_j,

$$A = \{(a_i, u_i) | i = 1 \cdots L\} \qquad B = \{(b_j, v_j) | j = 1 \cdots M\}$$

the degree of match(DM) between concept A and B is defined as the following equation.

$$DM(A, B) = \frac{\frac{s}{L} + \frac{s}{M}}{2}$$

Where, s is the number of same attributes $(a_i = b_j)$ between concept A and B. In this definition, the weights of the attributes are ignored.

The degree of match between concepts A and B using the weights of attributes is defined by the following equations.

$$DMW(A, B) = \frac{\frac{s_A}{n_A} + \frac{s_B}{n_B}}{2} \tag{1}$$

Where,

$$s_A = \sum_{a_i = b_j} u_i \qquad s_B = \sum_{a_i = b_j} v_j$$

$$n_A = \sum_{i=1}^{L} u_i \qquad n_B = \sum_{j=1}^{M} v_j$$

The equation 1 represents the mean value of s_A/n_A (which is the rate of the same attributes in concept A) and s_B/n_B.

The Degree of Association between Concepts Using the Weights of Attributes: DA

The Degree of Association between concept A and B using the weights of attributes is derived by the following algorithm.

(1) Assuming the number of attributes of concept A is less than or equal to the number of attributes of concept $B(L \leq M)$, fix the order of attributes of concept A.

$$A = ((a_1, u_1), (a_2, u_2), \cdots, (a_L, u_L))$$

(2) Make the matrix of the degree of match using the weights of attributes.

$$M_{DMW(A,B)} = \left[\begin{array}{c|cccc} & b_1 & b_2 & \cdots & b_M \\ \hline a_1 & dw_{11} & dw_{12} & \cdots & dw_{1M} \\ a_2 & dw_{21} & dw_{22} & \cdots & dw_{2M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_L & dw_{L1} & dw_{L2} & \cdots & dw_{LM} \end{array}\right]$$

Where, each element of this matrix is the degree of match between two corresponding attributes using the weights of their attributes,

$$dw_{ij} = DMW(a_i, b_j)$$

(3) Reorder each attribute of concept B so that the sum of the degree of match between corresponding attributes using the weights becomes maximum.

$$B_x = ((b_{x1}, v_{x1}), (b_{x2}, v_{x2}), \cdots, (b_{xL}, v_{xL}))$$

Where, $\{b_{xj} | j = L+1, \cdots, M\}$ are ignored.

(4) The Degree of Association between concept A and B using the weights of attributes is derived by the following equations.

$$DA(A, B) = \frac{\frac{s_A}{n_A} + \frac{s_B}{n_B}}{2}$$

$$s_A = \sum_{i=1}^{L} u_i DMW(a_i, b_{xi})$$

$$s_B = \sum_{i=1}^{L} v_{xi} DMW(a_i, b_{xi})$$

$$n_A = \sum_{i=1}^{L} u_i \qquad n_B = \sum_{j=1}^{M} v_j$$

first-order attributes

desk = {school, study, bookshelf}
chair = {study, classroom, tree}

second-order attributes

school = {college, student, wood}
study = {book, pen, teacher}
bookshelf = {book, wall, box}
classroom = {student, blackboard, college}
tree = {wood, leaf, green}

matrix of the degree of match

	school	study	bookshelf
study	0	$\boxed{1}$	1/3
classroom	$\boxed{2/3}$	0	0
tree	1/3	0	$\boxed{0}$

calculation

chair_x =(classroom, study, tree)
s = 2/3+1+0 = 5/3
DA(school, chair) = (5/3)/3 = 5/9

Fig. 3. An Example of calculating the Degree of Association

Figure 3 shows an example of the Degree of Association ignoring the weights.

To understand the contents of a document, we use the Concept Base, which expresses the semantic characteristics of a word with a word and a weight, and also the method of calculating the Degree of Association, which numerically calculates the semantic relationship between words.

3.3 Sense Judgment System[4,5]

The knowledge base for the sense and perception judgments has a structure like a thesaurus, and it contains sense and perception words that are associated with typical nouns, which have been entered manually. In cases when an unknown word not registered in the Sense Knowledge Base needs to be processed, the system calculates the Degree of Association with those known words registered in the knowledge base for the sense and perception judgments and chooses the one with the highest Degree of Association for processing. This lets the system obtain the rough corresponding sense and perception. In addition, the system refers to the attributes registered in the Concept Base to find the sense and perception particular to that word. Due to its structure, these attributes in the Concept Base contain some inappropriate words as senses and perceptions to be associated, and thus the system is carefully designed so that the correct sense and perception is selected using the Degree of Association.

4 Proposed Emotion Judgment Method

A speaker's utterance sentence includes a linguistic proposition and a linguistic modality. The linguistic proposition represents objective meaning contents of a sentence. The linguistic modality represents mind attitude of speaker. For example, a sentence "I get a cold." is just information. Such sentences including only the linguistic proposition were processed by the existing Emotion Judgment System. However, a sentence "I managed to get a cold." represents emotion "sadness" for speaker. In this way, the linguistic modality is an important factor to represent emotions for an utterance sentence.

Then, in this paper, an proposed emotion judgment method processes the linguistic modality in an utterance sentence. A humanlike emotion judgment is achieved by the linguistic proposition process and the linguistic modality process.

4.1 Classification of Linguistic Modality

The linguistic modality representing mind attitude is important for conveying emotions with conversation. In this paper, major five kinds of the linguistic modality: "Approach", "Expression", "Question", "Declarative" and "Exclamation" are used for classification of the linguistic modality. Moreover, four kinds of mark: "!", "?", "..." and "no mark" are used for subcategorization of the classification. In addition, the use of such signs assumes recognition by the voice recognition technology.

These classifications are achieved by using the knowledge base. The knowledge base is manually constructed based on Japanese grammar books. In addition, 503 words are registered in the knowledge base.

4.2 Redefined Emotion

Many psychologists, philosophers, etc., have studied human emotions[9,10,11]. However, these researchers have had different interpretations of emotion and devised different models for emotion, as emotions have no substance and are quite ambiguous. Therefore, in the existing Emotion Judgment System, emotion has been defined as "something one feels instantaneously when an action takes place" and has defined the following ten emotions to judge: "joy", "sadness", "anger", "ease", "fear", "disappointment", "shame", "regret", "sense of guilt", and "no emotion".

In this paper, the following six emotions to judge are redefined: "joy", "sadness", "fear", "anger", "hate", "surprise" and "no emotion". These emotions can be judged from face expressions that are common in all culture[12]. The facial recognition and the language processing can cooperate in the future by sharing emotions that can be judged from the face expression and the language.

4.3 Method of Emotion Judgment

Figure 4 shows an outline of the proposed emotion judgment method.

The existing Emotion Judgment System judges speaker's emotions from only linguistic proposition. In addition, the existing Emotion Judgment System is

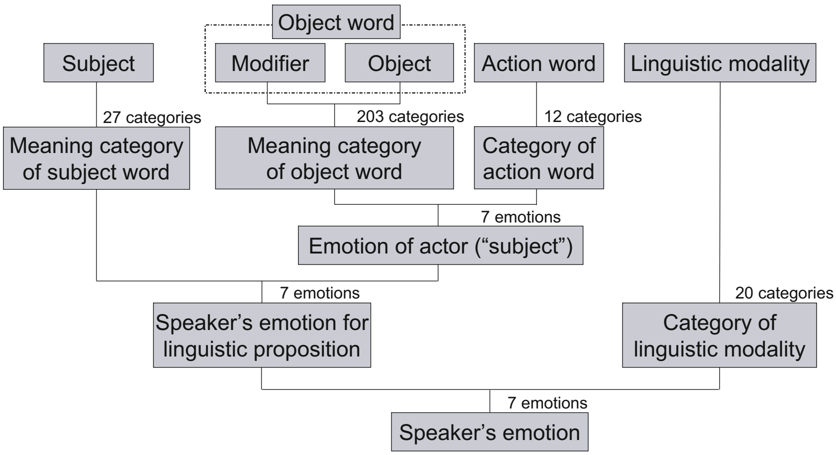


Fig. 4. Outline of proposed Emotion Judgment System

arranged to output seven kinds of emotion. Moreover, the proposed emotion judgment method judges the category of the linguistic modality from the linguistic modality. Then, the speaker's emotions are judged from combination of both such results. 140 rules which are combination of 7 emotions and 20 categories (combination of 5 categories and 4 subsategories for the linguistic modality) are registered in the knowledge base.

5 Performance Evaluation of the Proposed Emotion Judgment Method

5.1 Evaluation Data

In this paper, the purpose is to judge speaker's emotions from an utterance sentence in daily conversation. Then, dialogues of a movie that were natural conversation were used for an evaluation. 200 sentences about which two people were talking were evaluated.

The movie was shown to 20 test subjects, and they were asked to judge an emotion of utterer. As a result of the experiment, the difference was caused in the answer. Then, suitable emotions for the correct answer were defined by using the following expressions.

$$E_{n+1} > E_n \times (1 - 0.5^n)$$

Where, E_n is nth emotion. In short, E_n when the abovementioned expression consists is the correct answer (the suitable emotion).

5.2 Evaluation of Human Judgment

An accuracy rate of human judgments was evaluated based on the abovementioned the correct answer (the suitable emotions). 20 test subjects' judgments were evaluated. As a result, the accuracy rate of human judgments was 79.2% on the average. From this, it was reaffirmed that human cannot perfectly judge others' emotions. In this paper, the proposed emotion judgment method aims at this human's accuracy rate.

5.3 Evaluation of the Emotion Judgment Method

Table 1 shows the results of the proposed emotion judgment method and the existing Emotion Judgment System. As a result, the accuracy of the proposed emotion judgment method processing the linguistic modality was improved 29.5% compared with the existing method. An accuracy rate of the existing Emotion Judgment System processing only the linguistic modality was reported to be 85%. The reported result has the difference of 73% compared with this paper's result. This fact shows that the possibility of the linguistic modality to be included in an utterance sentence for natural conversation is high. Moreover, these results show that the proposed emotion judgment method processing the linguistic modality is very effective for an utterance sentence in conversation. In addition, the proposed emotion judgment method is able to come closer to human's judgment in 52.4%.

Table 1. Results of human judgment, proposed method and existing method

	Human judgment	Proposed method	Existing method
Accuracy rate	79.2%	41.5%	12.0%

6 Conclusions

Authors focused on the emotion of common sense and attempt to establish a method to judge the user's emotions based on utterances. A speaker's utterance sentence includes a linguistic proposition and a linguistic modality. The linguistic modality is an important factor to represent emotions for an utterance sentence. Therefore, in this paper, the method was proposed which judged speaker's emotions from the linguistic proposition and modality included in the utterance sentence. The proposed method used knowledge base and an Association Mechanism.

As a result, the accuracy of the proposed emotion judgment method processing the linguistic modality was improved approximately 30% compared with the existing method. In short, the proposed emotion judgment method processing the linguistic modality was very effective for an utterance sentence in conversation. In addition, the proposed emotion judgment method was able to come closer to human's judgment in 52.4%.

Acknowledgment

This research has been partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (Young Scientists (B), 21700241).

References

1. Tsuchiya, S., Yoshimura, E., Watabe, H., Kawaoka, T.: The Method of the Emotion Judgment Based on an Association Mechanism. Journal of Natural Language Processing 14(3), 119–238 (2007)
2. Tsuchiya, S., Yoshimura, E., Ren, F., Watabe, H.: Emothion Judgment based on Relationship between Speaker and Sentential Actor. In: Velásquez, J.D., Ríos, S.A., Howlett, R.J., Jain, L.C. (eds.) Knowledge-Based and Intelligent Information and Engineering Systems. LNCS (LNAI), vol. 5711, pp. 62–69. Springer, Heidelberg (2009)
3. NTT Communication Science Laboratory: NIHONGOGOITAIKEI (Japanese). Iwanami Shoten (1997)
4. Horiguchi, A., Tsuchiya, S., Kojima, K., Watabe, H., Kawaoka, T.: Constructing a Sensuous Judgement System Based on Conceptual Processing. In: Gelbukh, A. (ed.) CICLing 2002. LNCS, vol. 2276, pp. 86–95. Springer, Heidelberg (2002)
5. Watabe, H., Horiguchi, A., Kawaoka, T.: A Sense Retrieving Method from a Noun for the Commonsense Feeling Judgement System. Journal of Artificial Intelligence 19(2), 73–82 (2004)
6. Hirose, T., Watabe, H., Kawaoka, T.: Automatic Refinement Method of Concept-base Considering the Rule between Concepts and Frequency of Appearance as an Attribute. Technical Report of the Institute of Electronics, Information and Communication Engineers. NLC2001-93, pp. 109–116 (2002)
7. Kojima, K., Watabe, H., Kawaoka, T.: A Method of a Concept-base Construction for an Association System: Deciding Attribute Weights Based on the Degree of Attribute Reliability. Journal of Natural Language Processing 9(5), 93–110 (2002)
8. Watabe, H., Kawaoka, T.: Measuring Degree of Association between Concepts for Commonsense Judgements. Journal of Natural Language Processing 8(2), 39–54 (2001)
9. Hukui, Y.: KANJO-NO-SHINRIGAKU (Japanese). Kawashima Shoten (1990)
10. Rita, C.: NO-TO-KOKORO-NO-CHIZU (Japanese). Hara Shobo (1999)
11. Susan, G.: NO-NO-TANKYU (Japanese). Mumeisha (2001)
12. Ekman, P.: Argument for basic emotions. In: Cognition and Emotion, pp. 169–200 (1992)

Local Model Update with an Application to Sliding Window Protocol

Michael Kelly and Yan Zhang

Intelligent Systems Laboratory
School of Computing & Mathematics
University of Western Sydney
{mkelly,yan}@scm.uws.edu.au

Abstract. Local model update is a procedure whereby a property is found unsatisfied in some model and can be effectively made satisfied in the error reporting counterexample through a series of modifications. With this the counterexample can be reintegrated into the original model to derive a global fix. In this paper we put forward the case for local model update as a method for automatic update. We review theory underlying local model update and introduce the use of a tool `lmu` for automatically generating candidate updates, looking at general design principles. Finally we present a case study of the update process as applied to an abstraction of the sliding window protocol under a man-in-the-middle attack, showing the modelling process, update procedure and final derivation of a coherant model fix.

Keywords: Model checking, model update, intelligent systems, knowledge representation.

1 Introduction

Model checking is well known for its ability to verify complex system models [1,5]. Model checkers accepting ACTL specifications return counterexamples explaining the fault apparent in the given model [3,4]. These are useful to the software developer, but the task of performing the system update such that the model satisfies the given property is left unsupported. As such it is left to the developer to modify the model to achieve property satisfaction, which could lead to human fault in modifying the model or in worse cases fix their problem but cause other required properties to become unsatisfied in the model.

Methods of computer aided modification have previously been proposed [7,8,10]. These involve AI techniques such as belief revision and model update. One such approach proposed by Zhang and Ding [11] involves automated update on some Kripke structure which has been shown to fail over a temporal specification in CTL. While this technique proposes possible updates on the model structure it requires taking the entire model as input, which for large scale industrial domains is technically infeasible.

As is known, counterexamples generated from model checking sessions play a role in system repair. Counterexamples reveal the fault in the system in the

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 11–21, 2010.

context of the specification given [2,3,4]. One method to overcome complexity would be to develop a local model update system which performs updates on the counterexample, generating candidate modifications on the counterexample model and reintegrating it with the original system, working around having to process the entire model and irrevelant features of the system.

In this paper we review the groundings of local model update and look at some of the issues of design with `lmu`. Finally we consider a case study for update, applying the technique to an abstraction of the communications protocol sliding window with a man-in-the-middle attacker, varying window sizes to show computation time as model and property complexity increases.

2 Theoretical Background

To start we look at syntax of ACTL and background to local model update. Readers are referred to [5,9,11] for a deeper background into model checking, CTL and ACTL.

ACTL is a fragment of Computation Tree Logic (CTL) that has attracted considerable studies from researchers [3]. Besides Boolean connectives, ACTL provides both linear time operators X, F, G and U and the branching time operator A.

ACTL has the following syntax given in Backus-Naur form:

$$\phi ::= \top \mid \bot \mid p \mid \neg p \mid \phi \wedge \psi \mid \phi \vee \psi \mid \mathrm{AX}\phi \mid \mathrm{AG}\phi \mid \mathrm{AF}\phi \mid \mathrm{A}[\phi \mathrm{U} \psi],$$

where p is any propositional atom (variable). Let AP be a set of propositional variables. A *Kripke structure* M over AP is a triple $M = (S, R, L)$, where S is a finite set of states, $R \subseteq S \times S$ is a binary relation representing state transitions, and $L : S \rightarrow 2^{AP}$ is a labeling function that assigns each state with a set of propositional variables. A common method of visualising a Kripke structure is as a rooted graph (S, R) whose nodes are labelled by L.

We evaluate an ACTL formula over a Kripke structure. A *path* in a Kripke structure from a state is a(n) (infinite) sequence of states. Note that for a given path, the same state may occur an infinite number of times in the path (i.e., the path contains a loop).

Let $M = (S, R, L)$ be a Kripke structure and $s_0 \in S$. A *path* in M starting from s_0 is denoted as $\pi = [s_0, \cdots, s_i, s_{i+1}, \cdots]$, where $(s_i, s_{i+1}) \in R$ holds for all $i \geq 0$. If $\pi = [s_0, s_1, \cdots, s_i, \cdots, s_j, \cdots]$ and $i < j$, we denote $s_i < s_j$.

For any $s \in S$, the satisfaction relation between (M, s) and an ACTL formula ϕ, denoted by $(M, s) \models \phi$, is defined in a standard way as described in [9].

$M = (S,\ R,\ L)$, a Kripke interpretation $(M,\ s)(s \in S)$:

1. $(M,\ s) \models \top$ *and* $(M,\ s) \not\models \bot$ *for all* $s \in S$.
2. $(M,\ s) \models p$ *iff* $p \in L(s)$.
3. $(M,\ s) \models \neg\phi$ *iff* $(M,\ s) \not\models \phi$.

4. $(M,\ s) \models \phi_1 \wedge \phi_2$ *iff* $(M,\ s) \models \phi_1$ *and* $(M,\ s) \models \phi_2$.
5. $(M,\ s) \models \phi_1 \vee \phi_2$ *iff* $(M,\ s) \models \phi_1$ *or* $(M,\ s) \models \phi_2$.
6. $(M,\ s) \models AX\phi$ *iff* $\forall s_1$ *such that* $(s,\ s_1) \in R, (M,\ s_1) \models \phi$.
7. $(M,\ s) \models AG\phi$ *iff* $\forall \pi = [s_0,\ s_1, \ldots](s_0 = s)$ *and* $\forall s_i \in \pi, (M,\ s_i) \models \phi$.
8. $(M,\ s) \models AF\phi$ *iff* $\forall \pi = [s_0,\ s_1, \ldots](s_0 = s), \exists s_i \in \pi$ *such that* $(M,\ s_i) \models \phi$.
9. $(M,\ s) \models A[\phi_1 \cup \phi_2]$ *iff* $\forall \pi = [s_0,\ s_1, \ldots](s_0 = s), \exists s_i \in \pi$ *such that* $(M,\ s_i) \models \phi_2$ *and* $\forall j < i, (M,\ s_j) \models \phi_1$

Now we introduce the concept of tree-like Kripke structures [3]. Let G be a directed graph. A *strongly connected component* (SCC) C in G is a maximal subgraph of G such that every node in C is reachable from every other node in C. C is *nontrivial* iff either it has more than one node or it contains one node with a self-loop. The *component graph* $c(G)$ of G is the graph where the vertices are given by the SCCs of G, and where two vertices of $c(G)$ are connected by an edge if there exists an edge between vertices in the corresponding SCCs. Then we say a graph G is *tree-like* if (1) all its SCCs are cycles; and (2) $c(G)$ is a directed tree.

We stress condition (1) as SCCs may not be cycles. For instance, in a graph $G = (V, E)$, where $V = \{s_1, s_2, s_3\}$ and $E = \{(s_1, s_2), (s_2, s_3), (s_3, s_3), (s_3, s_2)\}$, the subgraph
$G' = (\{s_2, s_3\}, \{(s_2, s_3), (s_3, s_3), (s_3, s_2)\})$ is a SCC, but it is not a cycle because edge (s_3, s_3) also forms a self-loop.

Consider a Kripke model (M, s_0), where $M = (S, R, L)$ and $s_0 \in S$. We say that (M, s_0) is a *tree-like Kripke model* if its corresponding graph $G(M) = (S, R)$ is tree-like. In this case, we call the initial state s_0 the *root* of this tree-like model. Since a tree-like Kripke model may not be a strict tree (e.g. it may contain some cycles along a branch), we cannot follow the traditional notions of *child and parent* in a tree-like model. Instead, we define the following new concepts. We say state s is an *ancestor* of state s', if there is a path $\pi = [s_0, \cdots, s, \cdots, s', \cdots]$ where s' does not occur in the part $[s_0, \cdots, s]$.

s is a *parent* of s' if s is an ancestor of s' and $(s, s') \in R$. In this case, we also call s' is a *successor* of s. A state s is called *leaf* if it is not an ancestor of any other states. A tree-like model (M', s') is called a *submodel* of (M, s), if $M' = (S', R', L')$, $s' \in S'$, $S' \subseteq S$, $R' \subseteq R$, for all $s^* \in S'$, $L'(s^*) = L(s^*)$, and in M, s' is an ancestor of all other states in M'. Figure 1 from [3] shows an example of a tree-like model that represents a counterexample for a specific ACTL formula.

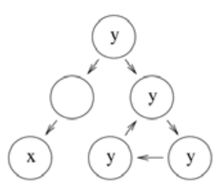


Fig. 1. A counterexample for AG$\neg x \vee$AF$\neg y$

Clarke et al. [3] proved an important result regarding ACTL model checking stating that if an ACTL formula is not satisfied in a Kripke structure, then this Kripke structure must contain a tree-like counterexample with respect to this formula.

Theorem 1. [3] ACTL has tree-like counterexamples.

Tree-Like Model Update

One major obstacle restricting the application of CTL model update in practical domains is that often we have to deal with a large Kripke model, and the update may also lead to a model explosion [11]. This motivates us to focus on the tree-like counterexample update. Since we can view a tree-like counterexample as a partial Kripke structure which is usually small, the update on this counterexample actually provides a computer aided approach for effective system modifications.

For convenience, in the rest of the paper, we will call (M, s) a *tree-like Kripke model* without explicitly mentioning the corresponding tree-like Kripke structure $M = (S, R, L)$ where $s \in S$. We also define $\mathsf{Diff}(X, Y) = (X \setminus Y) \cup (Y \setminus X)$ where X, Y are two sets.

Definition 1 (Weak bisimulation). *Let (M, s) and (M_1, s_1) be two tree-like Kripke models. where $M = (S, R, L)$ and $M_1 = (S_1, R_1, L_1)$ and $s \in S$ and $s_1 \in S_1$. We say that a binary relation $H \subseteq S \times S_1$ is a* weak bisimulation *between (M, s) and (M_1, s_1) if:*

1. *$H(s, s_1)$;*
2. *given $v, v' \in S$ such that v is a parent of v', for all $v_1 \in S_1$ such that $H(v, v_1)$, the condition holds: (a) if v_1 is not a leaf, then there exists successor v'_1 of v_1 such that $H(v', v'_1)$, or (b) if v_1 is a leaf, then $H(v', v_1)$ (forth condition);*
3. *given $v_1, v'_1 \in S_1$ such that v_1 is a parent of v'_1, for all $v \in S$ such that $H(v, v_1)$, the condition holds: (a) if v is not a leaf, then there exists a successor v' of v such that $H(v', v'_1)$, or (b) if v is a leaf, then $H(v, v'_1)$ (back condition).*

Definition 1 is inspired from the concept of bisimulation on Kripke models in classical modal logics. It is observed that for any two tree-like models where $S \neq \emptyset$, there exists at least one weak bisimulation between them, occurring at the root state s_0 of each. Usually, there are more than one such weak bisimulations.

Definition 2 (Bisimulation ordering). *Let (M, s), (M_1, s_1) and (M_2, s_2) be three tree-like models, H_1 and H_2 be two weak bisimulations between (M, s) and (M_1, s_1) and between (M, s) and (M_2, s_2) respectively. We say that H_1 is* as similar as *H_2, denoted by $H_1 \leq H_2$, if for all nodes $v \in S$, the following condition holds:*

1. *there exists an ancestor v' of v such that for all $v_1 \in S_1$ and $v_2 \in S_2$ satisfying $H_1(v', v_1)$ and $H_2(v', v_2)$, $\mathsf{Diff}(L(v'), L_1(v_1)) \subset \mathsf{Diff}(L(v'), L_2(v_2))$; or*
2. *for all $v_1 \in S_1$ and $v_2 \in S_2$ satisfying $H_1(v, v_1)$ and $H_2(v, v_2)$, $\mathsf{Diff}(L(v), L_1(v_1)) \subseteq \mathsf{Diff}(L(v), L_2(v_2))$.*

We write $H_1 < H_2$ iff $H_1 \leq H_2$ but $H_2 \not\leq H_1$.

Definition 2 specifies how we compare two weak bisimulations among three tree-like models. Intuitively, if H_1 and H_2 are two weak bisimulations between (M, s) and (M_1, s_1), and between (M, s) and (M_2, s_2) respectively, then $H_1 \leq H_2$ means that M_1 represents at least the same information about M as M_2 does under H_1 and H_2 respectively.

Note that if (M_1, s_1) and (M_2, s_2) are identical, then it is still possible to have two different weak bisimulations between (M, s) and (M_1, s_1). Hence, we are always interested in that H_1 where there is no other H_1' between (M, s) and (M', s') such that $H_1' < H_1$. We call such H_1 a *minimal weak bisimulation* between (M, s) and (M', s'), which, as should be noted, is not necessarily unique.

Definition 3 (Tree-like model update). *Let ϕ be an ACTL formula and (M, s) a tree-like model such that $M \not\models \phi$. A tree-like model (M_1, s_1) is called a result of updating (M, s) with ϕ, if and only if*

1. *$(M_1, s_1) \models \phi$;*
2. *there is a weak bisimulation H_1 between (M, s) and (M_1, s_1) such that there does not exist another tree-like model (M_2, s_2) satisfying $(M_2, s_2) \models \phi$ and a weak bisimulation H_2 between (M, s) and (M_2, s_2) such that $H_2 < H_1$. In this case we say that (M_1, s_1) is an update result* under *H_1.*

Condition 1 in Definition 3 simply states that after the update, the resulting tree-like model should satisfy the updating formula. Condition 2 ensures that the resulting tree-like model is minimal from the original model under some weak bisimulation.

3 System Design

To automatically generate model modifications on the fly we have constructed an implementation $\texttt{lmu}_c$ (local model update core/counterexample) which utilises NuSMV as a front-end to generate counterexamples for models found to not satisfy the given property.

In `lmu` the program $\mathcal{M}$ is checked against the desired system property φ in NuSMV to verify its satisfaction. If unsatisfied in $\mathcal{M}$, `lmu` parses the given property and counterexample into discrete structures and applies the $\texttt{lmu}_c$ algorithm to have $\mathcal{M}'$ satisfy φ. $\mathcal{M}'$ is reintegrated into $\mathcal{M}^\dagger$, the global model, and rechecked by NuSMV.

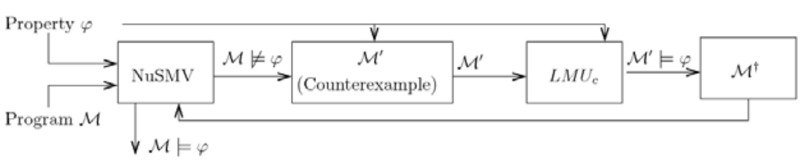


Fig. 2. Process graph of `lmu`

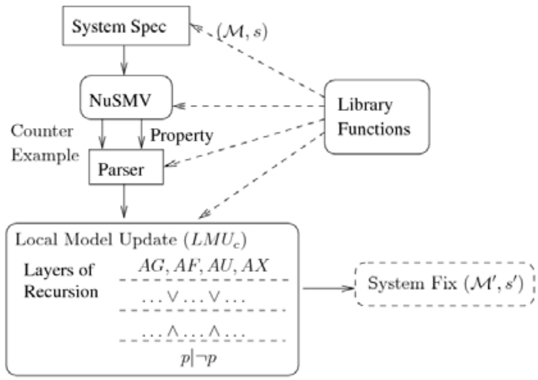


Fig. 3. Implementation graph

In determining the candidate model fix, the $\texttt{lmu}_c$ algorithm was designed with a top down recursive approach with respect to the given model and properties, similar in nature to the approaches of Huth, Zhang and Ding in [9,11] respectively.

Developed in the python language, `lmu` uses parsers built in the `pyParsing` environment and other library functionality to interpret the given properties into parse tree structures for use in update. Similiarly counterexamples are parsed and built into tree-like data structures for update. Execution of the update process begin from the root node of both the system property and model data structure. The model is parsed and operated upon based on the operational semantic instructions given by the ACTL property parse tree.

4 Case Study: Sliding Window Protocol (SWP)

To demonstrate an application of local model update we apply the process to a model of the SWP containing an attacker deriving a minimal fix to the transmission protocol.

SWP is a common protocol used for point to point packet-based communication in networked environments. Notable examples of SWP variants are point-to-point protocol (PPP) and general TCP connections [6].

In SWP the process occurs between three entities, a sender, a receiver and their given medium. Messages with unique identifiers are passed through the medium to the recipient who then sends an acknowledgement token back through the medium along with the original messages identifier. Both sender and receiver maintain state by keeping a copy of the last identifier value sent.

4.1 Methodology

The applied method for determining the worst case complexity and efficacy of local update involves modelling SWP with increasing window sizes to determine computation time for update as model and formula size increases. For the purposes of this test we check window size $n = 2..8$ and log computation time.

For the purpose of modelling we use the NuSMV language. With this we define a liveness property we wish to check based on the window size specified. For each successive model we find $\mathcal{M} \not\models \varphi$ and derive a counterexample T with which we apply $update(T,\ \varphi)$ such that $T \models \varphi$. All experiments are completed on a system with an AMD Phenom 8450, 2.10GHz, 3-core processor and 2GB 800MHz RAM.

4.2 Modelling Sliding Window

For the purpose of simplicity we abstract out many of the details of SWP in an attempt to verify a single property in the model.

We define two processes sender and receiver which accept a two arguments representing the packet id sent and received either way, each called send and receive respectively. The second argument is the window size for the given module. The incrementing of each id represents the shifting window buffer performed on the sender module when it receives an id value one higher than its last sent message. With this it sends out a message with the next id. For the purpose of these experiments n represents the window size designated for the current test case of SWP.

```
MODULE sender(receivedId, n)
  VAR
    sentId : 0..(n - 1);
  ASSIGN
    next(sentId) :=
      case
        (sentId + 1) mod n = receivedId :
            ((sentId + 2) mod n) union sentId;
        1 : (sentId + 1) mod n;
      esac;
  FAIRNESS running
```

```
MODULE receiver(sentId, n)
  VAR
    receivedId : 0..(n - 1);
  ASSIGN
    next(receivedId) :=
      case
        (sentId = receivedId) : receivedId;
        1 : ((receivedId + 1) mod n);
      esac;
  FAIRNESS running
```

Fig. 4. Sender, Receiver modules

The receiver sends an acknowledgement when it receives an id higher than the last message, in effect returning the id received. To cause the model to violate the given property we introduce a man-in-the-middle attacker which takes the identifier and returns it altered. Finally we declare how these processes interact in the variable declaration of the main module and establish an initial value for the variables of the process.

```
MODULE attacker
        (interceptId, n)
    VAR
      returnId : 0..n;
    ASSIGN
      next(returnId) :=
        case
          (interceptId = 2) : 4;
          1 : interceptId;
        esac;
    FAIRNESS running
```

```
MODULE main
    VAR
      s : process sender(r.receivedId, n);
      a : process attacker(s.sentId, n);
      r : process receiver(a.returnId, n);
    ASSIGN
      init(s.sentId) := 0;
      init(r.receivedId) := 0;
      init(a.returnId) := 0;
```

Fig. 5. Attacker, Main module

4.3 Properties and Checking

In this case we wish to establish a liveness property over the identifier being sent through the transmission medium. We say that at every state in the model on every computational path and for every value of the window n sent ($s.sentId$) implies that it will be eventually received ($r.receivedId$). We express this with

$$\bigwedge_i AG(s.sentId_i \rightarrow r.receivedId_i).$$

With the model and property clearly defined we apply a model checking session in `NuSMV` to ascertain the models satisfaction of the property. From `NuSMV` we receive a report explaining the fault in the model by the specification[1].

To simplify the property for program use we use equivalences and transform it to DNF. This allows the update algorithm to evaluate negation as applied to propositional atoms and evaluate possible update sets as disjunctions of terms, turning the update process into a search problem for the cheapest set of updates.

In the case of a window size equal to five, this explodes the count of propositional atoms in the formula from 10 to 160 and creates 128 terms connected by 31 nested clauses.

$$\begin{gathered} AG((\neg(s.sendId = 0) \wedge \neg(s.sendId = 1) \wedge \neg(s.sendId = 2) \\ \wedge \neg(s.sendId = 3) \wedge \neg(s.sendId = 4)) \vee \ldots \vee \\ (AF(r.receivedId = 0) \wedge AF(r.receivedId = 1) \wedge AF(r.receivedId = 2) \\ \wedge AF(r.receivedId = 3) \wedge AF(r.receivedId = 4))) \end{gathered}$$

Fig. 6. Processed formula

4.4 Counterexample to Kripke Structure Translation

In usage the returned counterexample with window size five translates to the Kripke structure seen in 7. Comparing the given states in the Kripke model

[1] The full counterexample as returned by NuSMV has been omitted for brevity.

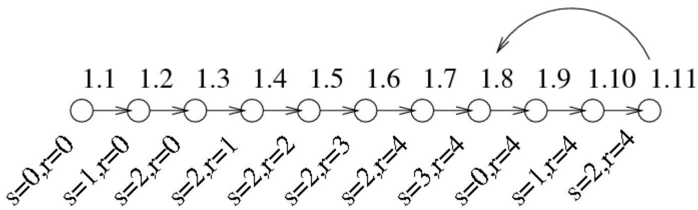


Fig. 7. local model for window size $n = 5$. s represents $s.sentId$ and r $r.receivedId$.

$1.1 \ldots 1.5$	$\models \varphi$
$1.6 \ldots 1.7$	$\not\models (\neg(s.sendId = 2)) \vee AF(r.receivedId = 2)$
1.8	$\not\models (\neg(s.sendId = 3)) \vee AF(r.receivedId = 3)$
1.9	$\not\models (\neg(s.sendId = 0)) \vee AF(r.receivedId = 0)$
1.10	$\not\models (\neg(s.sendId = 1)) \vee AF(r.receivedId = 1)$
1.11	$\not\models (\neg(s.sendId = 2)) \vee AF(r.receivedId = 2)$

Fig. 8. Counterexample states satisfying clauses of the property

against the clauses of the model property we can see where each state violates the property. With the given Kripke structure representing the local model and the property found unsatisfied in the original model we apply an update session using `lmu`.

4.5 Deriving Update

For each operation type in the property there will be semantics applied which determine how the model can be manipulated in its satisfaction. For the property in the example the ACTL tokens are AG, AF, $\vee$, $\wedge$ and $\neg$[2]. Based on the property operation and state passed as argument the update will be routed to the method created to enact the update.

In lmu_c a high level update will be performed for the AG operator using the method $lmu_{AG}()$. From here the nested disjunctive clauses will be evaluated with the $lmu_{\vee}()$ method, evaluating all of the lower level conjunctive terms to determine which is the cheaper update. $lmu_{\wedge}()$ will evaluate terms of atoms and return the cost to perform the set of updates.

At the propositional level $lmu_{AF}()$ is called in the cases where *r.receivedId* is being evaluated. The temporal operation will be satisfied when at some future state its operand is satisfied. As given earlier in the present case of a cycle the last state of the cycle will be updated.

4.6 Derived Fix

In applying the update algorithm `lmu` returned finding the most minimal update could be performed by modifying the variable *r.receivedId* at future states by

---
[2] For further information on semantic satisfaction of CTL temporal operators see [9].

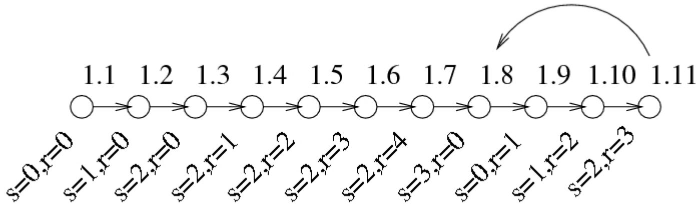


Fig. 9. Updated Local Model Fix $(\mathcal{M}', s')$

the values `0`, `1`, `2`, `3`, giving a total update count of `4`. This evaluates to the model being updated by the formula

$$(AF(r.receivedId \ = \ 2)) \wedge (AF(r.receivedId \ = \ 3)) \wedge \\ (AF(r.receivedId \ = \ 0)) \wedge (AF(r.receivedId \ = \ 1)) \\ \wedge(\neg(s.sentId \ = \ 4))$$

4.7 Results

From Figure 10 we can see that counterexample size increases in a relatively proportional manner with respect to window size. Property Atom count grows in $2n$, where n is window size. Large sets of clauses with AF tokens over AG tokens computation time explodes quickly, as a combinatorial approach is required to find the most minimal set of updates for such a model.

Window Size	Formula Atom Count	Counterexample State Count	Total State Count	Reachable State Count	Computation Time (ms)
2	4	7	8	6	344.00
3	6	13	27	27	1938.00
4	8	12	64	45	5813.00
5	10	19	125	96	22531.00
6	12	31	216	175	94531.00
7	14	32	343	288	278844.00
8	16	33	512	441	1020688.00

Fig. 10. Results of Update on sliding window n = 2..8

5 Final Remarks

In this paper we have shown an example of how ACTL tree-like local model update can be used as an effective method of generating local fixes from counterexamples for reintegration back into a candidate model. The experimental results show the feasibility of this approach in applications.

A fruitful direction to take local model update is the inclusion of constraint automata in the update procedure. In practice, system behaviour usually follow given constraints based on model actions and precedence between events occurring in the model. With added constraint information we can further restrict the types of updates that can occur at a given state and better ensure update correctness.

References

1. Berard, B., Bidoit, M., Finkel, A., Laroussinie, F., Petit, A., Petrucci, L., Schnoebelen, P., Mckenzie, P.: Systems and software verification: model-checking techniques and tools. Springer, New York (1999)
2. Buccafurri, F., Eiter, T., Gottlob, G., Leone, N.: Enhancing model checking in verification by AI techniques. Artificial Intelligence 112(1-2), 57–104 (1999)
3. Clarke, E., Lu, Y., Com, B., Veith, H., Jha, S.: Tree-like counterexamples in model checking. In: Proceedings of the 17 th Annual IEEE Symposium on Logic in Computer Science (LICS 2002). IEEE Computer Society, Los Alamitos (2002)
4. Clarke, E.M., Veith, H.: Counterexamples revisited: Principles, algorithms, applications. In: Verification: Theory and Practice, pp. 208–224 (2003)
5. Clarke Jr., E.M., Grumberg, O., Peled, D.A.: Model checking. MIT Press, Cambridge (1999)
6. Forouzan, B.: Data Communications and Networking. Mc Graw Hill, New York (2003)
7. Griesmayer, A., Bloem, R., Cook, B.: Repair of boolean programs with an application to c. In: Ball, T., Jones, R.B. (eds.) CAV 2006. LNCS, vol. 4144, pp. 358–371. Springer, Heidelberg (2006)
8. Harris, H., Ryan, M.: Theoretical foundations of updating systems. In: 18th IEEE International Conference on Automated Software Engineering (ASE 2003), p. 291 (2003)
9. Huth, M.R.A., Ryan, M.D.: Logic in Computer Science: Modelling and Reasoning about Systems. Cambridge University Press, Cambridge (2000)
10. Staber, S., Jobstmann, B., Bloem, R.: Finding and fixing faults. In: Borrione, D., Paul, W. (eds.) CHARME 2005. LNCS, vol. 3725, pp. 35–49. Springer, Heidelberg (2005)
11. Zhang, Y., Ding, Y.: Ctl model update for system modifications. Journal of Artificial Intelligence Research 31, 113–155 (2008)

Development of RP^{3}CA-EMP, a Knowledge-Based System for Applying Environmental Management Plan (EMP) in the Malaysian Construction Industry

Leila Ooshaksaraie[1], Noor Ezlin Ahmad Basri[1], Azuraliza Abu Bakar[2], and Khairul Nizam Abdul Maulud[1]

[1] Department of Civil and Structure, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia, Selangor, Malaysia
[2] Center for Artificial Intelligent Technology, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia, Selangor, Malaysia
{Leila.Ooshaksaraie,Noor.Ezlin.Ahmad Basri, Azuraliza.Abu.Bakar, Khairul.Nizam.Abdul.Maulud}nakisa_ooshaksaraei@yahoo.com

Abstract. The Environmental Management Plan (EMP) is a tool for preventing environmental pollution from a proposed construction project. EMP development requires specialized construction site data and information from multiple sources. Recently, knowledge-based system (KBS) has been used extensively in the environmental field, especially in cases where human expertise and data are limited. In this study, a KBS computer application was developed to guide the application of EMP for housing and new township development projects in Malaysia. This system, called River Pollution Prevention Plan during Construction Activity- Environmental Management Plan, or "RP^{3}CA-EMP" was developed in the Microsoft Visual Basic environment. RP^{3}CA-EMP predicts the impacts of housing and new township development projects on surface water and details necessary mitigation measures. Additionally, RP^{3}CA-EMP is an educational tool for individual developing EMP for housing and new township development projects.

Keywords: Water Resources Protection, Construction Industry, Knowledge-Based System, Microsoft Visual Basic.

1 Introduction

The rapid urbanization occurring in most developing countries presents major environmental problems due to city expansion, conurbation development, and population growth. The problems that are typically associated with such large urban centres are expected to appear in developing countries [1, 4]. In most major cities in Asia , Africa, and Latin America, a significant proportion of the population lives in shelters and neighbourhoods lacking both adequate access to clean water and system for the safe disposal of solid and liquid wastes [9, 14]. Although water is a critical natural

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 22–31, 2010.

resource that plays a vital role in national progress and development [2], surface water quality may be adversely affected by construction activities [13] such as trenching, surface preparation, demolition, pouring concrete, and material and waste management [8]. Construction projects also require managing labour and materials, preparing the land and foundation, installing mechanical and electrical systems and appliances, erecting and enclosing the walls and roof, applying interior and exterior finishes and trimming [10, 16]. Typical pollutants resulting from these activities include displaced soil, sediment, debris, wood products, concrete slurry, hazardous materials, and recyclable materials [8].

Urban development is particularly rapid in Malaysia and has caused adverse environmental effects. Specifically, construction sites, especially those that have been cleared of vegetation but have not yet been developed, frequently lose excessive amount of soil. In some cases, this has caused deterioration of water courses [5, 6]. In Malaysia, the high mean annual rainfall and frequent intense storm make water erosion the most significant cause of soil loss, and the rate of this erosion increases when vegetation disturbed or removed [21]. Malaysia experiences more intense and frequent rainstorms than most developed countries and therefore requires more stringent control measures with both structural and non-structural components [6].

Environmental study is a time-consuming process that requires special tools or support systems because a wealth of dependent and independent variables must be considered. Various computer-assisted systems can help to meet the challenges of collecting, processing, analyzing and reporting this complex environmental information [15, 19]. Knowledge-based systems or expert systems have been used for environmental planning, impact assessment, regional environmental impact analysis, environmental assessment, and environmental management since the 1980s. These systems provide advice and support for decision makers based on a database of expert knowledge in combination with heuristic (rule of thumb) reasoning, [3, 19]. Expert systems are promising technologies for enhancing data, and information management, and access to specialized expertise. Therefore, they appear to be well suited for many of the tasks associated with environmental management study. In this application, they could provide structured approach to environmental studies and help users cope with large volumes of environmental management studies [12, 19].

In this study, a KBS called the River Pollution Prevention Plan during Construction Activities (RP^{3}CA) was developed with a particular emphasis on housing and new township development projects in Malaysia. RP^{3}CA has three sub-systems: Environmental Management Plan (EMP), River Water Quality Monitoring (RWQM), and Construction Site Inspection (CSI). The EMP described in this paper has the capacity to advise the end-users regarding Environmental Management Plan for the construction industry, especially related to housing and new township development projects in Malaysia.

2 Methodology

RP^{3}CA-EMP was developed according to standard KBS development methods. The KBS design and development process is guided by a five-step process: task analysis, knowledge acquisition, system development, expansion and refinement, and verification and validation [11, 7, 20, 22].

2.1 Knowledge Acquisition

The Best Management Practices (BMPs) identified in this study were drawn from multiple sources of expertise in the field of construction engineering in order to minimize river pollution during construction activities. Pre-requisite basic knowledge was acquired from textbooks and manuals and served as the basis for the development of the initial prototype modules [1]. Journal publications and conference proceedings were used as recent and more specialized data sources within the domain. Current research publications facilitated a more practical approach to various environmental concerns, including technological considerations, legislative enforcement, and finances aspects [1]. Experts in the field were contacted directly to obtain additional information. This interaction consisted of a series of systematic consultations, over a period of a few months. According to Yialouris [24], two meetings were organized with three experts to elicit opinions about developing an integrated expert geographical information system for soil suitability and soil evaluation. In this study, three meetings were organized with ten experts to discuss construction management, soil erosion, sedimentation, environmental management, and BMPs. After the third meeting, the experts were asked to provide their recommendations on the following topics:

- The construction stages and activities that should be considered during the development of the KBS,
- The properties of construction sites that are most important for preventing river pollution,
- BMPs that should be considered for construction sites, and
- BMPs and construction sites that should be inspected.

Expert responses to these questions are summarized in Table 1, 2, and 3.

Table 1. Recommended stages of construction activity

Construction stages	Substage
Site construction facilitie	Access road and stream crossing, Temporary building, Site utility
Site clearance	Demolition, Site clearance
Site formation	Earth work, Import and export of fill, Drainage work, Road development
Infrastructural work	Operation of heavy equipment, Construction of building and utility

Table 2. Recommended construction site characteristics for pollution prevention

Construction site information	
Season condition	Dry, rainy
Slope of streambed for entrance the site	$< 1:3$, $\geq 1:3$
Flow condition of site upslope	Concentrate, Non-concentrate
Drainage area	< 0.4 , 0.4-2, 2-4, > 4 ha
Soil type	C, D, F, More than one type
Waste type	Steel, Aluminium, Gypsum plasterboard, timber, concrete, bricks and tiles, plastic, glass, carpet
Flow condition at site	Sheet, rill, gully, channel
Duration for soil stabilization	< 3, 3-12, ≥ 12 months
Slope range for soil stabilization	$\leq 1:4$, 1:4 – 1:2, $\geq 1:2$
Flow velocity of surface runoff	< 0.6 m/s, ≥ 0.6 m/s
Slope of streambed within the site	$< 15^{\circ}$, $\geq 15^{\circ}$
Slope range of cut and fill area	$< 1:2$, $\geq 1:2$

Table 3. Recommended Best Management Practices (BMPs) for construction sites

BMPs	
BMPs for contractor activities	Construction practice, Material management, Waste management, Vehicle and equipment management, Contractor, employee, and subcontractor training
BMPs for erosion and sedimentation	Site planning consideration, Soil stabilization, Tracking control, Velocity reduction structures, Sediment trapping structures

The primary goal of this stage of the study was to acquire expert knowledge based on information obtained during the previous stage and a series of structured interviews. Ten distinct sets of expert questionnaires were designed in order for the experts to share answers or other information, and interviews were organized separately with each expert.

In general, knowledge acquired from experts includes some conflicting information. Conflicts were resolved by applying Certainty Factors (CF), which quantify the confidence of an expert's belief. The minimum value for the certainty factor was -1.0 (definitely false), and the maximum value was +1.0 (definitely true). A negative value represented a degree of disbelief, and a positive value represented a degree of belief. CF values are interpreted as "Definitely not" (-1.0), "Almost certainly not" (-0.8), "Probably not" (-0.6), "Maybe not" (-0.4), "Unknown" (-0.2 to +0.2), "Maybe" (+0.4), "Probably" (+0.6), "Almost certainly" (+0.8), and "Definitely" (+1.0) [17].

In KBS, the knowledge base consists of a set of rules that have a common syntax of conjunctive and disjunctive rules. Conjunctive forms are written as follows: IF <evidence E_1> AND <evidence E_2> … AND <evidence E_n> THEN <hypothesis H> {cf}. The net certainty of hypothesis H is established as shown in Equation 1. Disjunctive rules are written as follows: IF <evidence E_1> OR <evidence E_2> … OR

<evidence E_n> THEN <hypothesis H> {cf}. The certainty of hypothesis H is established by Equation 2 [17].

$$cf\ (H, E_1 \cap E_2 \cap \ldots \cap E_n) = \min\ [cf\ (E_1), cf\ (E_2), \ldots, cf\ (E_n)] \times cf \tag{1}$$

$$cf\ (H, E_1 \cup E_2 \cup \ldots \cup E_n) = \max\ [cf\ (E_1), cf\ (E_2), \ldots, cf\ (E_n)] \times cf \tag{2}$$

2.2 Task Analysis

EMPs for housing and new township development projects incorporate many distinct tasks in diverse subject areas. Each of these tasks defined a knowledge base module that contributed to the system. The tasks used to organize BMPs are as follows:

Task 1: access road and stream crossing
Task 2: temporary building
Task 3: site utilities
Task 4: demolition
Task 5: clearance
Task 6: earth work
Task 7: import and export of fill
Task 8: drainage work
Task 9: road development
Task 10: operation of heavy equipment
Task 11: construction of utilities and buildings

2.3 Design

Microsoft Visual Basic software (VB) was used to develop the RP^3CA-EMP and a rule-based technique was used to represent knowledge. The RP^3CA-EMP reasoning process was controlled by a forward-chaining control strategy which represented collections of facts. In this control strategy, rules are formatted to read "IF 'condition is true', THEN 'perform action'." If the conditions are true within the context of the available knowledge base, they will be stored in the fact portions of the system knowledge engine to perform procedure [18].

A pilot prototype system was built to validate the project [11] and to provide the guidance for future work. Initially, it was used to solve one of the domain tasks in order to test and explore the concepts of problem definition, scoping and domain representation [1].

Next, the system was further refined to meet the project objectives. Expert knowledge was acquired and codified in the form of production rules. The overall RP^3CA-EMP system consists of the BMPs module, data input, and accessories. The BMPs module consists of eleven sub-modules that contain specialized knowledge bases for the system. The data input object facilitates collection of the input that is required to begin the RP^3CA-EMP process, and the accessories object provides support features for the system. Interface development began with RP^3CA-EMP prototype development. A standard Microsoft Windows interface was used so that the users familiar with similar programs could easily adapt to the interface [1]. VB provides a Session Window in which the system developer can create images that allow the end-user to interact with the knowledge base. VB also permits user interaction through forms. The Windows dialogue box guides the user through system operation [18], by presenting the user with a series of questions. The answers provided by the user are used to determine the factor or element [23]. System development was guided by information obtained through knowledge elicitation sessions and tests.

The basic organizational structure of the computer-assisted EMP system is given in Fig. 1.

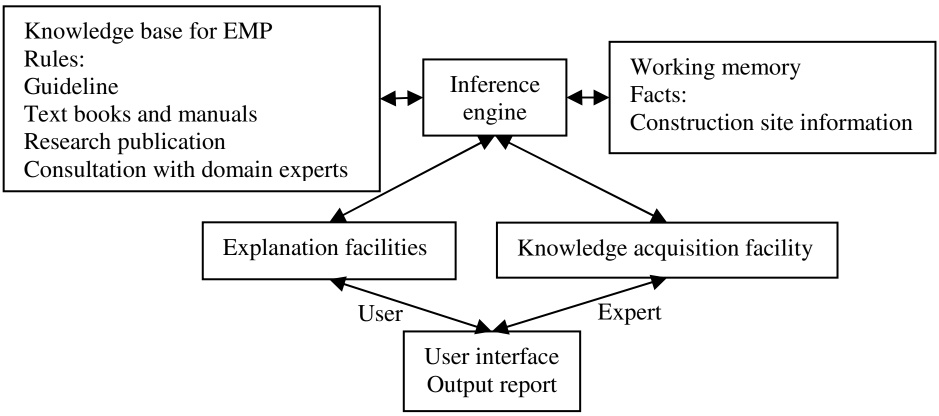


Fig. 1. Basic organizational structure of a computer-assisted EMP system

2.4 Testing and Validation

A construction site in Malaysia was used as a case study for testing and validating modules of the RP^{3}CA-EMP. The objective of the case study was to evaluate the performance of the RP^{3}CA-EMP consultation process when it was applied to a construction site. An Erosion and Sediment Control Plan (ESCP), and EMP must be submitted to Malaysian Drainage and Irrigation Department (DID) by a qualified consultant for all housing and new township development projects. The results of the RP^{3}CA-EMP were found to be comparable and similar with the project ESCP, and EMP, indicating that the RP^{3}CA-EMP performed as well as the human experts.

Finally, the user-friendliness of the user interface was evaluated. The main screen of the RP^{3}CA-EMP consists of several buttons with images similar to other software applications for the user to select with a mouse. Therefore, very few computer commands are involved, and the user is only required to perform a minimum amount of work with the keyboard. The RP^{3}CA-EMP also provides a teaching feature that is designed to aid an inexperienced engineer by guiding him or her through the process of selecting BMPs, preliminary design steps, and explaining inspection and maintenance requirement associated with BMPs. These features are available to the user through the *User Guide*, and *General Information* menus. Furthermore, the user is able to access the help features if he or she is having difficulty at any stage during the consultation process [11].

3 The RP³CA-EMP System

System output is primarily intended to provide engineers, consultants and EMP decision-makers with the following information:

a. General information about appropriate EMP content,
b. Potential environmental impacts of housing and new township development projects on surface water,
c. Information about mitigation measures for minimizing river pollution during construction activities,
d. BMPs for different stages of construction activities based on construction site characteristics,
e. A BMP checklist for site inspection, and
f. Locations for BMPs in a Geographical Information System (GIS) map

The user is advised on project steps through a series of choices from RP³CA-EMP. At each consultation step, the user can choose any of the construction stages presented in Figure 2, which illustrates the temporary building stage. After receiving data, the system delivers immediate results as shown in Figures 3, and 4. The system then presents all results and a preliminary design of the recommended BMPs in a report form.

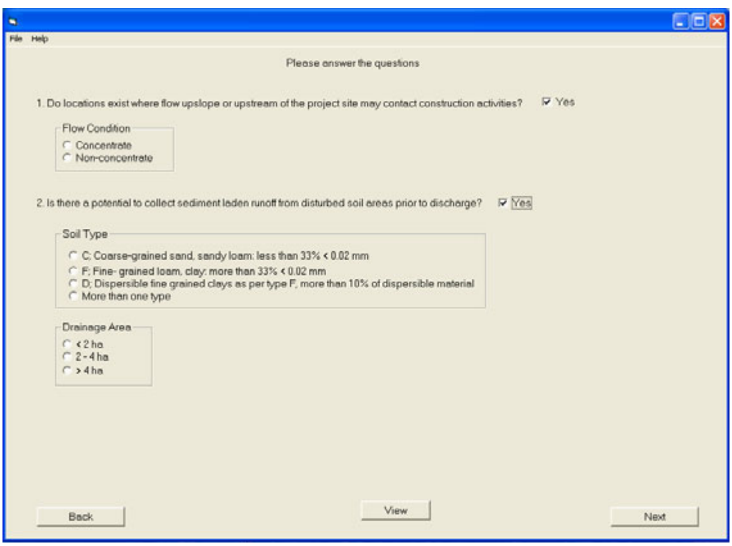


Fig. 2. Data input for the selected construction stage

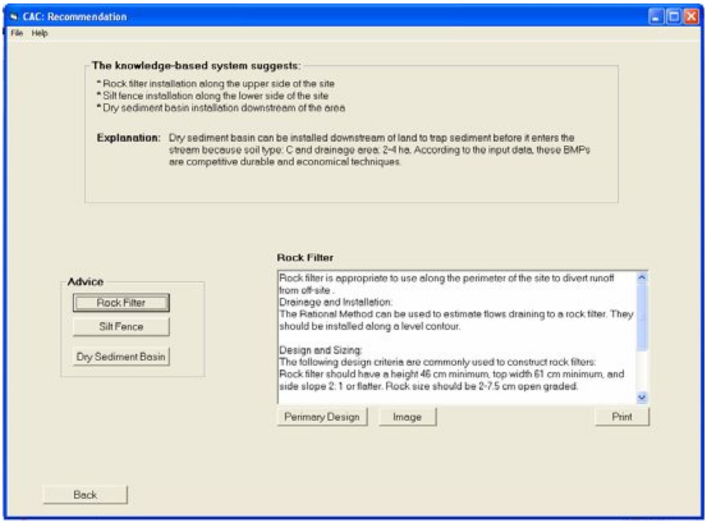


Fig. 3. RP^{3}CA-EMP recommendations for selection of BMPs

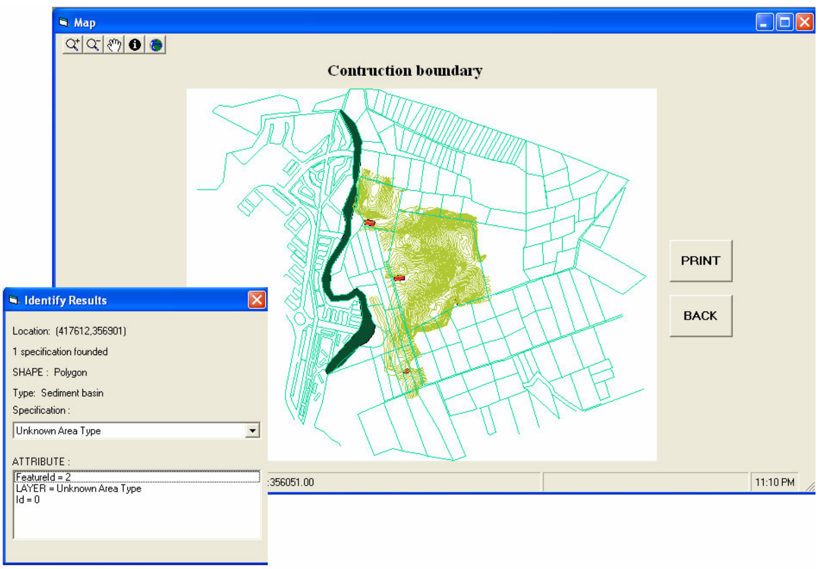


Fig. 4. Locations of BMPs in GIS format

4 Conclusions

The construction industry contributes significantly to river pollution. Therefore, it is essential to apply the most appropriate techniques and BMPs at construction sites to minimize river pollution. This paper presented a knowledge-based system, called the

RP^3CA-EMP that was developed to minimize river pollution during housing and new township development projects in Malaysia. This system provides advice based on input data and, together with its "help" feature, serves as useful educational tool. The system database compiled texts, expert knowledge about common BMPs, figures, GIS maps, potential environmental impacts, mitigation measures, and site inspection procedures. All of this information can be applied by both experts and non-experts during housing and new township development projects. The system accounts for relevant properties of the project area in order to generate site-specific recommendations and to guide the selection of appropriate BMPs. This system will be encourage the creation of EMPs in the construction industry and will thereby help relevant agencies make more accurate and objective decisions that will reduce river pollution during housing and new township development projects.

Acknowledgement

The authors would like to thank the Malaysian Ministry of Science, Technology and Innovation; this project was completed under research grant 01-01-02-SF0363.

References

1. Ahmad Basri, N.E.: An Expert System for the Design of Composting Facilities in Developing Countries, PhD thesis, School of Civil Engineering, University of Leeds (1999)
2. Alizadeh, A., Keshavarz, A.: Status of Agricultural Water Use in Iran. J.Water Conservation, Reuse and Recycling, 94–105 (2005)
3. Barrow, C.J.: Environmental and Social Impact Assessment. Centre for Development Studies. University of Wales/Wiley, Swansea/New York (1997)
4. Bartone, C.R., Bernstein, J.D.: Improving Municipal Solid Waste Management in Third World Countries. J. Resources, Conservation and Recycling 8, 43–54 (1993)
5. Department of Environment: Guidelines for Prevention and Control of Soil Erosion and Siltation in Malaysia, Malaysia (2008)
6. Department of Irrigation and Drainage.: Urban Stormwater Management Manual for Malaysia, Malaysia (2001)
7. Dym, C.L.: Implementation Issues in Building of Expert System. In: Maher, M.L. (ed.) Expert Systems for Civil Engineers: Technology and Applications, pp. 35–48. ASCE, New York (1987)
8. Environmental Services Department.: Jurisdictional Urban Runoff Management Program. San Diego Unified Port District, San Diego (2005)
9. Hardoy, J.E., Mitlin, D., Satterthwaite, D.: Environmental Problems in Urbanizing World: Finding Solutions for Cities in Africa, Asia and Latin America. Eathscan, London (2001)
10. Hassell, S., Wong, A., Houser, A., Knopman, D., Bernstein, M.: Building Better Homes, RAND (2003)
11. Islam, M.D.N.: Development of an Expert System for Making High Performance Concrete Using Hybrid Knowledge Representation System, PhD thesis, Faculty of Engineering, Universiti Kebangsaan Malaysia (2004)
12. Lohani, B.N., Evans, J.W., Everitt, R.R., Ludwig, H., Carpenter, R.A., Liang Tu, S.: Environmental Impact Assessment for Developing Countries in Asia. Overview Asian Development Bank 1, 356 (1997)

13. Barrett, M.E., Malina Jr., J.F., Charbeneau, R.J., Ward, G.H.: Effects of Highway Construction and Operation on Water Quality and Quantit. In: An Ephemeral Stream In The Austin, Texas Area, Austin, Technical Report (1995)
14. Mruma, R.O.: Implementation of the National Environment Policy: a Case of the Local Government Authorities in Dar Es Salaam City-Tanzania, MS thesis, Faculty of Social Sciences, University of Bergen (2005)
15. Muthusamy, N., Ramalingam, M.: Environmental Impact Assessment for Uurban Planning and Development Using GIS. In: Proceedings of the Third International Conference on Environment and Health, Chennai, India (2003)
16. NAHB Research Center.:Cycle-Time Reduction in the Residential Construction Process, Upper Marlboro, Md (1993)
17. Negnevitsky, M.: Artificial Intelligence: A Guide to Intelligent System, 2nd edn. Addison-Wesley, Reading (2005)
18. Pauziah, H.G., Mohd, K.Y., Latifah, M., Mohamed, D.: Knowledge-Based System for River Water Quality Management. J. Scientific Research 33(1), 153–162 (2009)
19. Say, N.P., Yucela, M., Yılmazerb, M.: A Computer-based System for Environmental Impact Assessment (EIA) Applications to Energy Power Stations in Turkey: C- EDINFO. J. Energy Policy 35, 6395–6401 (2007)
20. Terry, P.C.: Expert Systems: Manager's Perspective. J. Management in Engineering 7(1), 119–131 (1991)
21. University of Malaya Consultancy Unit.: Environment Impact Assessment Guideline for Housing and New Township Development Project, Department of Environment, Malaysia (2003)
22. Waterman, D.A.: A Guide to Expert Systems. Addison-Wesley, Reading (1986)
23. Yang, B., Li, L.X., Xie, Q., Xu, J.: Development of a KBS for Managing Bank Loan Risk. Knowledge-Based System, 299–302 (2001)
24. Yialouris, C.P., Kollias, V., Lorentzos, N.A., Kalivas, D., Sideridis, A.B.: An Integrated Expert Geographical Information System for Soil Suitability and Soil Evaluation. J. Geographic Information and Decision Analysis 1(2), 89–99 (1997)

Reasoning with Multiple Points of View: A Case Study*

P. Bouché, C. Zanni-Merk, N. Gartiser, D. Renaud, and F. Rousselot

LGECO – INSA de Strasbourg – 24 bld de la Victoire – 67084 Strasbourg – France
{philippe.bouche,cecilia.zanni-merk,nathalie.gartiser,
dominique.renaud,francois.rousselot}@insa-strasbourg.fr

Abstract. This article presents our approach to reasoning with diversified and voluminous knowledge sources that can eventually be, contradictory. In fact, knowledge sources coming from management sciences are inherently rich and, sometimes, conflicting. We choose to exploit the entire range of this diversity to improve business advice to small and medium enterprises (SMEs).

Keywords: multi-agents system, ontologies, rule-based reasoning.

1 Introduction

One of the major difficulties encountered by the smallest companies, especially today, in a crisis context, is how to manage their evolution. To do so, they need the capacity to perform a global analysis of all their aspects (economic, production, organization, human resources, sales, etc.). They also need to keep the sufficient distance to put this analysis in the perspective of their evolution. Change management becomes a key success factor for many firms facing strong competition.

Very often, the SMEs getting involved in this approach look for the help of consultancy services. In this context, there is a recurrent question that arises: how to access existing knowledge to diagnose the SME while thinking about its evolution.

Here, there are several important aspects regarding knowledge management. The main problem is that the volume of, both theoretical and "expert", knowledge is huge and sometimes, much more detailed than needed. We face a triple issue about knowledge capitalization: structuring to manage large quantities of knowledge, organizing it to permit different access levels, and linking different sources to develop valuable knowledge.

This knowledge management may lead to two different types of results. Firstly, it may allow the consultant and the company to more easily put typical solutions into practice. These solutions may already be known, but they might not have been implemented in the company yet. This fact will allow the company to go further in its evolution, integrating innovation in this way, even if this innovation is minor because it only concerns the company itself. Secondly, efficient knowledge management with the help of reasoning tools might allow the consultant to propose original solutions, thus generating high level innovations, because they are radically new.

---

* The authors would like to thank Mr. Henri-Pierre Michaud from AEM Conseil for his support and funding.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 32–40, 2010.

This article is structured as follows: Section 2 presents the project itself along with the ontological framework and the architecture of our development. Section 3 describes an example to show our results in a real industrial case study and, finally, Section 4 presents our conclusions.

2 The Project

MAEOS is a project dealing with the modelling of the organizational and strategic development support of SMEs. The main objective of MAEOS is to improve the efficiency and performance of business advice to SMEs. The whole description of the project can be found at [9].

To achieve this objective, a multidisciplinary team was created. Three main research areas are represented: artificial intelligence, software engineering and management sciences. This work aims at establishing a set of methods and software tools for analysis and diagnosis of SMEs. The software tools will have to evolve according to the state of the art of SMEs and, in particular, their administrative or legal environments. In addition, they must also be able to reflect the richness and contradictions that are inherent to the models coming from management sciences. Finally, they must permit the consultants to access knowledge coming from diverse sources in an efficient and pertinent way.

As presented in the introduction, one of the major difficulties is the manipulation of huge quantities of knowledge. The implementation choices were, therefore, directed by these issues: a multi-agent system[1] or MAS [12] is at the heart of the development, where each agent reasons (following a rule-based approach) with a specific modular ontology.

2.1 Ontological Framework of Our Development

An ontology is a formal explicit description of concepts in a domain of discourse [7]. Associated to an ontology, we may also define a set of rules that will permit reasoning about facts. Rules are expressed as implications.

An ontological study was conducted to provide the theoretical foundations necessary for the development. Several ontologies have been studied.

Our main sources were the MASON [5], TOVE [3], [4] and ENTERPRISE [11] ontologies. Some parts of specific ontologies have also been considered. These ontologies cover different areas such as Professional Learning and Competencies with the ontology of FZI-Karlsruhe [10], organization modelling with UEML-1 [1], or Service Oriented Architectures with the SOA Open Group ontology; among others.

Beyond the use of existing ontologies, we have developed our own about certain relevant fields for our SME context (organization, quality, production, innovation, etc.). Nevertheless, we risk being confronted by a double issue concerning contradictions at the knowledge level: they may appear among different knowledge sources and also within the same knowledge source. It is because of these issues that we have decided to develop separate ontologies and to separately process these two sources of contradictions.

[1] Like a kind of a committee of experts.

Firstly, we have developed an ontology about the organization models based on the main works of Henry Mintzberg [6]. This ontology integrates the concepts that describe the company structure and models, the relationships among concepts and the restrictions to those concepts according to the company characteristics (its size, for example, which is relevant for this project). We have chosen to use this source because it is a clear reference in the organizations theory field, at least, at the concept level. The works of this author have been widely quoted, commented and refined.

Secondly, we have developed an ontology about production systems based on [2]. This reference is a choice of our industrial partner.

2.2 Architecture of the Multi-Agent System

The system is being implemented as a MAS. Each agent will reason with a specific ontology and the set of associated rules. We use a blackboard for communication among agents (Fig. 1)

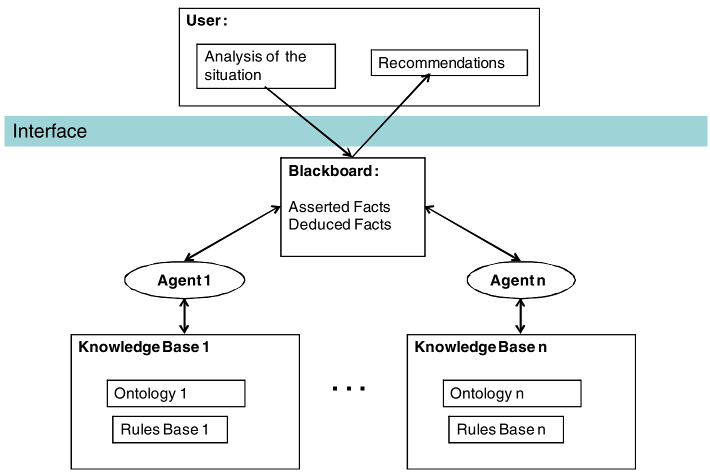


Fig. 1. The MAS system

The user (a consultant, for instance) translates his own situation analysis into “asserted facts” through the interface.

Each agent has the same behaviour: It examines the blackboard to identify the facts corresponding to concepts it knows. In this case, it launches reasoning about these facts and, in the end, it writes the deduced facts on the blackboard.

Finally the consultant uses these deductions to organize them into recommendations.

This situation implies the production of many pieces of results related to a limited topic. To be well understood by an external user, all the produced results must be combined, at least, by topic. In previous publications [8], we have presented an approach to aggregate results coming from different ontologies. This approach includes a strategy to solve the many limitations resulting from the use of ontologies whose contents are closely related.

3 Example: An Industrial Case Study

This is a real case study that has been manually solved by our industrial partner. Our goal here is to test if our reasoning approach based on modular ontologies produces the same results or others.

The case study is about a SME whose main activity is metal sheet bending. The initial facts identified by the consultant were about problems in organization and security (for confidentiality reasons we cannot describe the situation more thoroughly). These problems were reflected by the disorganization in work-orders, the chaotic layout, and the misinformation among actors. The situation had begun to cause accidents.

With this situation in mind, we have decided to test the example with only two agents: Agent 1 will reason on the production ontology based on [2]; while Agent 2 will reason on the organizations theory ontology based on [6].

3.1 Initialisation

The first task of the consultant in the company is to have an overview of the situation. This overview is obtained through one or several interviews and permitted us to identify a set of concepts and facts associated to one of our ontologies.

Facts will be expressed as predicates in first order logic. Arguments to predicates are concepts in our ontology. Facts will be written in lower case while concepts will be capitalized.

The analysis of the initial situation of this company led to the following facts:

```
intermittent-production
produce-what-is-sold
small-serial-production
raw-material(Sheet)
delete(Operation-not-generating-Saleable-Value)
```

Both agents examine the contents of the blackboard. Agent 1 is able to identify some of its own facts, so it launches its reasoning tasks. Agent 2 remains stand-by as it does not recognize any asserted fact as its own.

3.2 Operation of Agent 1

The facts identified by Agent 1 allow the identification of the relationships and rules in the production ontology that may be launched to predict plausible evolution, thanks to the inference engine. The initial set of facts permits the launching of the following rules by Agent 1 (see Fig. 2 for the inference tree):

$R_{C1.1}$: delete(Operation-not-generating-Saleable-Value) ⇒
simplify(Physical-Flow)

$R_{C1.2}$: intermittent-production ⇒
decrease(Stock) AND
decrease(Work-In-Process)

$R_{C1.3}$: intermittent-process ⇒ is-difficult(Good-Layout)

$R_{C1.5}$: small-serial-production ⇒ functional-layout

The rules launched by Agent 1 produce new facts that are written on the blackboard.

```
decrease(Stock)
decrease(Work-in-Process)
reactivity
functional-layout
```
(This fact is equivalent to the fact `homogeneous-section-layout`[2])

We can iterate this process:

- Find rules which can be launched,
- Deduce new facts (for example, `decrease(Delay)` after applying rule $R_{C2.1}$)

When no more rules can be launched by Agent 1, the process is finished and the agent is able to write its final set of deduced facts on the blackboard:

```
decrease(Delay)
decrease(Work-In-Process)
need(Circulation-Graph)
simplify(Product-Flow)
```

3.3 Links between the Ontologies

This first analysis was made with the production ontology only.

But, ontology experts have identified semantic equivalences among concepts and predicates in both ontologies.

For instance, `Reactivity` (defined in the production ontology) is semantically equivalent to `Speed-of-Response` in the organization ontology. Therefore, we can also define semantic equivalence between facts:

reactivity ≡ increase(Speed-of-Response)[3]
¬reactivity ≡ decrease(Speed-of-Response)

After this first reasoning step, Agent 1 has written the fact `reactivity` on the blackboard. A "translation" agent will be able, then, to write `increase(Speed-of-Response)` and, with this new fact, Agent 2 can begin its work.

3.4 Operation of Agent 2

Agent 2 is able to trigger $R_{M1.1}$:

$R_{M1.1}$: increase(Speed-of-Response) ⇒ increase(Unit-Grouping)AND increase(Behaviour-Formalization)

---

[2] This is because the corresponding concepts `Functional-Layout` and `Homogeneous-Section-Layout` have been defined as equivalent classes in OWL.

[3] A concept is not the same thing as a fact, this is why we have written `reactivity` in lower case, to identify it as a fact or predicate.

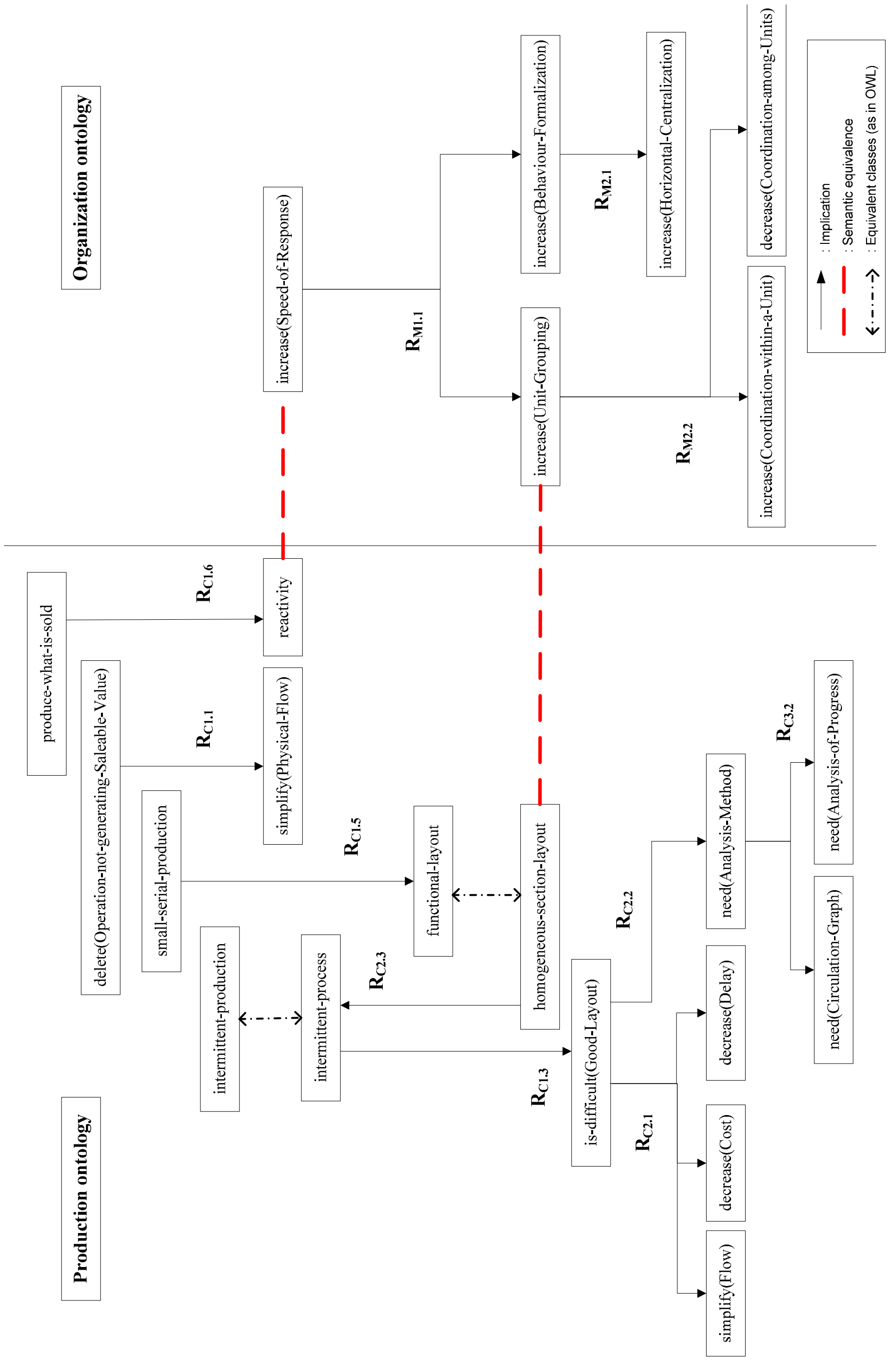


Fig. 2. A sub-set of the inference tree for the agents

This rule leads to the definition of two facts which will be written on the blackboard. There are, then, new rules that can be launched by the agent:

```
R_M2.1: increase(Behaviour-Formalization)⇒
          increase(Horizontal-Centralization)
R_M2.2: increase(Unit-Grouping)⇒
          decrease(Coordination-among-Units) AND
          increase(Coordination-within-a-Unit)
```

The final set of deduced facts by Agent 2 includes:

```
increase(Unit-Grouping)
increase(Behaviour-Formalization)
increase(Horizontal-Centralisation)
increase(Coordination-within-a-Unit)
decrease(Coordination-among-Units)
```

3.5 Aggregation of Results

As explained in Section 2.2, reasoning with multiple ontologies implies the production of many pieces of results related to a limited topic.

The strategy we proposed to tackle this problem [8] is articulated around two key points: the choice of the combining method and the partitioning of ontologies.

In this case study and using these two ontologies, the system has deduced the following set of facts:

```
increase(Unit-Grouping)
increase(Behaviour-Formalization)
decrease(Delay)
decrease(Work-In-Process)
need(Circulation-Graph)
```

The MAS was then able to produce five sets of linked concepts. For each concept in the deduced facts, a close exploration in the ontologies is done, to identify the set of linked concepts to that one.

These sets of concepts are related to the following topics: `Flow`, `Reactivity`, `Type-of-Production`, `Delay` and `Stock`. The use of this strategy of aggregation reduces the number of sets of concepts and adds new ones. The overall number is therefore reduced to two: `Flow`, `"Reactivity / Type-of-Production / Delay / Stock"`. This situation could be further improved by the addition of other instances of concepts. These new instances are suggested to the consultant for identification and validation. The consultant can then choose to extend the number of instances to complement its initial model or restart the MAS with a new query. This fact therefore needs interaction between the system and the consultant. This part of our work is focused on a GUI solution and is currently being evaluated.

3.6 Recommendations

The set of facts after aggregation is at the heart of the report that the consultant will hand to the company.

The case we have described in this example has already been manually processed by our industrial partner without the help of the knowledge based system. We have compared the consultant's conclusions and the system's conclusions.

It is interesting to see that there is a set of proposals common to both reports (`increase(Unit-Grouping)` or `need(Circulation-Graph)`, for instance) but also that the knowledge based system has been able to propose new ideas, such as `need(Analysis-of-progress)`, or `increase(Behavior-Formalization)`. These new propositions have been positively evaluated by the consultant.

Even if this example remains quite simple, the interest of the approach is the use of close modular ontologies to propose new solutions to the consultant and assist him in his work.

4 Conclusions

In this article we have presented our first results about a knowledge-based system to assist consultants in their analysis and diagnosis of SMEs. This system consists in a set of knowledge bases whose contents are heterogeneous and this fact raises several issues.

Firstly, there is the management of knowledge structures and backgrounds. It takes the form of complex cycle phases including acquisition, use and maintenance. There is, also, the manipulation of huge quantities of knowledge by the consultant; his expectation is to obtain innovative solutions with several alternatives coming from the system. Also, as he has to be able to explain his recommendations to his client, he needs to understand the suggestions and analysis provided by the system.

To solve these issues, our approach is to split the different fields of knowledge into smaller modular and homogeneous bases. This approach is supported by the exclusive use of formal ontologies. This point is motivated by the possibilities of manipulation of knowledge and a strict framework imposed by this type of ontology. In this way, we hope to combine the advantages of rigorous knowledge manipulation, adaptability to particular cases and readability of the results.

This approach by close modular ontologies is possible because the ontologies reflect different points of view on the same reality and because these discrepancies among points of view (usual in companies) only question the logical coherence in certain parts of the model. Arbitration among the different points of view will be, in the end, the responsibility of the consultant.

The main objective of our project is to improve the efficiency and performance of business advice to SMEs (and neither to propose solutions to a problem, nor to have the consultant's work done by the software).

To do this, the targeted system aims at providing suggestions to help address those areas in which a consultant is less efficient.

All the suggestions are related to the point of view of an expert; these suggestions may be close, complementary and sometimes contradictory.

To introduce the notion of multiple points of view, each software agent is associated with a particular knowledge base (KB) and ontology. This situation implies the production of many pieces of results related to a limited topic. To be well understood by an external user, all the produced results must be combined, at least, by topic.

The innovative results come when some parts of these different KBs are combined at the end of an automatic study.

The main interest of our development is the capacity to aggregate results coming from different ontologies or previous case-studies (no matter if the sources are obsolete), as we are more interested in the differences among those sources rather than in the consensual aspects. These differences would lead the consultant to a more thorough reflection about the results.

References

1. Berio, G., Jaekel, F.-W., Mertins, K.: Common Enterprise Modelling Framework for Distributed Organisations. In: 16th IFAC World Congress, Prague, Czech Republic (2005)
2. Courtois, A., Martin-Bonnefous, C., Pillet, M.: Gestion de Production. Editions d'organisation (1989), ISBN 2708136488
3. Fox, M., Grüninger, M.: Enterprise Modeling. AI Magazine, 109–121 (1998)
4. Fox, M.: The TOVE project: A Commonsense Model of the Enterprise. In: Belli, F., Radermacher, F.J. (eds.) IEA/AIE 1992. LNCS (LNAI), vol. 604, pp. 25–34. Springer, Heidelberg (1992)
5. Lemaignan, S., Siadat, A., Dantan, J.-Y., Semenenko, A.: MASON: A proposal for an Ontology of Manufacturing Domain. Distributed Intelligent Systems, Collective Intelligence and Its Applications. In: IEEE Workshop on DIS 2006, 15-16, pp. 195–200 (2006)
6. Mintzberg, H.: The Structuring of Organizations. Prentice-hall, Englewood Cliffs (1979), ISBN 0-13-8555270-3
7. Noy, N.F., McGuinness, D.L.: Ontology Development 101: A Guide to Creating your First Ontology (2001), `http://www.ksl.stanford.edu/people/dlm/papers/ontology101/ontology101-noy-mcguinness.html`
8. Renaud, D., Zanni-Merk, C., Bouché, P., Gartiser, N., Rousselot, F.: A strategy for reasoning with close knowledge bases: Application to knowledge bases on SMEs. In: Data Mining and Business Intelligence. Publishing House of the University of Economics, Wroclaw, Edition en 2010 (2010)
9. Renaud, D., Bouché, P., Gartiser, N., Zanni-Merk, C., Michaud, H.P.: Knowledge Transfer for Supporting the Organizational Evolution of SMEs A Case Study. In: International Conference on Innovation through Knowledge Transfer 2009 (InnovationKT), Hampton Court Palace, London (2009)
10. Schmidt, A., Kunzmann, K.: Sustainable Competency-Oriented Human Resource Development with Ontology-Based. In: Cunningham, M., Cunningham, P. (eds.) Competency Catalogs, eChallenges (2007)
11. Uschold, M., King, M., Moralee, S., Zorgios, Y.: The Enterprise Ontology. The Knowledge Engineering Review 13, 31–89 (1998)
12. Wooldridge, M.: An introduction to MultiAgent Systems, 2nd edn. John Wiley & Sons, Chichester (2009), ISBN-10: 0470519460

Syntax and Semantics for Business Rules

Xiaofan Liu[1 2], Natasha Alechina[1], and Brian Logan[1]

[1] School of Computer Science, University of Nottingham, Nottingham, NG8 1BB, UK
[2] School of Computer and Communication, Hunan University, Hunan, 410081, P.R. China
{lxx,nza,bsl}@cs.nott.ac.uk

Abstract. Business processes are increasingly being expressed in declarative terms, as *business rules* which express obligations concerning conduct, action, practice or procedure regarding a particular business activity or sphere of operation. However, the leading standard for business rules, Semantics for Business Vocabulary and Business Rules (SBVR), leaves certain decisions concerning the precise syntax and semantics of business rules unspecified, for example, the scope and nesting of modalities and first order quantifiers, and the precise semantics of alethic and deontic modalities. In this paper, we propose a precise syntax and semantics for SBVR and present some complexity results for business rules with the proposed syntax and semantics.

1 Introduction

A *business rule* is a statement that defines or constrains some aspect of a business. Business rules express complex constraints on the business states and how the state of a business or business process should evolve. They talk about necessity and possibility, obligation and permission, and impose constraints on temporal sequences of states of affairs. As such they lie outside the scope of standard ontology languages.

Business processes are increasingly being specified in terms of such rules. In addition to capturing the 'knowledge of a company', the codification of business processes in the form of rules is key to many agile software development strategies. Rule-based approaches offer significant advantages to the application developer: their focus on the declarative representation of small, relatively independent, knowledge units makes it easier for developers and even end users to rapidly develop applications. In addition, the clear separation of application logic (represented as rules) from the bulk of the application code facilitates rapid modification in response to changing organizational goals (e.g., changes in the business process).

Business rules can be expressed either informally in English or formally in terms of predicates and logical connectives. Clearly, the formal language of business rules requires a precise formal semantics, so that the users of the language can understand what the rules mean and the conditions under which rules are violated.

The leading standard for business rules is *Semantics for Business Vocabulary and Business Rules* (SBVR) [1]. SBVR defines a very rich logical language for expressing business rules which includes full first order logic, and both alethic (necessity) and deontic (obligation) modalities. However, SBVR leaves some decisions concerning the precise syntax and semantics of business rules unspecified, while other parts of the

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 41–50, 2010.

specification appear ambiguous or contradictory. For example, SBVR states that the formal language of business rules corresponds to a fragment of full first order modal logic. However while the alphabet of this language is given and some general guidance regarding syntax (e.g., *usually*, each rule contains at most one modality, which is its main connective [1, p.101]), there is no definition of what constitutes a well-formed formula. Similarly, SBVR does not define a precise semantics for business rules. For example, at various points SBVR says that the alethic modality means 'true in all states' (which corresponds to an S5 modality), that it could be S4 [1, p.102] and at other points that the exact modal logic is left open [1, p.101]. SBVR also leaves open the precise semantics of the interaction of first order quantifiers and modalities, while at the same time assuming some logical equalities, such as Barcan formulas, that permute universal quantifiers and alethic modalities. The semantics of deontic modalities is also not completely specified [1, p.103], but some equalities of deontic logic are assumed, e.g., $(p \supset \mathbf{O}q)$ $\mathbf{O}(p \supset q)$ [1, p.103], which says that some sentence p implies that some other sentence q is obligatory, iff it is obligatory that p implies q.

How these questions are resolved has a significant impact on the expressive power of the business rules formalism defined in SBVR and on the meaning of the rules themselves, and hence whether they accurately model the business processes they purport to formalise.

In this paper, we propose a precise syntax and semantics for modalities in business rules and their interaction with first order quantifiers. We show that the implication problem for business rules is undecidable, and characterise complexity of checking whether a state of the business process and its history violates a business rule.

The structure of the paper is as follows: in the next section, the main concepts used to define logical foundations of business rules in SBVR are briefly summarized. In the following sections we discuss syntax and semantics for static and dynamic business rules. In the last section, we propose a restricted syntax of static and dynamic business rules and analyse the compexity of reasoning with business rules.

2 Preliminaries

In this section, the main notions defined in SBVR and used later in the paper are briefly summarized.

A *ground fact* is a "proposition taken to be true by the business". Some facts state that a certain property holds for a specific individual or that some relation holds between individuals . An example would be "Employee 123 works for the sales department". Other facts declare the existence of an individual like "There is a company that has the company name 'Microsoft" [1, p.87-88]. Further, a *fact type* is a kind of ground fact such as "Employee works for Department" [1, p.86]. A ground fact "Employee 123 works for the sales department" is based on or instantiates the fact type "Employee works for Department". Below, we will simplify the account in SBVR and use a single relation symbol for fact types such as $WorksFor(x\ y)$, without the additional variable sort information used in SBVR $WorksFor(Employee\!:\!x\ \ Department\!:\!y)$; this does not make any formal difference but makes formulas more readable.

A *model* or *knowledge base*, as defined in SBVR, is a structure adopted to describe the "universe of discourse" or "business domain" indicating those aspects of the business that are of interest. A model consists of a sequence of *states* over time. A state, in the sense used here, contains ground facts at a particular point of time [1, p.87].

Business rules are constraints which are used to define conditions on business states. *Static* constraints apply to individual states. An example given in [1, p.86] is

> *Each* employee was born on *at most one* date.

Business states which contain facts of the type "Employee born on Date" stating that some individual was born on two or more different dates are ruled out by this constraint.

Dynamic constraints impose restrictions on transitions between business states. An example in [1, p.86] is

> A person's marital status may change from single to married, but not from divorced to single.

The facts "a person's marital status is single", "a person's marital status is married", and "a person's marital status is divorced" are all based on the same fact type and transitions between them are 'controlled' by the dynamic constraint above, which is also constructed from the same fact type.

We consider the syntax and semantics of static business rules in the next section, and dynamic rules in the subsequent section.

3 Static Business Rules

The language for expressing static business rules given in SBVR contains atomic formulas, boolean connectives, first order quantifiers, alethic modalities and $\Diamond$, and deontic modalities **O**, **F** and **P**. A definition of a formula in this language is not given in SBVR. We state here a standard definition of a formula of a full first order language with alethic and deontic modalities:

$$P^n(t_1 \qquad t_n) \;\; t_1 \quad t_2 \;\; \neg \qquad \wedge \qquad \vee$$
$$x \qquad x \qquad \Diamond \quad \mathbf{O} \quad \mathbf{P} \quad \mathbf{F}$$

where P^n is an n-ary predicate symbol and $t_1 \qquad t_n$ are terms (variables or constants), $\neg$, $\wedge$, $\vee$ and are classical boolean connectives (not, and, or, implies), and are existential and universal quantifiers. This part of the definition corresponds to full first order logic. (referred to as necessity, or box modality) and $\Diamond$ (referred to as a possibility, or diamond modality) are called alethic modalities. Intuitively, means that is true in all states and $\Diamond$ means that is true in at least one state. $\Diamond$ is definable as $\neg \quad \neg$. Deontic modalities include **O** (it is obligatory that), **P** (it is permitted that) and **F** (it is forbidden that). **P** is definable as $\neg$**F** and **F** is definable as **O**$\neg$.

In addition to the ordinary existential quantifiers x, SBVR also adopts counting quantifiers ^{n}, $^{0..n}(n \quad 1)$, $^{n..}$, and $^{n..m}(n \quad 1 \; m \quad 2)$ which mean "there are exactly n", "there are at most n", "there are at least n" and "there are at least n and at most m" respectively. These counting quantifiers can be defined in terms of the standard quantifiers in a standard way.

To avoid the paradoxes of standard deontic logic (SDL) and use the same semantics as for the static alethic rules, we propose to define deontic modalities using alethic ones and a sentential constant V for 'undesirable state-of affairs'.[1] Intuitively, $\mathbf{O}$ will be interpreted as saying that all states of the knowledge base satisfying $\neg$ are 'forbidden', or satisfy V:

$$\begin{array}{lll} \mathbf{O} & {}_{def} & (\neg \quad V) \\ \mathbf{F} & {}_{def} & \mathbf{O}\neg \quad (\quad V) \\ \mathbf{P} & {}_{def} & \neg\mathbf{F} \quad \Diamond(\quad \wedge \neg V) \end{array}$$

A similar idea is suggested in SBVR [1, p.104-106] but it involves a predicate 'forbidden' applied to objectified facts.

3.1 Semantics

In giving a semantics for business facts and rules, we need to determine (a) whether constants denote the same object in different worlds or not, and (b) whether different possible worlds have the same domain of quantification or not, and if not, how the domain can change. Different answers to these questions result in different formulas being valid. For example, the converse Barcan formula $x \quad x$ and Barcan formula $x \quad x$ are only valid at the same time if the domain of quantification is the same in all worlds.

We assumpe that constants (such as 'John Smith') should denote the same object in different states of the knowledge base (this is referred to as Rigid Designation in first order modal logic) and that the set of objects which actually exist (and should be in the domain of quantification) changes from one state of the business to the other: new individuals appear and old ones are deleted from the database. In particular, the denotation of 'John Smith' may not 'exist' in some of the states. However, the second assumption seems to clash with the rigid designation assumption. An approach which accommodates both assumptions is to have an 'active domain' of quantification in each state and a general domain which is used for giving meaning to terms and predicates.

Formally, we interpret the language of first order modal logic in a possible worlds structure where each world intuitively corresponds to a state at a given point in time. Each world has its own domain of quantification which intuitively corresponds to the objects actively existing in that state. The union of all of the world domains corresponds to the 'universal domain' of all possible objects. In accordance with the intuition that 'necessary' means 'true in all possible worlds' [1, p.100], we define to be true if is true in all possible worlds. This makes an S5 modality (see, for example, [3]).

Definition 1. *A model for the language of static business rules is a tuple* $W \; D \;\; D_w : w \quad W \;\; v\rangle$ *where*

[1] The sentential constant V was first proposed in [2]. The key idea is that "it is obligatory that p" can be taken to mean "if not-p, then V" where V is some bad state-of-affairs. For example, "It is obligatory that students go to bed on time" can be expressed as " If students don't go to bed on time, they won't feel energetic the next day", which is a bad state-of-affairs.

W is a non-empty set of possible worlds;
D is the non-empty universal domain of the model (the union of all world domains);
D_w is the domain of quantification in each w W. Objects in D_w are said to exist in w and for all w, D_w D.
v is the valuation function which interprets constants and predicate symbols in each w W, namely for a constant c, $v(w\ c)$ D and for an n-ary predicate symbol P, $v(w\ P) \subseteq D^n$.

The models satisfy the following condition (Rigid Designation):

$$w_i\ \ w_j(w_i\quad W \wedge w_j\quad W)\qquad v(w_i\ c)\qquad v(w_j\ c))$$

An inductive truth definition relative to a possible world w and a variable assignment a is given below. w a means 'w satisfies with variable assignment a' or ' is true in the world w with variable assignment a'.

w a $P^n(t_1 \quad t_n)$ iff $(v_a(w\ t_1) \quad v_a(w\ t_n))$ $v(w\ P^n)$ (where $v_a(t_i)$ $v(w\ c)$ if t_i is a constant c and $v_a(t_i)$ $a(x)$ if t_i is a variable x)
w a t_1 t_2 iff $v_a(w\ t_1)$ $v_a(w\ v_2)$;
w a ¬ iff w a ;
w a $\wedge\,\psi$ iff w a and w a ψ;
other propositional connectives are defined as usual;
w a x (x) iff for each d D_w, w $a[d\ x]$ (x), where $a[d\ x]$ is an assignment obtained from a by assigning d to x;
w a x (x) iff there is a d D_w such that w $a[d\ x]$ (x);
w a iff for each w′ W, w′ a ;
w a ◇ iff there is a w′ W such that w′ a .

The definitions of truth in a model and logical entailment are standard. A formula is true in a model, , if w a for all *w* and *a*. A set of sentences *Γ* entails a sentence , *Γ* , if is true in all models where all the sentences of *Γ* are true.

Note that in the models defined above, both the Barcan formula $x\quad (x)\qquad x\ \ (x)$ and its converse $x\ \ (x)\qquad x\quad (x)$ are not valid. A simple counterexample to $x\ \ P(x)\qquad xP(x)$ is a model where W $w1\ w2$, D $d_1\ d_2$, D_{w_1} d_1 , D_{w_2} $d_1\ d_2$, $v(w_1\ P)$ $v(w_2\ P)$ d_1 . Then w_1 $x\ \ P(x)$ but w_1 $xP(x)$. A counterexample to the converse is with the same W and D, but where D_{w_1} $d_1\ d_2$, D_{w_2} d_1 , $v(w_1\ P)$ $d_1\ d_2$ and $v(w_2\ P)$ d_1 . So, if our proposal for the semantics of alethic modalities and quantifiers is accepted, the equivalence $x\quad (x)\qquad (x)$ suggested in SBVR must be given up.

Also, the deontic equivalence suggested in SBVR, $(p\quad \mathbf{O}q)\quad \mathbf{O}(p\quad q)$ is not valid if $\mathbf{O}q$ is interpreted as $(\neg q\quad V)$.

As noted above, the truth definition for given above makes it an S5 modality. This runs contrary to the suggestion in SBVR that if a choice should be made, S4 [1, p.102] is preferred. However, this is an inevitable consequence of defining as 'true in all possible worlds'.

We now give some examples of static business rules from SBVR *Annex E: EU-Rent Example* and their translations into the first order modal logic defined above.

Example 1. It is necessary that each rental has exactly one requested car group:
$x(Rental(x) \quad {}^{1}y\ CarGroupOf(x\ \ y))$

Example 2. It is permitted that a rental is open only if an estimated rental charge is provisionally charged to a credit card of the renter that is responsible for the rental:
$\diamond\ \ x\ \ y(Rental(x) \wedge Open(x) \wedge DriverOf(y\ \ x) \wedge ChargedFor(x\ \ y) \wedge \neg V)$
$x\ \ y(Rental(x) \wedge Open(x) \wedge DriverOf(y\ \ x) \wedge \neg ChargedFor(x\ \ y) \quad V)$

Example 3. It is obligatory that each driver of a rental is qualified: $x\ \ y(Rental(x) \wedge DriverOf(y\ \ x) \wedge \neg Qualified(y) \quad V)$

4 Dynamic Business Rules

According to the SBVR specification, dynamic business rules or constraints restrict possible transitions between business states, and there are two types of dynamic constraints: those which simply compare one state to the next and those which may compare states separated by a given period. Dynamic business rules essentially place constraints on temporal sequences of states. For example, they assert that there is no sequence where a state where *MaritalStatus*(*john single*) is true is immediately followed by a state where *MaritalStatus*(*john divorced*) is true. Similarly, a rule that says that the price charged for the rental should be calculated using the lowest tariff which was applicable during the term of the rental says that for all sequences $w_1 \quad w_n$, if the rental started in w_1 and was charged in w_n, then the tariff charged in w_n should be the lowest value of the tariff in any of $w_1 \quad w_n$. SBVR discusses some possibilities for formal treatment of the semantics of dynamic business rules, including the use of temporal logic, but ultimately defers decisions on their semantics to a later version of the standard [1, p.120].

Our approach to formalizing dynamic rules is based on first-order timed linear time temporal logic (first-order timed LTL) using a similar syntax to [4]. Linear time temporal logic is well suited for stating constraints on discrete linear sequences. Its language contains temporal operators X (X means that is true in the next state) and U ($U(\quad \psi)$ means that in some future state, ψ holds, and until then holds). Similarly to for example [5,6], we introduce timed constraints for the Until operators, of the form $U_{\sim n}$, where and n is a natural number. The meaning of $U\ _n(\quad \psi)$ is that ψ will become true within n time units from now, and all states until then satisfy . Having timed constraints makes it a lot easier to express properties such as ‘within 30 days’ (which can be expressed in standard LTL using X operators but in a rather awkward way).

4.1 Syntax

In addition to the language of static business rules introduced in the previous section, for dynamic business rules we add temporal operators.

As we did for the static business rules, we distinguish alethic and deontic dynamic business rules; the former say what should always be or might be the case after executing a particular change, the latter specify those conditions leading to a ‘bad state-of-affairs‘. As before, we do not use the deontic modalities explicitly, but use the proposition V (forbidden) instead.

The definition of a formula is as for the static business rules, but there is an additional clause for formulas with temporal modalities: if and ψ are formulas, then X and $U_{\sim n}(\ \ \psi)$ are formulas.

4.2 Semantics

Definition 2. *A model for static and dynamic business rules is a tuple $W\ D\ \ D_w$: $w\ \ W\ \ v\ R\rangle$ where $W\ D\ \ D_w : w\ \ W\ \ v\rangle$ is as in Definition 1 and $R \subseteq W\ \ W$ is a temporal transition relation (where is the duration of the transition). We assume that W is linearly ordered by $dR(w\ \ w'\ \ d)$ (that is, each world has a unique successor and corresponds to a linear flow of time), that there is an initial state (moment of time) $w_0\ \ W$, and the flow of time is unbounded to the right (for every w, there exists a w' and a d such that $R(w\ \ w'\ \ d)$).*

To make some expressions more intuitive, we will sometimes denote a transition $(w_1\ \ w_2\ \ d)\ \ R$ by $w_1 \overset{d}{\ } w_2$.

$\pi\ \ w_1 \overset{d_1}{\ } w_2 \overset{d_2}{\ } w_3$ is called a path in starting in w_1. Each model contains a single maximal path π_0 starting in w_0, and all other paths are suffixes of π_0. We define a function $t : W$ which assigns timestamps to the possible worlds in : $t(w_0)\ \ 0$, and $t(w_n)\ \ d_0\ \ \ \ d_{n\ 1}$, where $w_0 \overset{d_0}{\ } w_1\ \ \overset{d_{n\ 1}}{\ } w_n$.

All clauses of the truth definition are the same as in Definition 1, with the following additions:

- *$w\ \ a\ \ \ X$ iff there exists w' and d such that $w \overset{d}{\ } w'$ and $w'\ \ a$*
- *$w\ \ a\ \ \ U_{\sim n}(\ \ \psi)$ iff there exists w_k with $k\ \ 2$ such that for some w_2 and d_1 , $w \overset{d_1}{\ } w_2 \overset{d_2}{\ } \overset{d_{k\ 1}}{\ } w_k$ and $w_k\ \ a\ \ \psi$, $t(w_k)\ \ t(w)\ \ n$ and for all $w_i\ \ w\ \ w_1\ \ \ w_{k\ 1}$, $w_i\ \ a$.*

The definitions of truth in a model and logical entailment are standard.

As usual, other operators can be defined in terms of $U_{\sim n}(\ \ \psi)$. Namely, the plain (without a subscript) $U(\ \ \psi)\ \ _{df}\ U\ _0(\ \ \psi)$, $F_{\sim n}\ \ \ U_{\sim n}(\top\ \ )$ (holds at some time in the future satisfying n), and $G_{\sim n}\ \ \ \neg F_{\sim n}\neg$ (suppose w_i is the current time point, then holds in all future points w_j, where $t(w_j)\ \ t(w_i)\ \ n$).

Below are some examples of dynamic business rules:

Example 4. It is necessary that the marital status of a person does not change from married to single: $x(MaritalStatus(x\ \ married)\ \ \ \neg X\ MaritalStatus(x\ \ single))$

Example 5. It is obligatory that the tariff charged for the rental is not greater than the lowest tariff applicable during the term of the rental: $x\ \ y_1(Rental(x) \wedge FCharged(x\ \ y_1) \wedge\ \neg U(\neg\ \ y_2(y_2\ \ \ y_1 \wedge Tari\ \ (x\ \ y_2))\ \ Charged(x\ \ y_1))\ \ \ V)$

Example 6. It is possible that the notification of a problem during a rental occurs after the return of the rental: $\diamond\ \ x\ \ y(Rental(x) \wedge Returned(x) \wedge F\ Problem(x\ \ y))$

Example 7. It is obligatory that each invoice which was issued on Date_1 is paid on Date_2 where $\mathrm{Date}_2\ \ \ \mathrm{Date}_1$ + 30 Days: $x(Invoice(x) \wedge \neg F\ _{30}Paid(x)\ \ \ V)$

5 Complexity of Reasoning

Turning to the complexity of reasoning with business rules, we observe that it does not seem to make sense to state the results for full first-order modal temporal logic. Clearly, business rules correspond to a proper fragment of it. For example, all business rules we have seen so far contain a single occurrence of the modality. From empirical analysis of the examples given in SBVR and other collections of business rules, we conclude that it is reasonable to constrain the logical syntax of static business rules to the following general form:

Static necessity constraints $\bar{x}(\quad \psi)$
Static possibility constraints $\Diamond \; \bar{x}(\; \wedge \psi)$

where $\bar{x}$ are all the free variables of ψ, is a conjunction of atomic formulas, and ψ is an existentially quantified formula in disjunctive normal form (where existential quantifiers may be counting quantifiers).

Note that some of the constraints above, namely the second sentence in Example 2 and Example 3 were stated in a slightly different form, for readability. However, they can be easily rewritten to conform to the pattern above:

Example 2' $x \; y(Rental(x) \wedge Open(x) \wedge DriverOf(y \; x) \quad ChargedFor(x \; y) \vee V)$

Example 3' $x \; y(Rental(x) \wedge Driver(y \; x) \quad Qualified(y) \vee V)$

For the dynamic business rules, we suggest the form

Dynamic necessity constraints $\bar{x}(\quad \psi)$
Dynamic possibility constraints $\Diamond \; \bar{x}(\; \wedge \psi)$

where and ψ are as before, but in addition are allowed to contain -free formulas starting with a temporal modality. The latter may be a somewhat overgenerous definition, but on the other hand we do not want to preclude a possibility that the antecedent or the consequent of a rule may contain nested temporal modalities.

Before analyzing the complexity of reasoning with business rules, we translate them into the language of first order logic, for the convenience of proving the corresponding results. In the translation, each predicate acquires an extra argument place corresponding to a state where the formula is evaluated. The translation ST^w is analogous to the Standard Translation from modal logic to classical logic defined by van Benthem [7] (w is the parameter corresponding to the free state variable). Since we use the universal modality to interpret , we do not need to introduce an explicit accessibility relation in the translation. However, we need to restrict quantification to the domains of the possible worlds. We do this by introducing a binary predicate $D(x \; w)$ which corresponds to 'x is in the quantification domain of w'. We also need predicates R and R corresponding to temporal relations (where R is a strict transitive closure of $d(R(x \; y \; d))$ and a functional symbol t corresponding to timestamps.

Formally, the translation ST^w is as follows:

$ST^w(P(x_1 \qquad x_n)) \quad P(x_1 \qquad x_n \; w)$
ST^w commutes with the booleans

$ST^w(\ x\ )\quad x(D(x\ w)\quad ST^w(\ ))$
$ST^w(\ x\ )\quad x(D(x\ w) \wedge ST^w(\ ))$
$ST^w(\ \ )\quad w\,ST^w(\ )$
$ST^w(X\ )\quad w'\ \ d(R(w\ w'\ d) \wedge ST^{w'}(\ ))$
$ST^w(U_{\sim n}(\ \ \psi))\quad w'(R\ (w\ w') \wedge ST^{w'}(\psi) \wedge t(w')\quad t(w)\quad n \wedge\ \ w''(w''\quad w' \wedge$
$\quad R\ (w\ w'') \wedge R\ (w''\ w')\quad ST^{w''}(\ )))$

Proposition 1. *ST^w is a truth preserving translation, that is $\quad w\ a\quad$ iff $M'\ a$ $ST^w(\ )$ where $\quad W\ D\ \ D_w : w\quad W\ \ v\ R\rangle$ and $\ '\quad W\quad D\ D^2\ R\ R\ \ v'\ t^1\ w_0\rangle$ is a structure corresponding to $\quad$, where the domain is $W\quad D$, a binary relation D^2 corresponds to $D_w\ :\ w\quad W$, R is a strict transitive closure of $\ dR(w_1\ w_2\ d)$, v' is a set of relations given by v but with an extra worlds parameter, t^1 is a function $W\quad D\quad$ where for $w\quad W$, $t(w)$ is the distance from w_0 and for $e\quad D\ t(e)\quad 0$, and w_0 is the initial world.*

Proof. It is immediate from the Definition 2 that the translation is just a reformulation of the truth conditions for the formulas of modal temporal logic. It is easy to prove by induction on the complexity of the formula that the translation is truth preserving.

The first problem we would like to consider is the implication problem for business rules. By the implication problem for business rules we mean the question whether a finite set of business rules Γ entails a business rule . Clearly, it is very important to be able to check whether a set of rules is consistent (entails a trivial business rule). It may also be useful to check whether a set of rules is redundant (some rule is actually entailed by other rules). Both questions can be answered if we have an algorithm to solve the implication problem. Unfortunately, even if we restricted the syntax of business rules as proposed above, the problem is undecidable.

Proposition 2. *The implication problem for business rules is undecidable.*

Proof. The translation of business rules into first-order logic includes all formulas in the prenex class which is undecidable [8].

We next turn to another natural question: are any constraints violated by a run of the system? Note that this problem only has practical meaning for necessity constraints, and even then only for a subset of them (basically the ones which translate into a purely universal first order sentence; we will refer to them as purely universal business rules). This is not a classical model-checking problem: whether a model satisfies a sentence (business rule) . Rather, we are talking about *run-time* model-checking: whether a finite run of a process violates an invariant property. We view the current state of a business and its history as an initial prefix of a real (infinite) model. To see why it does not make sense to check for violations of possibility constraints, consider the following example of a simple possibility constraint $\diamond\ \ x(Employee(x) \wedge PartTime(x))$. It says that a state where someone is a part-time employee is possible — in principle; it does not say that such a state should have actually existed in the past history of the business, so if it does not hold in the current state or any of the preceding states, we still should not consider it violated. Similarly, a necessity constraint which says that each invoice should be paid within 30 days is not violated — yet — if there exists an unpaid invoice, but it has just been issued. However, this constraint can be reformulated as 'All invoices issued 30 or more days ago should have been paid', which is a universal statement and can be meaningfully checked for the finite history and the present state of the business.

Formally, we define as a 'state and history of the business' a finite prefix of a model . It corresponds to a finite linearly ordered sequence of states with the initial state w_0 (when the records began) and the current state w_n. We define the problem of checking whether a state and history of the business satisfies a universal business rule as the problem of checking whether w_0 .

Proposition 3. *The problem of checking whether a state and history of the business satisfies a purely universal business rule is solvable in time polynomial in the size of .*

Proof. Using the translation ST^{w_0} and Proposition 1, we can reduce this problem to the question whether a finite first order structure $'$ satisfies a first-order sentence $ST^{w_0}(\ )$. It is well known, see for example [9], that this problem can be solved in time polynomial in the size of $'$. Note that the size of is the same as the size of $'$ apart from the explicit representation of R and t. However R and t can be computed in polynomial time and only result in a polynomial increase in the size of $'$.

6 Conclusions

In this paper, we made a proposal for the precise formal syntax of business rules and their logical semantics. We have also shown that the implication problem for business rules is undecidable, and a natural problem of checking for violation of universal constraints can be solved in time polynomial in the size of the business state and history. In future work we plan to investigate the design and implementation of efficient algorithms for solving the latter problem.

References

1. OMG: Semantics of business vocabulary and business rules (SBVR) (2008)
2. Anderson, A.R.: Some nasty problems in the formal logic of ethics. Noûs 1, 345–360 (1967)
3. Bull, R., Segerberg, K.: Basic modal logic. In: Gabbay, D., Guenthner, F. (eds.) Handbook of Philosophical Logic. Extensions of Classical Logic, vol. II, pp. 1–81. Kluwer, Dordrecht (2001)
4. Kare, J., Kristoffersen, C.P., Andersen, H.R.: Checking temporal business rules. In: Proceedings of the First International REA Workshop (2004)
5. Alur, R., Courcoubetis, C., Dill, D.L.: Model-checking in dense real-time. Inf. Comput. 104(1), 2–34 (1993)
6. Laroussinie, F., Markey, N., Schnoebelen, P.: Efficient timed model checking for discrete-time systems. Theor. Comput. Sci. 353(1-3), 249–271 (2006)
7. van Benthem, J.: Modal logic and classical logic. Bibliopolis (1983)
8. Börger, E., Grädel, E., Gurevich, Y.: The Classical Decision Problem. Springer, Heidelberg (1997)
9. Ebbinghaus, H.-D., Flum, J.: Finite model theory. Springer, Heidelberg (1995)

Dialect Recognition Method Using Emotion Judgment

Noriyuki Okumura

Department of Electronics and Information Science, Nagano National College of Technology, Tokuma 716, Nagano, Nagano, Japan
noriyuki_okumura@ei.nagano-nct.ac.jp

Abstract. This paper proposes a method to recognize dialect on personal computers. Proposal method pays attention to dialects' emotion content of speakers. Dialects have much emotional expression compared with standard dialect. Therefore, this system judges sentences including some dialects using Emotion Judgment System inputting dialects as standard dialects. Emotion Judgment System answers wrong emotion as standard dialect if inputs include some dialect. This paper treats dialect used in Nagano prefecture, Japan. This paper shows that using speaker's emotion can solve the disambiguation of the dialect and the standard dialect.

Keywords: Dialect, Emotion Judgment, Commonsense, Concept-base, Degree of Association, Morphological Analysis.

1 Introduction

This paper proposes a method to recognize dialect on personal computers. We must solve words sense disambiguation if dialect and standard dialect are homonyms. proposal method pays attention to Dialect's Emotion content. Dialects have emotional expression compared with standard dialects. Therefore, this system judges sentences including some dialects using Emotional Judgment System inputting dialects as standard dialects.

Generally, there are some methods such as using Corpora [1], heuristics [2], and neural networks [3] and so on in solving words sense disambiguation. However, these methods are based on statistical processing, they do not pay attention to speakers' emotion.

In this paper, we propose a method that is able to aid morphological analysis using speakers' emotion. This paper proposes a method that converts from the dialect into the standard dialect based on speakers' emotion.

2 Purpose

The purpose of this study is to specify the meaning by using speaker's emotion when the dialect and the standard dialect are homonyms. This paper treats especially dialect used in Nagano prefecture, Japan. For example, 'KO-WA-I' means 'afraid' in standard dialect, but in Nagano prefecture this word means 'hard', 'firm', and so on. Fig.1 shows that we use both meanings of 'KO-WA-I'.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 51–59, 2010.

私は彼がこわい (Watashi ha Kare ga KO-WA-I)
(In English : I'm afraid of him.)

この野菜はこわい (Kono Yasai ha KO-WA-I)
(In English : This vegetable is firm.)

Fig. 1. The example of dialect and standard dialect (KO-WA-I)

Thus, this paper constructs the system that enables the interpretation of the dialect by correctly specifying the meaning of words and phrases that the meaning is different because of the dialect and the standard dialect, and emotion that the speaker holds are different.

3 Associated Techniques

In this section, we explain associated techniques used in this paper. I explain of each technique of the concept base in 3.1, of the related level calculation method in 3.2, of the feelings judgment system in 3.3, of the morphological analysis machine in 3.4.

3.1 Concept-Base

Concept-base [4] is a large-scale relational database constructed from Japanese electronic dictionaries and newspapers. Headwords in dictionaries are assumed to be concepts and content words in explanation sentences are assumed to be attributes for headwords (concepts). Concept-base is a very important factor to Emotion Judgment. Concept-base consists entirely of sets of concepts. A concept (A) consists of sets of pairs of attributes (a_i) which characterizing concept (A) and weights (w_i) which mean the importance of each attributes ('i ' is a natural number for each concepts)(eq.1).

$$A = \{(a_i, w_i) \mid 0 < i < znum + 1, \textit{znum} \text{ is a number of attributes}\}\ . \tag{1}$$

Attributes for each concept are also defined in Concept-base as concepts. Therefore, one concept is defined as attributes chain model of n-th-order dimension. In this

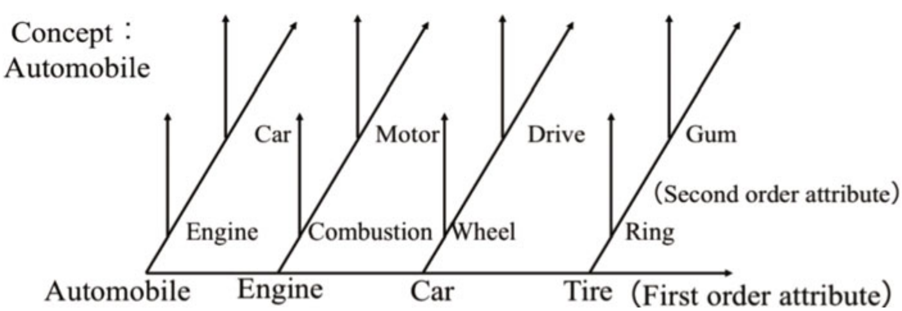


Fig. 2. Concept (automobile) is extended to second order attribute

paper, Concept-base has about 90,000 concepts, and each concept has 30 attributes on average. Fig.2 shows the example of concept (automobile). ‘Automobile’ has attributes (engine, car, tire, and so on). ‘Engine’, ‘Car’, ‘Tire’ are also defined in Concept-base. Therefore, ‘Engine’ has attributes (Combustion, Motor, and so on).

3.2 A Calculation Method of Degree of Association

For measure method between concepts, Calculation Method of Degree of Association [4] is proposed. In Calculation Method of Degree of Association, concepts are defined as attributes sets with weight, and Degree of Association is calculated in consideration of Degree of Match. For the calculation of Degree of Match, pairs of attributes set which are Match or Similar are defined between the concepts. After the process, for the pairs of attributes, it is calculated in consideration of weights of the attributes.

3.2.1 Degree of Match

For two concepts: *A* and *B*, the first-order attributes and the weights are defined as follows (eq.2). u_i and v_j in eq.2 and eq. 3 are weights of attributes (a_i, b_j).

$$A = \{(a_i, u_i) \mid i = 1 \quad L\}, B = \{(b_j, v_j) \mid j = 1 \quad M\} \tag{2}$$

Then, it is defined of Degree of Match ($MatchWR(A,B)$) between concepts: *A* and *B*. In addition, each attributes’ weights are normalized to the total of sum of weights equal 1.0.

$$MatchWR(A,B) = \sum_{a_i = b_j} Min(u_i, v_j), \quad Min(u_i, v_j) = \begin{cases} u_i (u_i < v_j) \\ v_j (else) \end{cases} \tag{3}$$

Eq.3 defines degree of Match, because if common attributes are detected, then effective weight is smaller weight of both common attributes. In addition, Degree of Match is from 0.0 to 1.0.

3.2.2 Calculation Method of Degree of Association

Refined eq.2, *A* is assumed to be the concept that has lower number of attributes ($L \leq M$). The rank of first-order attributes of concept *A* draw in order.

$$A = \{(a_i, u_i) \mid i = 1 \quad L\} \tag{4}$$

First, it searches attributes which become $\{a_i = b_j \ (MatchWR(a_i, b_j) = 1.0)\}$. The case is searched which becomes $a_i = b_j$ (Spell Match), if u_i is greater than v_j, then $u'_i = u_i - v_j$, else $v'_j = v_j - u_i$, then using u'_i or v'_j , it is made to correspond to other attribute. The relation is more correctly appreciable.

Second, the rank of first-order attributes of concept *B* is decided with concept *B* as reconstructed with concept *A* by Eq.5 considered with Completely Match. The number of Completely Match is assumed to beα. In addition, $b_{L+\alpha+1}$,$b_{L+\alpha+2}$, b_M is not used.

$$Bx = \{(b_{x1}, v_{x1}), (b_{x2}, v_{x2}), \cdots, (b_{xL+\alpha}, v_{xL+\alpha})\} \tag{5}$$

x_k = {Number of attributes to which a_i is decided correspondence}

Third, *ChainWR*(*A*, *B*) is defined as follows, because *L*+*α* attributes correspond.

$$ChainWR(A,B) = \sum_{i=1}^{L+\alpha} MatchWR(a_i,b_{x_i}) \times \frac{u_i + v_{x_i}}{2} \times \frac{\min(u_i,v_{x_i})}{\max(u_i,v_{x_i})}$$
$$\min(u_i,v_{x_i}) = \begin{cases} u_i (u_i \le v_{x_i}) \\ v_{x_i} (otherwise) \end{cases} \quad \max(u_i,v_{x_i}) = \begin{cases} u_i (u_i \ge v_{x_i}) \\ v_{x_i} (otherwise) \end{cases} \tag{6}$$

ChainWR(*A*, *B*) corrects Degree of Association by multiplying average $((u_i + v_{xi})/2)$ of both attribute's weight and ratio $(\min(u_i,v_{xi})/\max(u_i,v_{xi}))$ of both attributes' weights to *MatchWR*(*A*, *B*). Degree of Association is from 0.0 to 1.0. Table 1 shows the example of Degree of Association.

Table 1. The example of Degree of Association (Automobile: Airplane, Train, Car)

Basic Concept	Object	Degree of Association
Automobile	Airplane	0.22
	Train	0.35
	Car	0.7

3.3 Emotion Judgment System

Emotion Judgment System [5] is a system that guesses speakers' emotion. This system receives a natural utterance as an input and this system outputs 10 kinds of speakers'

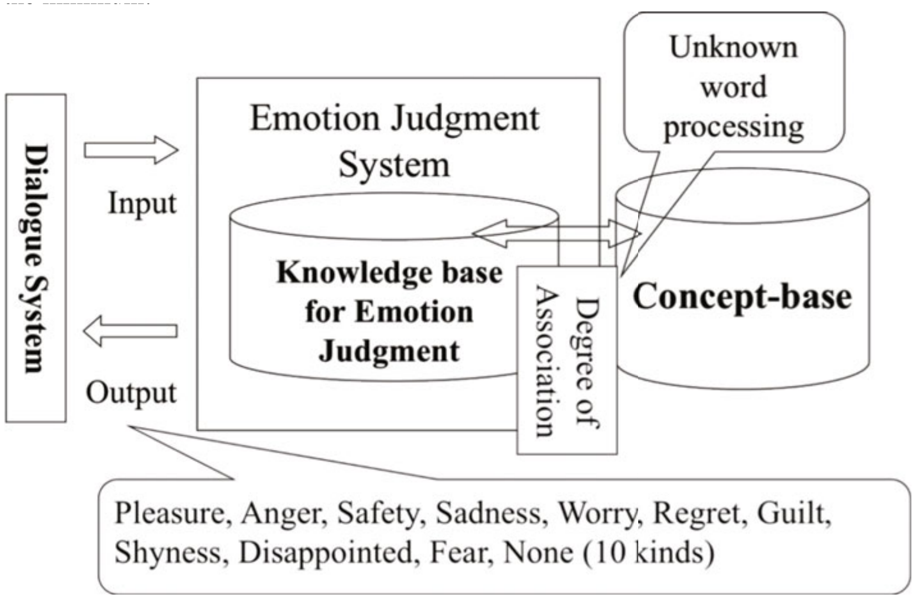


Fig. 3. Outline of Emotion Judgment System

emotion. For example, if the input sentence is "I got the beautiful jewelry." the system outputs 'Pleasure'. Fig.3 shows the outline of Emotion Judgment System. Concept-base in Fig.3 converts the word that doesn't exist in the knowledge base for Emotion Judgment into the word that exists in the knowledge base (Unknown word processing). Concept-base enables the size of the knowledge base to be suppressed to the minimum.

3.4 Morphological Analysis

Proposal method uses MeCab [6] for morphological analysis of judging that input is dialect or standard dialect. MeCab is a morphological analyzer that is known to be high-speed and highly accurate. In addition, users freely define how to treat unknown words.

In this paper, we add the dialect to MeCab's dictionary for analyzing sentences including the dialect.

4 Proposed Method

The word with another meaning in the same mark exists in a Japanese standard dialect and the dialect of Nagano Prefecture as shown in Fig.1. In this paper, if emotion are guessed by using Emotion Judgment System based on sentences just before the input sentence, and sentences immediately before are corresponding to emotion of the input sentence, it proposes the standard dialect and the system that judges that otherwise, there is a possibility of the dialect. Feelings of speaker *A* and speaker *B* are corresponding as shown in Fig.4 using Emotion Judgment System by assumption that inputs are standard dialect for example "I was gotten food poisoning eating this vegetable", "You ate a scary vegetable".

Next, when the dialect is input to the system as a standard dialect, the system judges the meaning of the dialect as a standard dialect. Therefore, the difference of Emotion is caused (Fig.5).

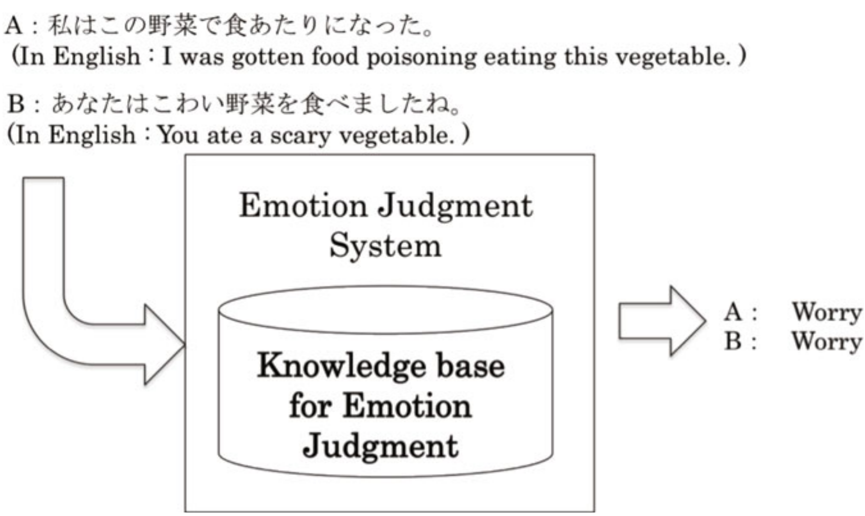


Fig. 4. Using Emotion Judgment System by assumption that inputs are standard dialect

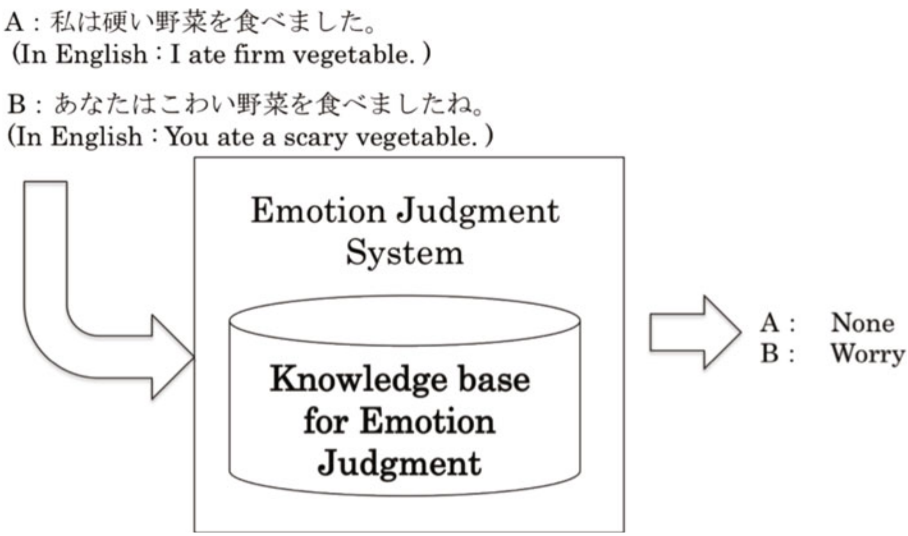


Fig. 5. Using Emotion Judgment System by assumption that inputs are standard dialect

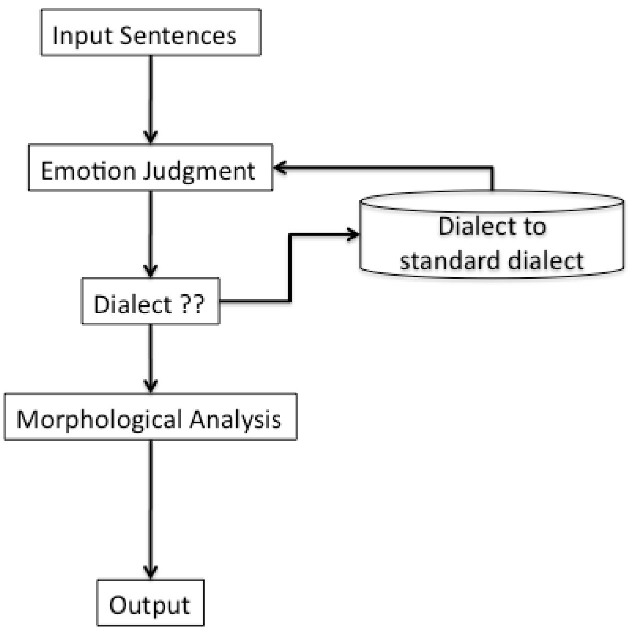


Fig. 6. Outline of proposal method

Thus, this system pays attention before and behind the input sentence, and judges whether the in put sentence is a dialect or a standard dialect from the flow of Emotion. Fig.6 shows outline of proposal method.

5 Results

To examine the tendency to the dialect used in Nagano Prefecture, we caught the questionnaire. The questionnaire was executed to about 200 people, and 130 effective answers were obtained. The dialect with a peculiar spelling that did not exist in the

Table 2. Dialect in Nagano Prefecture which have same spelling for standard dialect (32 words)

Dialect in Nagano Prefecture			
おい (OI) S : Hi D : You	おっしゃん (Osshan) S : Say D : Priest	おる (Oru) S : Fold D : be here	おれた (Oreta) S : Puckered D : You
かう (Kau) S : Buy D : Button	からい (Karai) S : Hot Flavor D : Salty	かる (Karu) S : Mow D : Button	かんじる (Kanjiru) S : Feel D : Cold
こうもり (Koumori) S : Bat D : Umbrella	こく (Koku) S : Rich D : Say	こげん (Kogen) S : Burn D : This much	たたる (Tataru) S : Curse D : Build
ねった (Netta) S : Formulate D : Sleep	やらず (Yarazu) S : Don't do D : Let's do	よせる (Yoseru) S : Get together D : Clean up	さら (Sara) S : Dish D : Case
だべ (Dabe) S : Say D : Be	おしかけ (Oshikake) S : Barge D : Wash rice	くれる (Kureru) S : Recieve D : Give	さわぎ (Sawagi) S : Pother D : Accident
しみる (Shimiru) S : Nip D : Cold	つく (Tsuku) S : Cue D : Push	つる (Tsuru) S : Fishing D : Carry	ほける (Hokeru) S : Kicked D : Grow up
まる (Maru) S : Circle D : Do the business	いってきました (Itte-Kimashita) S : Be gone D : Re-hi	へら (Hera) S : Paddle D : Tongue	えらい (Erai) S : Smart D : Tired
とぶ (Tobu) S : Jump D : Hurry up	ぼける (Bokeru) S : Tease D : Spongy	ずく (Zuku) S : Force D : Hustle	こわい (Kowai) S : Afraid D : Firm

standard dialect of 130 words and the dialect with the same spelling as the standard dialect of 32 words were acquired by the investigation. The dialect with a peculiar spelling can be specified that the dialect at the stage of the morphological analysis by MeCab. Therefore, this paper is verified for 32 words with the same spelling as the standard dialect. Table.2 shows 32 words.

National Institute of Japanese Language further has described many spellings in the dialect dictionary [7] though the dialect with the same spelling as the standard dialect was acquired by 32 words. However, because the operation frequency of the dialect is not clear in the material of the dictionary, 32 extracted words are verified by the questionnaire survey.

Next, the morphological analysis result as the standard dialect when information on the dialect was added to the dictionary of MeCab was investigated. If the following examples are made for the investigation, and information that shows that it is a dialect in most significant of the morphological analysis result is acquired, it is assumed that it answers correctly. Fig.3 shows the result.

Table 3. The result of Writing with a space between words and Morphological Analysis

	Normal	Dialect Added
Writing with a space between words	100%	100%
Morphological Analysis	100%	16%

It is understood that the accuracy of the morphological analysis has decreased by adding information on the dialect to the dictionary from Table.3. Then, the experiment to identify the dialect to the standard dialect by using Emotion Judgment System was investigated. It examined whether 12 testees were made to make the example sentences, and the evaluation by the system was corresponding to man's evaluation. The number of example sentences is 208. Table.4 shows the result.

Table 4. The result of using Emotion Judgment System

Coresponding to Emotion Judgment System	0
Not Coresponding	208
The one output by system in meaning of standard dialect	1
The number of outputs (None)	207

Emotions are not output in many cases.

6 Discussion

The accuracy of the morphological analysis as the dialect has improved by adding information on the dialect to the dictionary of MeCab. However, the problem that the discrimination capacity as the standard dialect decreased greatly, and an original performance could not be demonstrated became clear.

Then, when the example was verified by using Emotion Judgment System, it was understood not to agree to Emotions of an actual speaker when sentences including the dialect were interpreted as a standard dialect. As a result, it is thought that the polysemy of the dialect and the standard dialect can be flexibly canceled by recognizing it as a standard dialect by recognizing it as a dialect when speaker's feelings are not corresponding when agreeing.

7 Conclusions

In this paper, the technique to recognize the dialect appropriately by paying attention to the dialect of Nagano Prefecture, and using Emotion Judgment System was described. The accuracy of writing with a space between words clarified the problem that the accuracy of the morphological analysis decreases by enhancing the dictionary of the morphological analysis machine while the improvement. Moreover, processing that appropriately classified the dialect and the standard dialect was described in using Emotion Judgment System. I want to compare with not only Emotions but also man's common senses and to develop the system that can recognize various dialects at the same time in the future.

References

1. Yarowsky, D.: Word-sense disambiguation using statistical models of Roget's categories trained on large corpora. In: Proc. COLING 1992, pp. 454–460 (1992)
2. Yarowsky, D.: Unsupervised word sense disambiguation rivaling supervised methods. In: Proc. ACL 1995, pp. 189–196 (1995)
3. Veronis, J., Ide, N.: Word sense disambiguation with very large neural networks extracted from machine readable dictionaries. In: Proc. COLING 1992, pp. 359–365 (1992)
4. Noriyuki, O., et al.: An Association Method Using Concept-base. In: Apolloni, B., Howlett, R.J., Jain, L. (eds.) KES 2007, Part I. LNCS (LNAI), vol. 4692, pp. 604–611. Springer, Heidelberg (2007)
5. Seiji, T., et al.: Proposal of Method to Judge Speaker's Emotion Based on Association Mechanism. In: Apolloni, B., Howlett, R.J., Jain, L. (eds.) KES 2007, Part I. LNCS (LNAI), vol. 4692, pp. 847–854. Springer, Heidelberg (2007)
6. Taku, K., et al.: Applying Conditional Random Fields to Japanese Morphological Analysis. In: Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP 2004), pp. 230–237 (2004)
7. The National Language Research Institute, Tokyo, Japan: Texts of Tape-Recorded Conversation In Japanese Dialects, vol. 1 (1978)

Applying a Knowledge Based System for Metadata Integration for Data Warehouses

Dan Wu and Anne Håkansson

Department of Communication System, School of ICT,
KTH Royal Institute of Technology,
Electrum 418, 164 40 Kista
{dwu,annehak}@kth.se
http://www.kth.se/ict?l=en_UK

Abstract. Data warehouses is a typical example of distributed systems where diverse tools and platforms need to communicate to understand each other. For the communication, metadata integration is significant. Seamless metadata interchange improves the data quality and the system effectiveness. Metadata standards exist, for instance, Common Warehouse MetaModel (CWM), which have enhanced the metadata integration. However, it is far from solving the problem of metadata integration in data warehouse environment. This paper proposes an approach to apply a knowledge-based system that supports the metadata integration. By utilizing the knowledge of software engineers on Common Warehouse MetaModels and the metadata interchange models, the knowledge-based system can give metadata interchange model suggestions. Such a knowledge-based system intends to partly automate the metadata integration to improve the efficiency and the quality of metadata integration in data warehouses.

Keywords: Metadata integration, Knowledge based system, CWM, Data Warehouse.

1 Introduction

A data warehouse (DW) is "a subject-oriented, integrated, nonvolatile, time-variant collection of data in support of management's decisions[11]". In a DW environment, data is extracted from various heterogenous systems and integrated into the data warehouse for analysis purpose[8][11]. The Extraction, Transformation and Loading (ETL) tools automate the processes of data extraction, transformation and loading. DW is the target of the integrated data and applies multidimensional models, aggregation and selection applications on the data. The analysis tools such as OLAP and data mining apply further analysis on the data to generate results for the decision makers on different levels[8]. Since many tools, such as ETL tools, data modeling tools, DW repository tools and analysis tools, are involved in a DW environment, metadata integration is of critical importance to minimize the administration efforts and improve the extraction of information of DW[8].

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 60–69, 2010.

In the early 2000s, Object Management Group (OMG) adopted the Common Warehouse Metamodel (CWM) "to enable easy interchange of warehouse and business intelligence metadata between warehouse tools, warehouse platforms and warehouse metadata repositories in distributed heterogeneous environments"[9]. CWM specification defines the metadata as "data about data. Examples of metadata include data element descriptions, data type descriptions, attribute/property descriptions, range/domain descriptions, and process/method descriptions[9]". CWM has defined an abstract language for expressing metadata models (called metamodels) for the various DW tools[9], and is supposed to be the solution for the metadata integration in the data warehouse systems. However, the DW repository vendors are slow in implementing CWM[10]. In many cases, systems have to be designed with diverse metadata for various reasons, such as "best-of-breed", "most compatible" or "most economic" [1]. In fact, it is hard to find robust tools that completely support CWM[1]. CWM does not serve its purpose as it is designed in practice and the metadata integration is often an ad hoc solution.

We assume that CWM is a good specification of metamodel for DW. The expert of CWM has a good understanding of general issues of metadata integration in DW environment. We therefore present an approach of a knowledge-based system (KBS) for supporting the metadata integration. The knowledge base in the KBS represents the knowledge of the expert engineers on the specification of CWM. The KBS also stores the instances of CWM metadata models and the translation of metadata models. The purpose of the system is partly automate the metadata integration process to improve the quality and the efficiency of the metadata integration by reusing the knowledge of metadata integration and accumulating the instances of the metadata models. The following sections are arranged as: Section 2 is about the related work around CWM and the metadata management in data warehouses; Section 3 is about the CWM specification and its implementing; Section 4 is the description of the knowledge-based system (KBS) model; Section 5 is the conclusions and the future work.

2 Related Work

The research on CWM is foremost concerned DW domain and for the DW system architecture modeling. Mazón et al.[6] describe a complete model driven approach for the development of all the components of DWs. The authors apply the UML and CWM to model a DW repository. Medina and Trujillo[3] show how to transform the multidimensional properties with the CWM package, Online Analytical Processing (OLAP), to a CWM-based metamodel, therefore to realize the metadata integration in a CWM-based DW system. Auth and Maur[4] describe a software architecture based on CWM with which to enable the metadata integration in the systems. CWM has inspired other areas of modeling, e.g., the enterprise modeling and the information mediation systems. Neaga and Harding[2] contribute with a common knowledge enterprise model using Model-Driven Architecture (MDA) and Common Warehouse MetaModel (CWM). Zhao and Siau[7] propose a new architecture established on the CWM resolving the metadata problems for information mediation on the online information sources. While this paper suggests a new application with the CWM standard. The CWM specification is not only a model driven approach to DW systems, but also a source of knowledge on metadata integration for a KBS, which improves the metadata integration.

Another related area is the metadata management for data warehouses. Sen[1] who present a survey of the metadata management for the latest 50 years, proposes the concept of "metadata warehouse" to manage the metadata storing and the metadata changes. Zhao and Huang[12] describe how to use Description Logics to reason on CWM to improve the metadata consistencies and reduce metadata redundancies. Sen applied database technology in the implementation and Zhao and Huang tested their idea with the Description Logics. The research in this paper suggests the different approach of KBS and the focus of the KBS is the metadata integration.

3 Common Warehouse MetaModel Specification

CWM specification defines classes and associations abstracted from the common features and functions used in the DW domain. These classes and associations are organized in 21 packages and the packages are arranged in a 5-level-hierarchy. Each package represents a specific area in DW, e.g., relational resource. The architecture supports the model reuse and the understanding of the specification[9].

CWM conforms to Meta Object Facility (MOF), which is a standard technology for defining, constructing, managing, interchanging and integrating metadata[9]. MOF uses four levels architecture, named M3, M2, M1 and M0, where M0 is the lowest level with Object and M3 is the highest level with metadata object, see figure 1. CWM is an instance of MOF and a meta-metadata, so it is on the level M2. The models on M3 level describe the models on M2 level, the models on M2 level describe those on M1 level and the models on M1 level describe models on M0 level. The higher up the models are in the hierarchy, the more abstract the models become. Thus, the specific CWM metadata model instances are on the M1 level, describing specific models conformed to CWM. Our idea is to generalize CWM to integrate the models on M2 and M1 levels.

Meta-level	MOF terms	Examples
M3	Meta-MetaModel	The "MOF" model
M2	MetaModel, Metametadata	UML metamodel, CWM metamodel
M1	Model, Metadata	UML models, CWM metadata
M0	Object, data	Modeled systems, Warehouse data

Fig. 1. OMG Metadata Architecture[9]

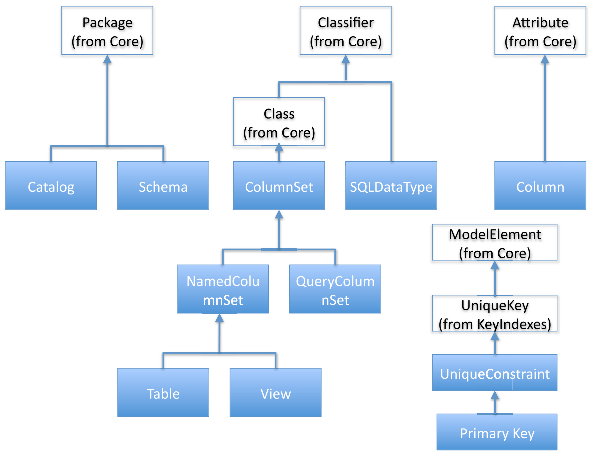


Fig. 2. Relational Package inheritances (partly) [9]

CWM uses the UML notation and the Object Constraint Language (OCL) to define the generic warehouse metamodels. By extending the Unified Modeling Language (UML), all the CWM meta-classes are inherited from the UML meta-classes. Figure 2 shows some of the relational package inheritances, e.g., the Schema inherits the Package and the Table inherits the Classifier.

CWM is designed for, e.g., DW tool vendors and DW end-users. figure 3 shows a CWM implementing instance that is developed by CWM users. Comparing with Figure 2, it shows how an instance follows the CWM specification. Figure 2 does not show the relationship between Schema and Table. However, in the Package of Core in CWM specification, it defines that a Package contains ModelElements such Packages and Classifiers. Therefore, the Schema Company contains several tables. Column inherits Attribute, which inherits Feature from the Core Package; Classifier is an abstract class and Feature is its contained elements. So the table Employees contains several columns. EmpKey is an instance of UniqueConstraint, an alternative is to define it as an instance of PrimaryKey. The CWM specification[9], the book[5] and the articles [3][4] are recommended for more detailed knowledge of CWM specification and its implementation.

4 The Approach of the Knowledge Based System

4.1 The KBS on CWM

The strength to integrate data warehouse metadata based on CWM metamodel is shown when data warehouse tool vendors implement the CWM specification. We define this as a passive solution. An active solution is that a KBS will apply the knowledge on CWM and the metadata integration to suggest models for the metadata integration. Our idea is that a KBS can present the definition and associations of model elements of CWM and store all the implementing

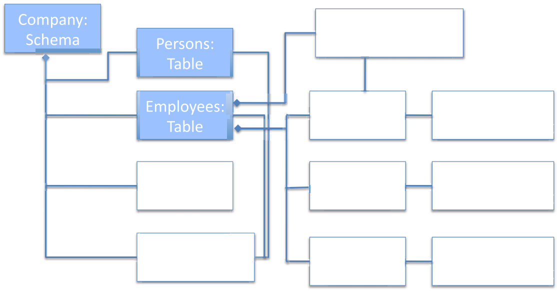


Fig. 3. CWM metadata instance of a relational database

instances. Thus the KBS can turn the CWM specification into a theoretical model, and turn the CWM implementing instances into heuristic cases. Thereby, the KBS will suggest metadata interchange models. The theoretical model of CWM will be used to verify the CWM implementing instance. The suggestions and verifications of the metadata interchange models are semi-automatic, hence change the passive status of CWM for the metadata integration.

The CWM specification is a generalization of the metadata of certain data warehouse tools. The sharing ability of information presented in CWM has been checked and refined by examining the metadata needs of several different, but representative, implementations as well as a broad range of representative warehouse configurations[9]. Therefore the KBS is designed to use logic rules to reason in two directions, i.e., from the specific metadata models to the generalization metamodel and vice versa. The reasoning is between the M2 and M1 levels in Figure 1. Moreover, with the help of the CWM implementing instances, the KBS can reason more efficiently. The CWM implementing instances can extend the definitions of the CWM metamodel because the CWM allows extending to most of the classes for specific model designs. By combining the CWM specification and the CWM implementing instances, a KBS is designed to support the metadata integration for CWM compatible and incompatible data warehouse metamodels. The KBS supports data warehouses to integrate metadata more easily and with more broader area of tools. Additionally, the KBS can improve the quality of the metadata management in general. The idea is illustrated in figure 4.

In figure 4, 'CWM based Metamodel Interchange Control' implements the CWM specification by describing the definition, the behavior and the associations of of classes in CWM specification. 'Metadata model A' is an instance of CWM. 'Metadata model translate between B and CWM' is an example of the translation of the metadata models made by the software engineers. With the knowledge of CWM specification and the collected instances such as A and B, the KBS is able to suggest the translations between CWM and the model C. The model of A, and the translated model of B and C are saved by the KBS and will enhance the KBS performance as the heuristic cases are increasing. Figure 4 also shows horizontal and vertical reasoning. The vertical reasoning includes

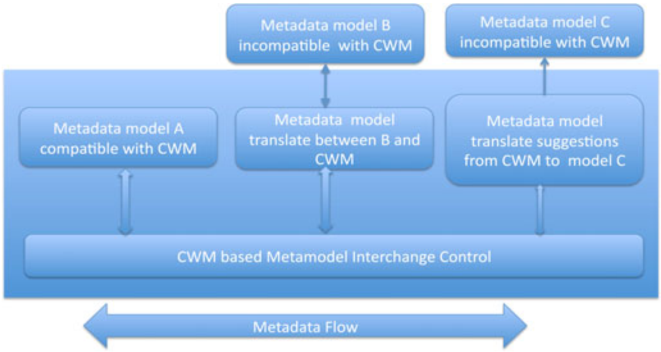


Fig. 4. The metadata integration idea of KBS

reasoning between the meta-metadata levels and the metadata levels (recalling M2 and M1 metadata levels in Figure 1), i.e., from CWM based metamodel to the specific models A, B and C and vice versa. The vertical reasoning checks, for instance, if implemented metadata models follow the CWM standard correctly. The horizontal reasoning is inside each level, for example checking the consistence between the CWM metamodels, and comparing the models among specific model A, B, and C.

The components of the KBS are shown in figure 5. The 'CWM metadata repository' stores the metadata models and the 'Metadata/metamodel model translated examples' stores the translated metadata models and the metamodels. The Knowledge base (KB) has the knowledge of the expert software engineers and can explain, reason and verify CWM metamodels and CWM instances. Ideally the KBS will be able to automatically suggest a CWM metadata interchange model.

The KB includes: definitions, associations and behaviors of CWM model elements; rules for generating CWM model instances; generalizing rules from a specific metadata model to metamodels; rules for checking consistence for levels of metamodels and the metadata models; rules for comparing metadata models and examining the similarity of the models; rules for extending CWM metamodels from the specific metadata models. With the KB, the system distinguishes elements by their associations, definitions and the behaviors. For example, Catalog in the relational package is the top level container. One of the Catalog's behaviors is to contain one or more Schemas. Schemas contain other elements, e.g. tables. The Catalog and the Schema interact as 'container' and 'ownedElement'. So KBS distinguishes Catalog and Schema by these definitions, elements' associations and behaviors. The 'relational package' supplies the context of this reasoning. Each package provides a reasoning context.

The goals of the KBS are: to implement a consistent CWM metadata management strategy; to distinguish CWM model elements and their instances by their definitions, associations and behaviors; to validate CWM metadata instances; to suggest the metadata interchange solutions to integrate data warehouse tools.

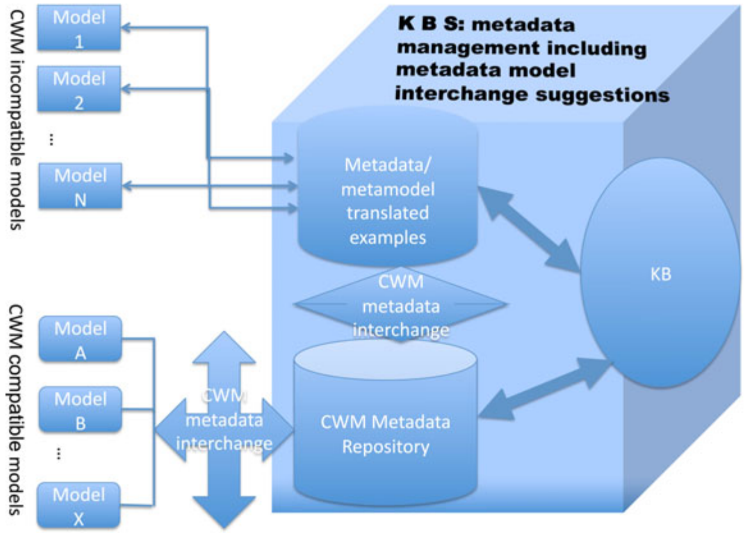


Fig. 5. The model of the Knowledge Based System

4.2 A Scenario

Whenever a data warehouse has changes in its data sources, targets or the business rules, a new metadata model might need to be integrated. To explain how the KBS can help with the metadata integration, a scenario is given. The UML diagram (in dark grey color) in figure 6 is a metadata model instance that needs to be integrated to CWM compatible models. The diagram in white color in the figure is an abstract description of the metadata model, i.e. metamodel of this model.

In the scenario, the KBS receives a task to translate a model of relational resource data into the CWM compatible system. The KBS compares it with the relational metadata instance as presented in figure 3 and the CWM metamodels. Then the KBS works to answer the following questions: (1) Is the Table and the Column the same as the CWM definition? (2) Is the PrimaryKey the same as the UniqueConstraint? (3) Is Datatype the same as the SQLSimpleType? (4) What is Domain and how to translate it into the CWM model?

To answer the first question, the KBS checks the definitions in the CWM specification and compare the definitions with the model and its metamodel. A simple rule is to assume that the same name of classes under the same context, i.e., the relational package, have the same definition, given that no conflicts between the classes are detected. To answer the second question, KBS follow the above mentioned simple rule, should suggest to translate the PrimaryKey to PrimaryKey in CWM. However, the question is to find the translation between the UniqueConstraint. KBS checks the definition of UniqueConstraint as "a condition to define uniqueness of rows in a table. An example of UniqueConstraint is a primary key[9]"; the definition of the PrimaryKey is "There is only

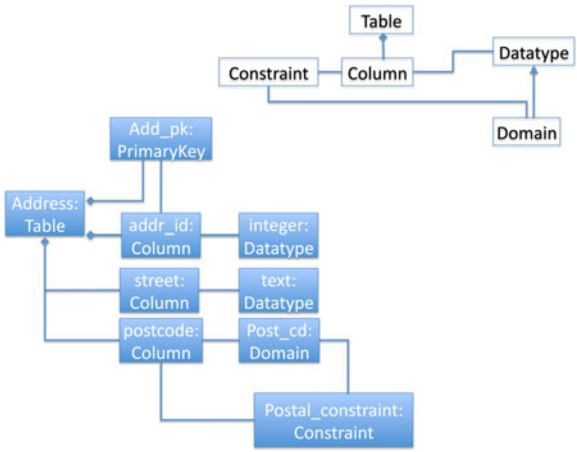


Fig. 6. An example of the CWM incompatible model

one UniqueConstraint of type PrimaryKey per Table." The CWM defines that UniqueConstraint is an ownedElement of a Table; and the UniqueConstraint is associated with the Column with UniqueKey Relationship. In figure 6, the PrimaryKey is an owned element of the table and associate to the column. Following these facts the KBS suggests to translate the PrimaryKey to an instance of the UniqueConstraint as it suggested in CWM. The KBS goes through the similar process to suggest that the Datatype can be translated into the SQLSimpleType. In the fourth question, the KBS meets the new metadata: Domain. However, the metamodel shows the relationship of the Domain to the Datatype, as well as to the Constraint, and to the Column, which are similar to the SQLDistinctType in CWM. The KBS searches the class of Column, since Column is the only known metadata related to the Domain. The KBS finds out the associations between the Column and the SQLSimpleType and the SQLDistinctType defined in CWM. KBS assumes that the Post_cd is either an instance of the SQLSimpleType or an instance of the SQLDistinctType. In this reasoning, the result from the third question, that the Datatype corresponds to the SQLSimpleType is used. Since Datatype is translated to the SQLSimpleType (result 3), the KBS assumes that the Domain relates more to the SQLDistinctType. Hence, comparing the relationship between Domain and Datatype and the relationship between SQLSimpleType and SQLDistinctType, the KBS suggests translating Domain to SQLDistinctType.

The KBS starts always from the closest concept of the models when reasoning. And then try to validate by detect logic conflicts with the integrated models. If no conflicts are found then the integrated model is verified and stored. The stored translated models are reused as heuristic instances in future reasoning.

5 Conclusions and Future Work

In this paper, we presented an approach of knowledge-based system that is applying the CWM specification and metadata model instances to support metadata integration between tools and systems. The KBS represents software engineers' knowledge of CWM and their experience of metadata interchange. The KBS also stores the translated metadata model instances. By reusing the knowledge and the stored metadata model instances, the KBS suggests metadata interchange models. Therefore, the KBS changes the ad-hoc solution into a semiautomatic process of metadata integration. The KBS also gives a better interaction with its users by explaining the suggestions of the metadata integration. With the support of the KBS, DW tool vendors and the data warehouse users will have more flexibility to choose best solutions without worrying about the metadata integration problem.

However, the KBS faces many challenges, e.g. the representation of the knowledge. The CWM specification generates many UML static models and XML files, which can be reused in the presentation of the knowledge. Hence, it is a difficult work to represent the Object Constraint Language (OCL), which is in text, with UML in the KBS. To reason from the abstract metamodel to a metadata model with the help of instances is also a challenge.

Another future work includes applying the KBS for CWM for metadata integration to other distributed systems, e.g., Internet. CWM is an example of metamodel standard. It would be very interesting to apply the KBS on other metadata standards and metadata model instances to solve the metadata integration problem.

Acknowledgements

This work is supported by Sweden-Korea Research Cooperation Program funded by the Swedish Foundation for International Cooperation in Research and Higher Education (STINT, http://www.stint.se/).

References

1. Sen, A.: Metadata management: past, present and future. Decision Support System 37, 151–173 (2004)
2. Neaga, E.I., Harding, J.A.: An enterprice modeling and integration framwork based on knowledge discovery and data mining. International Journal of Production Research 37(6), 1089–1108 (2005)
3. Medina, E., Trujillo, J.: A Standard for Representing Multidimensional Properties: The Common Warehouse Metamodel (CWM). In: Manolopoulos, Y., Návrat, P. (eds.) ADBIS 2002. LNCS, vol. 2435, pp. 232–247. Springer, Heidelberg (2002)
4. Auth, G., von Maur, E.: A software architecture for XML-based metadata interchange in data warehouse systems. In: Chaudhri, A.B., Unland, R., Djeraba, C., Lindner, W. (eds.) EDBT 2002. LNCS, vol. 2490, pp. 1–14. Springer, Heidelberg (2002)

5. Poole, J., Chang, D., Tolbert, D., Mellor, D.: Common Warehouse Metamodel, An Introduction to the Standard for Data Warehouse Integration. John Wiley & Sons, Inc., New York (2002)
6. Mazón, J.-N., Trujillo, J.: An MDA approach for the development of data warehouses. Decision Support Systems 45, 41–58 (2008)
7. Zhao, L., Siau, K.: Information Mediation Using Metamodels: An Approach Using XML and Common Warhouse Metamodel. Journal of Database Management 18(3), 69–82 (2007)
8. Staudt, M., Vaduva, A., Vetterli, T.: Metadata Management and Data Warehousing. University of Zurich (1999)
9. Object Management Group: Common Warehouse Metamodel (CWM) Specification, Version 1.1, vol. 1 (2003)
10. Kimball, R., Ross, M., Thornthwaite, W., Mundy, J., Becker, B.: The Data Warehouse Lifecycle Toolkit, 2nd edn. Wiley, Indiana (2008)
11. Inmon, W.H.: Building the Data Warehouse, 3rd edn. Wiley, New York (2005)
12. Zhao, X., Huang, Z.: A Formal Framework for Reasoning on Metadata Based on CWM. In: Embley, D.W., Olivé, A., Ram, S. (eds.) ER 2006. LNCS, vol. 4215, pp. 371–384. Springer, Heidelberg (2006)

A Ubiquitous Intelligent Tutoring System for Aiding Electronic Learning

Sergio Ciruela, Miguel Delgado, and Nicolás Marín

Dept. Computer Science and Artificial Intelligence
ETSIIT - University of Granada, 18071, Granada, Spain
{sciruela,mdelgado,nicm}@decsai.ugr.es

Abstract. In the new paradigm of Internet of Things, objects of the daily life will be part of the Web. In order to identify objects or communicate with them, tagging by radio frequency technologies or visual tags are used. The steady growth of mobile devices and increased functionality offered by these, allow them to access information into a tag using a camera or using a simple radio frequency reader. But such tags that identify things have a limited information space, so that the use of fuzzy logic and linguistic variables allow mobile devices to communicate with words compressing the data. This research presents an intelligent tutor that interacts with educational books to customize your self-assessment tasks through tags.

Keywords: Internet of things, NFC, linguistic variables, u-learning.

1 Introduction

The latest research on ubiquitous computing that can be found in the literature, focus on using different technologies to achieve new forms of communication between people and objects [8], as well as between objects only. This is intended to add a new dimension to the definition of ubiquitous computing[10], where information is accessible anytime, anywhere, to any user, but also for any object. This will create a new communications network, which will be dynamic and will consist of objects of everyday life. This new network will be known as the Internet of things.

The large number of mobile devices and wireless technologies are giving the opportunity to develop applications and innovative communications systems. A mobile device can be regarded as an environmental sensor because it has features like GPS location, movement capabilities as an accelerometer, a camera or a radio frequency reader. Although the start of some wireless technologies with mobile devices, its application in the "Internet of Things" will mean an important revolution in computing systems and communication architectures.

Under this concept of the "Internet of Things" is another psychological theory developed in the mid 80's by Edwin Hutchins [4] (1995) and it is known as "distributed cognition". This theory is based on the coordination between individuals and objects, and proposes that human knowledge and cognition are

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 70–79, 2010.

not confined to the individual but also on distributed memory spaces, facts, or knowledge of the objects in our environment. The distributed cognition may serve as a learning theory in which knowledge development is the result of a system consisting of the relationship between human agents and objects with which they interact [7].We can find an application example in the systems of distance learning through telematics platforms or other tools of computer-assisted learning. Distributed Cognition illustrates the process of interaction between people and technology, with the aim of determining how to represent, store and provide access to digital resources and other devices.

The distributed cognition makes communication with objects can be understood as the following metaphor [6]: "The data is a physical object". As an example of a metaphor of interaction, one can understand that a mobile device, "ask" or "answer" questions, "help" or "guide" a user, or simply "get information" that interacts with the environment.

Although the technologies for tagging the objects will help to deploy applications in the domain of the "Internet of Things", it is necessary to look for solutions that optimize the use of tags because they have a much reduced space. The theory of fuzzy sets and fuzzy logic can perform computing operations using words. A user can have an intelligent controller in his/her mobile device to communicate with objects inferring linguistic variables with fuzzy rules which are obtained from a tag. For example, a student may be interested in improving their skills in particular learning unit. Through tags in a book and using mobile devices can be improved the performance of a student in a particular subject with customized multimedia resources.

E-learning is one of the most innovative domains that makes use of new technologies in order to improve students' performance and facilitate the cognitive processes that are involved in learning. Through mobile devices and technologies for tagging the objects, it can provide ubiquitous learning solutions (u-Learning), which allow users to interact and communicate with objects. For example, a student who wants to know the name of an object and its meaning in another language, you simply have to take the camera from your mobile device or through a wireless card reader, and then interact with the tags that object. The labels usually save different types of information, such as a URL linking to a resource of an online encyclopedia such as Wikipedia.

In this research it is described a model for the development of ubiquitous applications using radio frequency tags that work with Near Field Communication (NFC) technology. The developed architecture enables mobile devices to communicate with wireless technologies such as NFC tags. In order to develop intelligent applications, taking into account the storage capacity offered by these wireless tags, it is necessary to use fuzzy logic and linguistic variables that allow compress the information. This intelligent tutoring system has a fuzzy inference engine that allow an approximate reasoning to a user from input variables and fuzzy rules that are retrieved from a tag of an object.

This research has the following framework, in section 2 there is a summary about NFC technology under the context of Internet of Things. Section 3 describes the research that is presented in this paper using linguistic variables with fuzzy logic, web services, mobile devices and NFC tags. Section 4 shows an illustrative example where there are code samples. Finally in section 5, the conclusions of this paper are discussed.

2 Internet of Things

Internet is growing and its subsequent evolution is through objects. Its main idea is to keep things interconnected through the Web. In this way objects may have hyperlinks, with which a user can access to multimedia resources that are on the Web. The goal of this research is to interact with objects from mobile devices using the radio frequency technology Near Field Communication (NFC).

2.1 RFID and NFC in Mobile Devices

The radio frequency identification technology (RFID) is implanted in a large number of areas, but probably will have more impact when it would be combined with mobile devices. The use of RFID and NFC technology on mobile devices offers a wide range of applications such as electronic payments, digital distribution strategies or interaction with objects of everyday life.

RFID technology covers a wide range of frequencies within the electromagnetic spectrum. In the case of mobile devices the range is 13.5 MHz, which limits the range at a distance of 3cm or almost a touch. The protocol Near Field Communication (NFC) is a radio frequency wireless streaming technology, which gives the possibility to interact with tags and interact with other devices. To connect two devices must be placed just a few inches away or put them in touch. The NFC protocol allows mobile devices to connect to a peer-to-peer (P2P) network automatically. Once the configuration data have been exchanged between the devices using NFC, then devices can use other technologies such as Bluetooth or Wi-Fi. Another of the advantages of NFC is that energy savings because it can maintain communications in an active way (MA) or a passive mode (MP). In MA, both the sender and the receiver generate a radio frequency field on which can transfer data. In MP mode only one device generates the RF field and the other device uses load modulation to transfer signal information. This enables communication between mobile devices by optimizing the use of their batteries. The information transfer rate is 106, 212, or 424 Kbits/s which should be sufficient for applications where a mobile device needs to communicate with objects.

The RFID/NFC can be implemented by Java ME technology that is supported on the most mobile devices. Therefore RFID/NFC can join the real world with the digital via the Web. With the interaction of a mobile device and a NFC tag, a mobile device might get and share information in a simply way.

3 An Intelligent Tutoring System

This section will describe an intelligent tutoring system [1] in the domain of the Internet of Things and it will use mobile devices for ubiquitous learning.

3.1 A Ubiquitous Learning System

This research presents a ubiquitous intelligent tutoring system aimed at allowing a user to carry out self-assessment at the end of a unit using a mobile device, to enhance their skills in a modern language.

For this reason, students can access to multimedia resources from a mobile device with an Internet connection. They can download customized educational files from the Web to enhance their skills.

To communicate with an object, it is necessary that this has a tag. For example, a book could have different NFC tags (see Figure 1) that might be accessible from a mobile device that has a NFC tag reader. The system can obtain a set of rules from this tag. Then if tags are used as repositories of information is necessary to optimize their storage space. Fuzzy logic with linguistic variables allows the system to compress the information by the granularity. In addition, using the approximate reasoning the mobile device resources are optimized.

Therefore, a user can perform a self-assessment of a lesson, by entering in a mobile device a numeric value between 1 and 100 from a discipline of writing, reading and listening, and speaking. Trough the interaction between the mobile device with a NFC reader and a NFC tag placed in the book, the system retrieves a set of fuzzy rules that serve to infer a result with the data previously entered. This result provides access to multimedia resources on a server via the Internet, which will support to improve the skills of the user on that topic in the disciplines most in need. This process is done through the contact between a mobile device and a book.

The system get a set of fuzzy rules and with input variables which are entered by a user, it will infer a result. Then the system will request to a web service with this result a set of multimedia files in a readable format for a mobile device. The web services can improve the mobility of multimedia resources that students need. The format of retrieved documents is text, audio or video.

3.2 Fuzzy Logic and Linguistic Variables

Are variables whose values are represented by linguistic terms [5,12]. The meaning of these linguistic terms is determined by fuzzy sets [2,11]. The linguistic variables are characterized by a quintuple, (v, T, X, g, m) where:

- v is the name of the variable,
- T is the set of linguistic terms of v,
- X is the universe of discourse of the variable v,
- g is a syntactic rule for generating linguistic terms, and
- m is a semantic rule that assigns to each linguistic term t its meaning $m(t)$, which is a fuzzy set in X.

An example of a linguistic variable is shown in figure 2. Its name, is the level of writing a modern language such as English, and it expresses the level achieved by a student in that subject for that discipline. This variable is defined by five basic linguistic terms {***very low, low, medium, high, very high***}. Each of

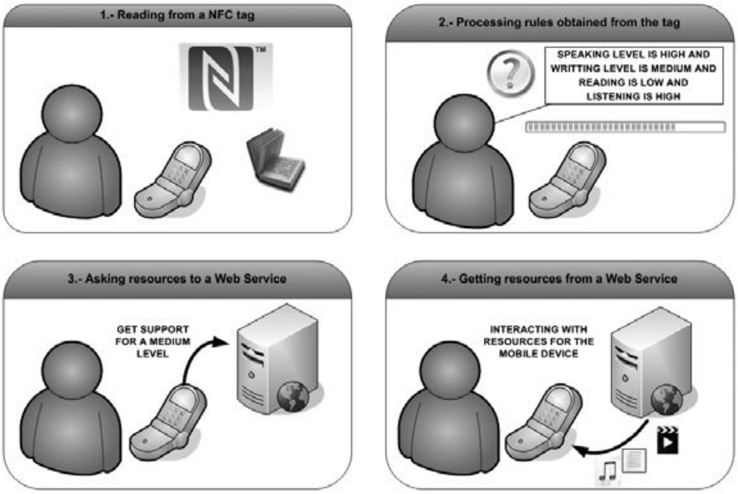


Fig. 1. Process of the ubiquitous intelligent system

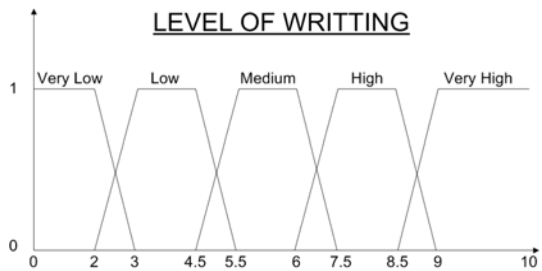


Fig. 2. Example of a linguistic variable in fuzzy logic

these basic linguistic terms are defined in an underlaying domain of knowledge from 0 to 10 that a student can improve in that discipline of a modern language.

Linguistic variables use the granulation to compress information, by including many values in a linguistic label. This is done to optimize the limited space with the labels that have NFC tags.

3.3 A Fuzzy Controller for Mobile Devices

The expert system for mobile devices is an open-source fuzzy inference engine for java, which was developed by Edward Sazonov [9]. It was adapted to Java ME. The inference engine enables the evaluation of fuzzy rules with linguistic variables that are introduced by users. The fuzzy rule base is retrieved from a NFC tag. The fuzzy engine the trapezoidal membership function to represent the most common forms of representation of fuzzy sets: the rectangular function, trapezoidal function and triangular function.

Quantifiers of a fuzzy rule can express different degrees of membership in a fuzzy set. The fuzzy controller defines three quantifiers:

- **"not"**-$X_{not} = 1 - X$
- **"very"**-$X_{very} = X^2$
- **"somewhat"**-$X_{somewhat} = \sqrt{X}$

The linguistic variables, membership functions, quantifiers and fuzzy operators are referenced by their symbolic names, represented by text strings in the following format:

```
<label> if LV1 is <hedge ...> MF1 <and/or LV2 is <hedge ...>
MF2 and/or ...> then LVN is <hedge ...> MFN ...
< and rule label weight is <hedge ...> weightMF>

<label> set LV1 is <hedge ...> MF1 <and LV2 is <hedge ...>
MF2 ...><and rule label weight is <hedge ...> weightMF>
```

- **"label"** defines a text label of a rule. The same label can be assigned to more than one rule.
- **"if"** indicates the start of the evaluation of a fuzzy rule.
- **"LV"** defines a linguistic variable and they are identified by their names. Each linguistic variable has one or more associated membership functions.
- **"is"** is a keyword that separates linguistic variables, linguistic quantifiers and membership functions.
- **"hedge"** this inference engine has three predefined quantifiers, “not”, “very” and “somewhat”. The user can use with the rules arbitrarily. Example: *if written is not very high ...*
- **"MF"** is a membership function. The membership functions are identified by their names. Each membership function is defined as a trapezoidal function.
- **"and/or"** these keywords indicate the logical operations that will be made in fuzzy expressions.
- **"then"** is a reserved word that separates antecedents from consequents of a fuzzy rule. Example: *... then support is very low.*
- **"rule"** is a reserved word that indicates the weight of a rule, and it can be modified in the expression. Example: *... then start rule: speaking is low ...*
- **"weight"** is a reserved word whose purpose is to change the weight of the rule.
- **"set"** is a reserved word for unconditional assignments.

3.4 An Architecture for Interacting with Objects

The architecture defined for NFC tags (see Figure 3) is divided into the following components:

- A **tag reader** that establish communications with radio frequency tags. From these tags are obtained a set of fuzzy rules. The fuzzy rules are stored as a xml file.

– A **fuzzy controller** to infer a result using the linguistic variables that are introduced by a user and the fuzzy rules retrieved from an NFC tag.
– A friendly **graphical user interface** for interacting between a user and a mobile device.

Fig. 3. Architecture for interacting with objects

This simple architecture allows the development of ubiquitous applications that are accessed from a mobile device with an NFC tag reader.

4 An Illustrative Example

Having seen the technologies that are used in this intelligent system, this section shows an example of how to implement this research.

Students will use the NFC tags in textbooks to improve their skills. A book could have NFC tags with relevant information about speaking, writing, reading and listening skills using fuzzy rules. The NFC tag will contain a set of fuzzy rules as shown in Table 1, and an identifier for each unit of a book to link with several multimedia resources.

A NFC card can use a markup language as XML for data format. This file will be processed in a simple way. To process the XML file is used the *k*XML parser for Java ME technology.

Table 1. Example of the rule base for improving in a modern language

IF				THEN
SPEAKING	WRITTEN	READING	LISTENING	SUPPORT
Low	Low	Low	Low	High
Low	High	High	Low	Medium
Medium	High	High	Medium	Low
Medium	Low	Low	Medium	High
High	High	High	High	Low
High	Low	Low	High	High

```
<?xml version="1.0" encoding="UTF-8"?>
<NFC>
<id>1039</id>
<rules>
<rule id="1">If Speaking is Low and Written is Low and Reading is Low and
 Listening is Low then Support is High</rule>
<rule id="2">If Speaking is Low and Written is High and Reading is High and
 Listening is Low then Support is Medium</rule>
<rule id="3">If Speaking is Medium and Written is Low and Reading is Low and
 Listening is Medium then Support is High</rule>
<rule id="4">If Speaking is Medium and Written is Low and Reading is Low and
 Listening is Medium then Support is High</rule>
<rule id="5">If Speaking is High and Written is Low and Reading is Low and
 Listening is Low then Support is High</rule>
<rule id="6">If Speaking is High and Written is High and Reading is High and
 Listening is High then Support is Low</rule>
...
</rules>
</NFC>
```

To implement the fuzzy controller is necessary to define the linguistic variables with the underlying numerical domain. Having defined the linguistic labels in the system, fuzzy rules are introduced in the application which have been obtained from the NFC tag. The following code shows an example:

```
// Create Lingustic variables and define membership functions
LinguisticVariable spoken = new LinguisticVariable("speaking");
spoken.add("high",7.5,8,8.5,10);
spoken.add("medium",4.5,5,6,8);
spoken.add("low",0,0,4,5);
LinguisticVariable written = new LinguisticVariable("written");
written.add("high",7.5,8,8.5,10);
written.add("medium",4.5,5,6,8);
written.add("low",0,0,4,5);
LinguisticVariable reading = new LinguisticVariable("reading");
reading.add("high",7.5,8,8.5,10);
reading.add("medium",4.5,5,6,8);
reading.add("low",0,0,4,5);
LinguisticVariable reading = new LinguisticVariable("listening");
reading.add("high",7.5,8,8.5,10);
reading.add("medium",4.5,5,6,8);
reading.add("low",0,0,4,5);
LinguisticVariable support = new LinguisticVariable("support");
support.add("high",7.5,8,8.5,10);
support.add("medium",4.5,5,6,8);
support.add("low",0,0,4,5);
// Create a fuzzy engine
fuzzyEngine = new FuzzyEngine();
// Register all Linguistic Variables
fuzzyEngine.register(speaking);
fuzzyEngine.register(written);
fuzzyEngine.register(reading);
fuzzyEngine.register(listening);
fuzzyEngine.register(support);
// Create a block of rules
String rule="if speaking is medium and written is high and reading is low";
rule+="and listening is medium then support is medium";
fuzzyBlockOfRules = new FuzzyBlockOfRules(rule);
// Register the block
fuzzyEngine.register(fuzzyBlockOfRules);
// Parse the rules
fuzzyBlockOfRules.parseBlock();
// Perform the evaluation
fuzzyBlockOfRules.evaluateBlock(); // faster execution
// Obtain the result
double result = support.defuzzify();
```

Then a user can introduce ratings about his/her skills in a modern language through a mobile device. The interaction with the device will be easy for the user. Once a result is inferred, a request to a web service via the Internet to get personalized multimedia resources. The REST [3] architectural style is ideal for deploying the Web service, because it is based on the principle of "hyperlink" to digital resources via URIs. To process the XML document the system uses the library kXML and it is used the library MMAPI for managing multimedia files in Java ME with mobile devices. Figure 4 shows a screenshot of the system.

Fig. 4. Example of the interaction between the user and the system

5 Conclusions

This research made possible to carry out a ubiquitous application for interacting with objects of the environment in an intelligent way. The ubiquitous intelligent system has been implemented under the electronic learning domain, which is one of the most innovative domains technologically. The following are the conclusions of this research:

– The NFC tags are suitable for use as repositories of information despite their limited space. For this reason, fuzzy logic and linguistic variables are an important tool to optimize the storage space of the tags. With NFC tags a student can interact with objects such as a textbook. Nfc labels as a complement to the Internet, support for developing intelligent mobile applications.
– In addition,fuzzy logic allows the development of cognitive systems through approximate reasoning, which allows working with devices that have limited features such as mobile devices.
– A Mobile device applied to e-learning is becoming more used because it can bring an expert system that simulates an intelligent tutor and it helps students to assess the difficulty of contents. A mobile device can recommend multimedia resources that are on the web to improve students skills.

Acknowledgments

This research was supported by the Andalusian Autonomous Government under Grants TIC02786 and TIC03175.

References

1. Chen, C.-M., Lee, H.-M., Chen, Y.-H.: Personalized e-learning system using Item Response Theory. Computers & Education 44(3), 237–255 (2005)
2. Delgado, M., Verdegay, J.L., Vila, M.A.: On aggregation operations of linguistic labels. Internat. J. Intell. Systems 7, 479–492 (1993)
3. Fielding, R.T.: Principled Design of the Morden Web Architecture. ACM Transactions on Internet Technology 2(2), 115–150 (2002)
4. Hutchins, E.: How a cockpit remembers its speeds. Cognitive Science 19, 265–288 (1995)
5. Klir, G.J., Yuan, B., Saint Clair, U.H.: Fuzzy Set Theory: Foundations and Applications. Prentice-Hall, Englewood Cliffs (1997)
6. Lakoff, G.: The contemporary theory of metaphor. In: Methapor and thought, pp. 202–251. Cambridge University, Cambridge (1993)
7. Perry, M.: Distributed Cognition. In: Carroll, J.M. (ed.) HCI models, theories, and frameworks: Toward an interdisciplinary science, pp. 193–223. Morgan Kaufmann, San Francisco (2003)
8. Satyanarayanan, M.: Pervasive computing: vision and challenges. IEEE Personal Communications 8(4), 10–17 (2001)
9. Sazonov, E.S., Klinkhachorn, P., GanfaRao, H.V.S., Halabe, U.B.: Fuzzy logic expert system for automated damage detection from changes in strain energy mode shapes. Non-Destructive Testing and Evaluation 18(1), 1–17 (2002)
10. Weiser, M.: The Computer for the Twenty-First Century. Scientific American 94 (1991), `http://nano.xerox.com/hypertext/weiser/SciAmDraft3.html` (Accessed March 10, 2010)
11. Zadeh, L.A.: Fuzzy Sets. Information and Control 8, 338–353 (1965)
12. Zadeh, L.A.: The Concept of a Linguistic Variable and its Application to Approximate Reasoning-I. Information Sciences 8(3), 199–249 (1975)

Towards a Proposal for a Vessel Knowledge Representation Model

I. Macía[1,2], M. Graña[2], and C. Paloc[1]

[1] Vicomtech, Visual Communications Technologies Centre
http://www.vicomtech.org
[2] Grupo de Inteligencia Computacional, UPV/EHU
http://www.ehu.es/ccwtinco

Abstract. We propose the development of a knowledge representation model in the area of Blood Vessel analysis, whose need we feel for the future development of the field and for our own research efforts. It will allow easy reuse of software pieces through appropriate abstractions, facilitating the development of innovative methods, procedures and applications. In this paper we present some key ideas that will be fully developed elsewhere.

1 Introduction

The vessel structure of the blood circulatory system is one of the most complex structures of the body. Blood vessel anatomy has been studied from castings and in-vivo examinations in order to build models that provide valuable insight into the normal and variant circulatory anatomy, helping to understand the causes, evolution and outcome of several vascular-related diseases. However, many answers to simple questions about vascular morphology and angiogenesis remain open [9]. Recent advances on medical imaging provide high resolution images of the vessel structures, so that the generation of accurate patient-specific geometric *in-vivo* vessel models [1] and related quantitative measurements has become feasible. This has resulted in a wide range of new applications for computer-assisted diagnostic, intervention and follow-up of vascular-related diseases.

The diversity of medical and biological applications and the availability of huge amounts of high-quality information for vessel analysis has raised the problem of vascular knowledge representation in its full multi-faceted complexity . The purpose of this paper is to present some key ideas for the development of appropriate knowledge representation and manipulation tools for vessel structures which could serve as a common ground for the development of compatible and reusable systems. We motivate this study in the diversity of applications found in the literature, and in our actual research experience [7,5,8]. We aim towards a *Vessel Knowledge Representation* (VKR) model that, due to its efficiency and versatility, may be used for a wide variety of image-based vessel extraction schemes and vessel analysis applications.

The structure of the paper is as follows. In Section 1 we provide a review of the topics of interest related to the definition of the VKR model: knowledge

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 80–87, 2010.

on vascular morphology, vessel-related diseases, angiographic diagnosis, vessel extraction and analysis techniques from these images and corresponding applications. As a corolary, in Section 2 we describe the requirements for a Vessel Knowledge Representation model, whose description is sketched in Section 3. Finally we provide final conclusions (Section 4).

2 Requirements of the Vessel Knowledge Representation (VKR) Model

The VKR model is being defined through a process of identification and abstraction of structural, geometrical and morphological properties of vessels in the literature and in our own research experience. This leads to the identification of data structures, operations and components used in the most common models and schemes for vessel extraction. This model can then be converted into an appropriate data representation, such as a mesh surface model, a refined segmentation or a symbolic visual representation. When rendered, these representations can be used for localization and for interactive exploration of the VKR model and underlying properties in some of the applications described above. Alternatively, these derived data representations can also be used, for example, for numerical studies, such as simulations of haemodynamics, structural analysis or other medical and research applications out of the scope of this paper. The VKR model must include the geometry and topology of vessel trees with constituting branches, bifurcations and sections, as well as vascular accidents such as stenoses, aneurysms and abnormal regions, such as those feeding neighboring tumors. Models of these physical entities and related concepts used in vessel analysis applications must be devised and structured by using object-oriented design techniques.

We can make more precise some desired properties of our VKR model design:

- *Versatility*:
  - Modelling of low level entities, such as vessel centerlines or sections, without compromising higher level elements, such as the global graph-based model of the vessel tree and its traversal mechanisms.
  - Allowing several coexisting representations of the same vascular system, providing easy transformation among representations.
  - Decoupling algorithms from underlying data structures. Abstract mechanisms must be provided for accessing, traversing and manipulating the data.
- *Efficiency*: as data amounts are huge in this kind of applications, and time requirements are increasingly tight, efficiency in terms of computational time and use of resources is highly desired.
- *Utility*: to be useful the VKR must take into account actual design practices and constraints from:
  - The vessel extraction algorithms used for generating the vessel data structure from the angiographic image data.

  - A broad range of clinical and research applications that will be increasing in complexity and response time requirements.
- *Complexity Hierarchy*: the framework should be able to provide different levels of complexity and abstraction in order to represent the vessel structures at different levels. The structures need to be represented at least at the tree, branch and section level and at each level geometric, topological and semantic information layers need to be managed.
- *Integrability and Specificity*: the framework needs to be designed so it can be easily integrated into pre-existing frameworks which deal with certain specific models, processes and data structures efficiently, such as the *Insight Toolkit* [13], for medical image segmentation, registration and analysis and the *Visualization Toolkit* [10], for visualization of resulting vascular structures together with image data.

3 Model Description

The VKR model is the core of the diverse operations and functions related with vessel analysis techniques, as shown in the workflow diagram depicted in Figure 1. The boxes in this diagram correspond to data types of some kind, while the labeled arrows correspond to transformations or manipulations of the data. We have omitted the closed operations, such as branch pruning or image filtering. The VKR vessel representation can be obtained directly from the angiographic image or volume or indirectly from the results of an intermediate image segmentation process. In the latter case, the segmentation detects the image/volume regions corresponding to the vessels, from which the vessel representation can be obtained by skeletonization, to obtain the centerlines, followed by section or boundary estimation. Alternatively, a set of disconnected volume vessel regions can be obtained by a global detection process of vessel features, followed by pruning and/or reconnection of centerline patches. We include in the diagram obvious storage and retrieval operation of the VKR to/from a file or database. The VKR model is the natural domain to perform measurements which can be added to it as an enrichment.

By assigning symbolic graphical representations or *glyphs* (such as lines, spheres, cones or more complicated shapes...) to the underlying components of the VKR model, a symbolic visual representation of the vessel tree can be obtained. This may be used as a roadmap, for agile exploration and interaction, or may be directly overlaid or projected onto the angiographic images, slices, or volumes in order to provide visual cues.

The VKR model could be the basis to build up a surface mesh of the vessel boundaries by several techniques such as contour sweeping of the cross-sections or by an explicit or implicit surface model. The VKR model could also be used to generate a mask or ROI on the CTA/MRA volume for further processes. The VKR data could then be converted into a mesh surface by iso-surface reconstruction [6]. Generated surface meshes could then be used for direct visualization and navigation, possibly mixed with other symbolic, surface, volume or slice renderings, in a kind of Augmented Reality computational environment. In the same spirit, the identified and labeled branches can be mapped into the CTA/MRA volume or mesh surface, allowing increased interaction via direct structure picking. The mapping can go both ways, allowing the access to the VKR model from the visualization of the CTA/MRA volume, and visualization of CTA/MRA data corresponding to VKR selections.

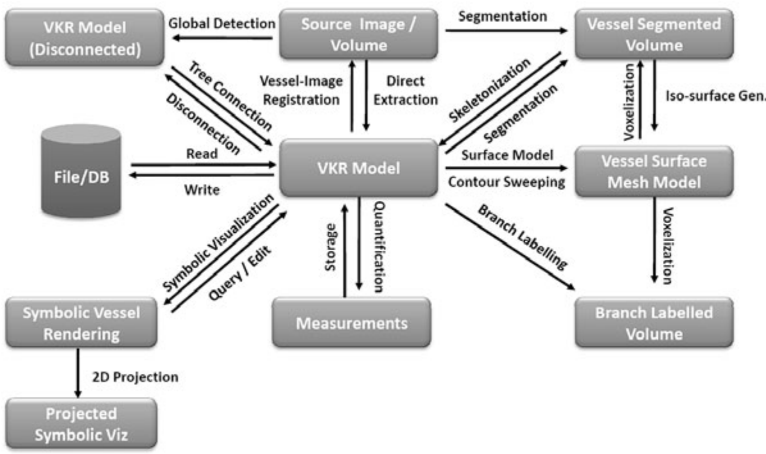


Fig. 1. VKR Workflow Diagram

3.1 Data Structures

Vessel Graph. In general, we can consider the vessel network as a binary tree structure since in most cases bifurcations split a branch into two [4], with some exceptions like the Circle of Willis in the brain [11]. Therefore, a graph representation is the natural choice for the structural representation in the VKR. A graph typically consists of nodes, representing the modelled concepts, and edges, that connect the nodes and represent their relationships, which is in terms of parent/child for tree structured graphs. In our case, a *Vessel Node* represents an abstraction of an element used for vessel representation and analysis at graph level. Such an element may be a vessel branch, bifurcation or vessel accident, among others. Anatomical vessel branches are modelled as nodes (*Branch Node*) and if we need to assign properties to the bifurcations, we can also explicitly model them (*Bifurcation Node*). In order to provide more modelling flexibility, we define also *Composite* nodes, which make use of the *Composite Pattern* [3] in order to group nodes. This way the group of nodes acts as a single entity, hiding their internal relationships and offering the possibility of building a hierarchy of several levels of abstraction complexity in the graph.

Vessel Branch. A virtual vessel branch is represented in VKR by a *BranchNode,* and it corresponds to the vessel segment that extends between consecutive bifurcations. A physical vessel branch may also be represented by several concatenated *BranchNode* instances. This would be useful when the user wants to make a difference between different parts along the length of a physical branch, for example by indicating that part of a branch is stenosed. This is performed by associating corresponding accident node representations, such as the *Stenosis Model*, to the *Branch Node*. The core of a vessel branch in our model is the *Centerline Model*.

Centerline Model. The vessel centerline or medial loci [2] is an important part of our model, since it is a good descriptor of elongated objects. Compared to other descriptors, such as boundary descriptors, the centerline captures better the vessel shape and provides a straightforward way of obtaining the relationships between the different branches of the vessel tree [12], since the centerline can be easily converted into a graph structure. Furthermore, it serves as a reference for calculating and storing local properties, both inside and on the boundaries of vessels. For example, the vessel length is measured along the centerlines and diameters are measured over sections whose center is the centerline. Therefore, we aim to provide an explicit, yet flexible and agile, representation of the centerline.

Section Model. *Vessel Sections* are localized at centerline points and they are assumed to vary along the vessel length. This variability is reflected in the parameters that define the section, for example, the diameter. Vessel sections, like centerlines, can also be defined at increasing levels of complexity. The simplest level is to define the section as a circle, giving its center and radius/diameter. Since our sections are defined at explicit centerline points, the center is already given. The next level of complexity is an elliptical shape. More advanced mathematical models include radial functions and B-spline contours.

3D Surface and Voxel Models of Vessels. So far we have dealt with explicit modelling of cross-sections. Another possibility, when dealing with 3D image data, is to directly generate a 3D surface mesh from the centerline. If the 3D mesh is generated for the complete vessel tree, it can be referred to branches or even to centerline points (and thus to sections) of the VKR model by proximity to the corresponding centerline. This reference can be direct, by splitting the model into surface patches and keeping references to them, or indirect, simply by associating a scalar value, acting as identifier, to the mesh points that corresponds to referred branches. This way a forth-and-back relationship may be kept between the VKR and surface models. Explicit sections may also be obtained by intersection with corresponding section planes.

If a segmentation is available, obtained either *a priori* or from the VKR model, it can be referred to corresponding branches by just labelling the mask pixels/voxels with corresponding branch identifiers. In this case, keeping references to separate volume “patches” seems to be more difficult to handle but it is a possibility that could be useful in cases where the source angiographic volumes are huge. The reason is that, in most software frameworks, only arrays corresponding to rectilinear volumes can be stored, and for sparse structures such as the vessels, sometimes many of these voxels are empty. Another possibility is to store these labelled voxels as sparse images, which is currently not implemented.

Vessel Bifurcations. Bifurcations may be represented explicitly in the VKR model by means of the *BifurcationNode* object that defined at graph level. The use of this node type would be optional, and may be required when we want to model special features of the bifurcation, when (quantification) operations need to be assigned to the bifurcation, such as estimation of branch angles, and when there may exist more than one parent branch.

Vessel Features. In the VKR model, vessel "features" (*Feature Node*) represent special characteristics of the vessels that need to be highlighted. Their definition may include models for vessel accidents or simply comments used for diagnostic. A feature may affect or may be associated to a part of a branch, a whole branch or a set of branches, entirely or partially. In order to make explicit these relationships, two mechanisms are devised:

1. *FeaturesNodes* are assigned as children (or alternatively as parents) of affected *Branch Nodes.*
2. *FeatureNodes* keep a *Vessel Region* structure that indicates which vessel branches are affected and to which extent. This is achieved by keeping a set of *Vessel Branch* node identifiers, and for each identifier, the starting and end indexes of the points in corresponding centerlines that comprise the area affected by the feature.

Since a feature may affect more than one branch, *Feature Nodes* are treated in a special manner and are not even visited when performing many operations that require traversal of the graph. In this sense, FeatureNodes can be treated as "hypernodes" and their relationship with *Vessel Branch* nodes (or possibly other nodes) is not that of a parent-child relationship but merely a reference.

An example of use of a *Feature Node* is to perform an annotation, such as a diagnostic remark in a application for computer aided vascular diagnosis. The clinician would choose the branches affected by a given feature, for example, those feeding a tumor or included on it, and assign them the corresponding nodes comment.

Models of Vessel Accidents or Disease. The VKR model offers the possibility of providing representation models for vessel accident or disease. Examples of these models are the *Stenosis Model* and *Aneurysm Model.* These models contain the quantitative morphological measurements and other properties that are typical of a given vessel accident or related disease. We aim to provide flexibility for defining application-specific models of this kind.

3.2 Supported Operations

Operations that can be performed on the VKR model data structures can be classified by their nature or by the type of object they operate on. For example, quantification operations can be performed at graph, branch, centerline or section level, among others. Based on their nature we distinguish the following types:

- *Access Operations*: these are abstract access mechanisms that allows to perform other types of operations. For example, graph traversal is an operation that allows to access nodes on the vessel graph and perform other operations on them.
- *Edition Operations*: allows to change the internal structure and properties of the model.
- *Quantification Operations*: evaluation of quantitative measurements over different elements of the model.
- *Input/Output Operations*: used to load and save the model data.
- *Data Transformation Operations*: include generation of the VKR model and transformation into another representation that can be useful for intended applications.
- *Model-specific Operations*: these are internal operations that are specific to certain elements of the model, such as the centerlines or sections.

4 Conclusions

The applications and techniques of Blood Vessel Analysis have produced a complex landscape of algorithms and data representations that hinders the composition of procedures, the reuse of software and the comparative analysis in terms of computational efficiency and quality of final results (visualization, measurement, edition, and others). We have detected the need of proposing a foundational Vessel Knowledge Representation (VKR) model that may allow the exchange of data among applications and users. One of the goals of VKR is the reuse of software pieces, providing a ground functional layer that may serve as the basis for new developments, thus alleviating development efforts. The model is intended to be used as an intermediate representation between image-based extraction schemes and clinical and research applications, to perform quantitative measurements on extracted vessel structures and to provide the necessary vessel representation and handling tools for the target applications. In this paper we have identified, from the literature and our own research work, some key knowledge representation items, as well as the key operations that are the building blocks for nowadays and future vessel analysis processes and applications. We are already applying the VKR model in vessel-related applications related to our current research areas.

References

1. Antiga, L., Ene-Iordache, B., Remuzzi, A.: Computational geometry for patient-specific reconstruction and meshing of blood vessels from mr and ct angiography. IEEE Transactions on Medical Imaging 22(5), 674–684 (2003)
2. Blum, H.: A transformation for extracting new descriptors of shape. In: Wathen-Dunn, W. (ed.) Models for the Perception of Speech and Visual Form, pp. 362–380. MIT Press, Cambridge (1967)
3. Gamma, E., Helm, R., Johnson, R., Vlissides, J.: Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley, Reading (2000)
4. Huang, W., Yen, R.T., McLaurine, M., Bledsoe, G.: Morphometry of the human pulmonary vasculature. Journal of Applied Physiology 81, 2123–2133 (1996)
5. Legarreta, J.H., Boto, F., Macía, I., Maiora, J., García, G., Paloc, C., Graña, M., de Blas, M.: Hybrid decision support system for endovascular aortic aneurism repair follow-up. In: Graña Romay, M., Corchado, E., Garcia Sebastian, M.T. (eds.) Hybrid Artificial Intelligence Systems. LNCS, vol. 6076, pp. 500–507. Springer, Heidelberg (2010)
6. Lorensen, W.E., Cline, H.E.: Marching cubes: A high resolution 3d surface construction algorithm. Computer Graphics 21(4), 163–169 (1987)
7. Macía, I., Graña, M., Paloc, C., Boto, F.: Neural network-based detection of type ii endoleaks in cta images after endovascular repair. Int. J. of Neural Systems (submited 2010)
8. Maiora, J., García, G., Macía, I., Legarreta, J.H., Paloc, C., de Blas, M., Graña, M.: Thrombus change detection after endovascular abdominal aortic aneurysm repair. In: Proc. of Computer Assisted Radiology and Surgery, CARS 2010 (2010) (in Press)
9. Santamore, W.P., Bove, A.A.: Why arteries are the size they are. Journal of Applied Physiology 104, 1259 (2008)
10. Schroeder, W., Martin, K., Lorensen, B.: The Visualization Toolkit: An Object-Oriented Approach to 3D Graphics, 4th edn., Kitware (2006)
11. Schwarz, M., Nguyen, M.P., Kiencke, U., Heilmann, C., Klemm, R., Benk, C., Beyersdorf, F., Busch, H.J.: Integration of the circle of willis into avolio's model of the arterial hemodynamics. In: 6th IASTED Int. Conf. on Biomedical Engineering (2008)
12. Siddiqi, K., Pizer, S.: Medial Representations: Mathematics, Algorithms and Applications. Springer, Heidelberg (2008), ISBN 978-1402086571
13. Yoo, T.S., Ackerman, M.J., Lorensen, W.E., Schroeder, W., Chalana, V., Aylward, S., Metaxas, D., Whitaker, R.: Engineering and algorithm design for an image processing api: a technical report on itk - the insight toolkit. Stud. Health. Technol. Inform. 85, 586–592 (2002)

Developing a Probabilistic Graphical Structure from a Model of Mental-Health Clinical Risk Expertise

Olufunmilayo Obembe and Christopher D. Buckingham

Aston University, Aston Triangle, Birmingham B4 7ET, UK
{obembeo,c.d.buckingham}@aston.ac.uk

Abstract. This paper explores the process of developing a principled approach for translating a model of mental-health risk expertise into a probabilistic graphical structure. The Galatean Risk Screening Tool [1] is a psychological model for mental health risk assessment based on fuzzy sets. This paper details how the knowledge encapsulated in the psychological model was used to develop the structure of the probability graph by exploiting the semantics of the clinical expertise. These semantics are formalised by a detailed specification for an XML structure used to represent the expertise. The component parts were then mapped to equivalent probabilistic graphical structures such as Bayesian Belief Nets and Markov Random Fields to produce a composite chain graph that provides a probabilistic classification of risk expertise to complement the expert clinical judgements.

Keywords: probability graph, knowledge representation, psychological modelling, clinical decision support system, mental health risk assessment, chain graph.

1 Introduction

Suicide is closely related to mental illness and the Department of Health estimates that 25 percent of people who commit suicide in the United Kingdom have been in contact with mental health services in the year prior to their deaths (this represents about 1500 cases per year) [2]. A reliable method for the assessment of suicide risk would therefore be a valuable tool for society at large in averting needless deaths. From a medical point of view, one of the main objectives of such a tool would be to classify a patient based on some risk factors into categories of high and low risk, such as suicide, self-harm, harm to others, and self-neglect.

Risk assessments have traditionally been undertaken in a variety of ways, which can be divided into two main approaches: structured clinical judgements based on human expertise and actuarial risk assessments, which use statistical calculations of probability [3]. Hence risk assessment models can be grounded in both subjective and empirical data. Combining them in a holistic approach is likely to provide the most efficient results and graphical probabilistic models can help facilitate this approach [4]. This paper will explore how the clinical judgements encapsulated by the

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 88–97, 2010.

Galatean mental-health Risk and Social care assessment Tool (GRiST) can inform a probabilistic graph that provides a mathematical analysis of risk data to complement the human expertise. The next section will introduce GRiST, discuss the formal specifications of its various hierarchical components, and consider their relationships to probability graphical structures. We then synthesize a GRiST probability graph from the building block structures and conclude by highlighting future directions for the research.

2 Galatean Risk Screening Tool, GRiST

GRiST is a clinical decision support system for mental health risk assessments [1] that represents knowledge in a hierarchical tree structure for generating risk evaluations. The encapsulated expertise was originally elicited from 46 domain experts [5] but subsequently validated and refined [6] with the involvement of over 100 multidisciplinary clinicians. The GRiST tree is an XML structure made up of a set of nodes, which can either be concepts or datum nodes. A datum node is a component value representing a physically measurable item of information or cue influencing risk evaluations (e.g. the seriousness of intention to commit suicide, as shown in Figure 1). Datum nodes equate to the information gathered during an assessment and are the input values to the tree (i.e. the leaf nodes). Concept nodes are the higher-level nodes in the tree consisting of two or more subcomponents that could be datum nodes or other concept nodes. They represent composite concepts underlying risk such as social context, feelings and emotions, or depression.

2.1 GRiST Uncertainty Representation and Propagation

In the Galatean model, uncertainty is represented by using two main measures, relative influences and fuzzy-set membership grades.

Relative Influence (RI) represents the influence or weight a node has on its parent concept, relative to its siblings. Within the GRiST knowledge structure (GKS) there is a constraint stating that the total sum of RIs across the siblings must equal one.

Membership Grade (MG) represents the degree of membership of an object in a node, where each node is considered to be a fuzzy set.

In the GKS, datum or leaf nodes match associated patient cues, which generates an MG from an MG distribution that has been defined by clinical experts as part of the elicitation process for the decision support system. These MGs feed through the concept hierarchy based on the RI values attached to nodes as illustrated by Figure 1 and equation (1) (see [1] for further details).

Equation (1) states that that the membership grade of a concept, *X,* such as ‘intention’ in Figure 1, is equal to the sum of the MGs of the datum nodes along all paths *p* that lead to X multiplied by all the corresponding RI values along the paths on each level *l* leading to X

$$MG(X) = \sum_{p=1}^{P} (MG(datum_p) \times \prod_{l=1}^{L} RI_{lp}) \tag{1}$$

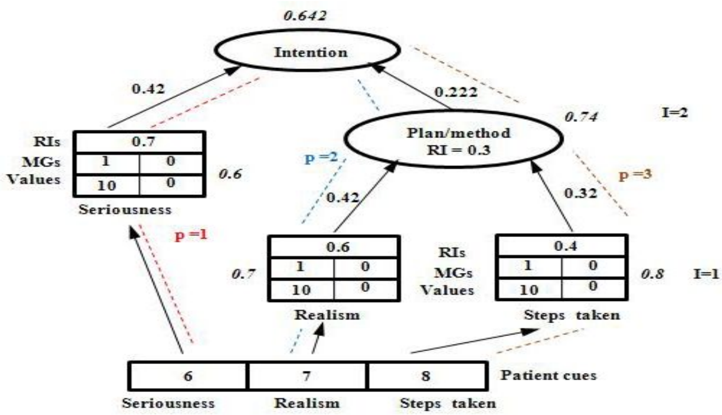


Fig. 1. Example of the propagation of membership grades in the GRiST model, highlighting the path to levels, p, for the concept node *intention*

2.2 GRiST Data Structures

Attempting to translate the GKS into a probabilistic equivalent requires understanding GRiST's uncertainty processing in conjunction with the semantics of its knowledge structures and any constraints operating on them. The concept and datum nodes are separated into the following types:

1. **Non-Generic Concept** – These are concept nodes occurring just once in the model.
2. **Generic Concept** – These are concept nodes that occur in more than one location but always with the identical structure. They may be one of two subtypes:
   a. **Non-Context Dependent Generic Concepts**, generic concepts (g) which have exactly the same RI and MG parameters wherever the concept node occurs;
   b. **Context Dependent Generic Concepts**, which are known as generic distinct (gd) nodes because their uncertainty values (RIs) are dependent on the location of the nodes in the hierarchical tree.
3. **Datum nodes** that are also subdivided into context dependent (generic distinct) and non-context-dependent nodes, which may also occur only once within a risk concept or several times (generic datum nodes). Generic distinct datum nodes have different MG distributions for the matching patient cue in different tree locations (e.g. the same level of assertiveness has a different risk for suicide than harm to others).

3 GRiST Component Structures and Their Relationships to Probability Graphical Structures

Now that the uncertainty processing and node structure types have been described, it is necessary to consider the different types of relationships or links existing between nodes in the GKS. This will enable the GKS to be divided into categories of equivalent graphical structures and each one can be translated into an associated probabilistic structure. The task then becomes one of constructing a single probability graph out of the smaller building blocks but an important issue to address is ensuring that conditional independence represented in the GSK is maintained during the translation process, as advised in [7] and [8]. However, unlike in [7] and [8], the variables in the components that we are aggregating do not overlap (i.e. we are not combining the overlapping data from different models). This distinction makes it easier for the integrity of the conditional independence relationships to be maintained during the conversion process.

3.1 Types of Probabilistic Graphical Models

There are two main types of graphical models, those based on undirected graphs, which are known as Markov Random Fields (MRFs) and model symmetric relationships between their variables, and those based on directed graphs known as Bayesian Belief Networks (BBNs) that model causal relationships between their variables. A BBN is a directed acyclic graph (DAG) consisting of nodes, which represent random variables in the domain being modelled, and directed arcs between the nodes, which represent direct dependencies. An MRF on the other hand is an undirected graph that, unlike BBNs, does not encode causal relations. However, in the MRF the probability distribution of each variable is dependent only on its graphical neighbors (local MRF Markov property). For more details see [9] for BBNs and [10] for MRFs.

3.2 Identification of Relationship Types between Variables

From a semantic perspective three main types of relationships were identified between the GKS variables:

1. **IS-A:** the *IS-A* link can have a variety of meanings with subtle differences depending on the variables being modelled [11]. Some of these meanings include a class relationship, generalisation/specialisation relationship, and so on. In the GKS the definition of the type of *IS-A* link found is the 'a kind of' definition. Examples of the *IS-A* link include the parent node *feelings and emotions* and its children nodes of *angry emotions*, *anxious emotions, helplessness,* and *sad* amongst others. With this kind of relationship the children nodes are associated through the parent node.
2. **Contribute-to:** This represents the case where the children nodes contribute to and influence the parent node. For example the parent node *constraints on suicidal behaviour* and its children nodes *insight and responsibility* and *religious values/beliefs affecting suicide risk*. In this type of relationship the children nodes have a direct influence on the parent node. For instance a person's religious

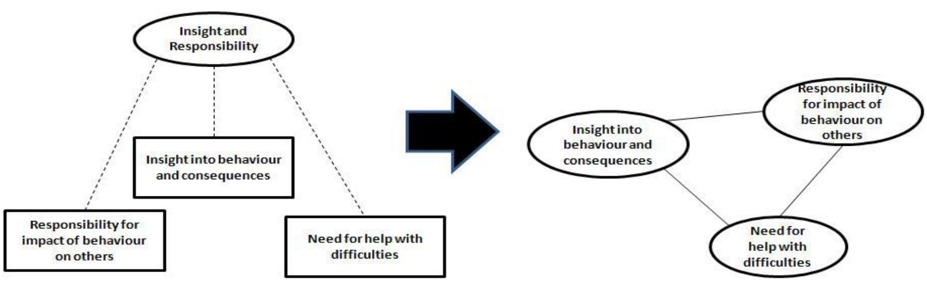
beliefs on suicide and their insight and responsibility will directly affect the constraints that they have when it comes to suicidal behaviour. Of the four identified relationship types identified in the GKS this is the only one that can be assumed to be causal.

3. **Wrapper:** This represents the case when a parent node serves as a sort of 'cover name' for a set of related children nodes (i.e. it 'wraps' the variables together). For instance parent node *pattern of suicide attempts* for children nodes *first time suicide attempt occurred*, *how many suicide attempts* and *suicide attempts escalating in frequency*. For these types of relations, there is a correlation between the children nodes, for instance whether or not the attempts are escalating in frequency correlates with the number of suicide attempts. However in this case the parent node is not made up of the children nodes and the inherent correlation cannot be assumed to be causal.

The next sections will take the categories of node structures and their relationships to the siblings to define some types of structures that can be translated into probabilistic equivalents.

3.3 The Fixed Generic Component Structure (FG)

The FG structure comprises of a generic root (parent) node and children nodes that can be of any type but the physical structure and uncertainty values are identical regardless of its location in the GKS. This means it can be considered to be independent of any other nodes not already contained within it because no matter where it occurs in the GKS, the MG of the root node and all sub-nodes are invariant. Figure 2(a) is an example of an FG structure taken from the GRiST GKS where sibling children nodes are related through their parents.



(a) (b)

Fig. 2. (a) Sample fixed generic component from GRiST knowledge structure and (b) its MRF equivalent structure

Variables that are symmetrically related can be represented in different ways, two of which are depicted in Figure 2. [12] describes the modelling of the associations between nodes of this form, as representing the association between the variables on each edge with weighted vectors as shown in figure 2(b). For associations of the form depicted in figure 2(a), they mention the use of a hidden node or variable, which is the generic parent node *Insight and Responsibility* in the figure, where the hidden variable is connected to each of the other variables. This hidden variable brings about a probability distribution for each value it takes over all the connected nodes [12]. An alternate viewpoint on these two ways of modelling symmetric associations is given by [13] where the children variables in figure 2(a) are said to be conditioned on the parent generic node *Insight and Responsibility* and in figure 2(b) are said to be marginalised over it. The above observations lead to the conclusion that for fixed generic components within the GKS, sibling children nodes are equivalent to neighbour nodes in a MRF, where neighbour nodes are nodes that are directly linked to each other (as in figure 2(b)). Bearing in mind our main objective, the relevant question at this point is then how can we relate the FG structure, which is independent of all nodes except its component nodes, to a probabilistic graphical structure?

We know that the parent (root) node of an FG structure is independent of all other nodes in the knowledge structure except for its children nodes and, as argued earlier, that the child sibling nodes of the parent are equivalent to neighbouring nodes in a MRF. We therefore conclude that for the FG structure, when the relationships between the children nodes and the parent node of this structure are of type *IS-A* or *wrapper* (section 3.1), then the structure maps to an MRF structure. However if the relationship is of type *contribute-to* then the structure maps to a directed graph with the cause to effect arrow pointing from each child node to the parent node. This is because of the causal nature inherent in the *contribute-to* type. In this directed graph each child node will have an effect on the parent node but between the children nodes, no causal connection will exist. This also tallies with BBN d-separation.

3.4 The Generic Distinct with Component Structures

In the generic distinct component structure, the parent node is of type generic distinct which implies that even though its structure is always the same, both its RI and MG values may vary with context. The behaviour of the internal nodes within these structures (i.e. the children and descendant nodes) is dependent on whether or not the parent node of the structure has any ancestor(s) that is a generic node. Based on the presence or absence of a generic ancestor we split this structure type into two different types as follows.

3.4.1 The Generic Distinct with a Generic Ancestor Component Structure (GDGA)

In the GDGA structure, the root (parent) type is of type generic distinct which would normally imply that even though its structure is always the same, its internal nodes RI values vary with context. However because it has at least one ancestor node that is of type generic, the presence of the generic ancestor defines the behaviour of both the root node and the internal nodes within the structure. In relation to the nearest generic ancestor all the nodes of the GDGA structure are internal nodes and make up a FG structure with the nearest generic ancestor as the FG's root node. This in turn means that the nodes will have fixed RIs and behave in a similar fashion to the FG structure. The mapping process for this component structure is obtained in a similar manner to that of the FG structure discussed earlier, with causal relationships represented as directed graphs or causally linked MRFs as appropriate and symmetric relationships represented by MRFs. The only difference is that the mappings are carried out within the context of the nearest generic ancestor node.

3.4.2 The Generic Distinct with No Generic Ancestor Component Structure (GD)

In the GD structure, the root node is of type generic distinct and, as it does not have any generic fixed ancestor, it can have varying MG values whereas the internal nodes may have both varying RI and MG values. These component structures are highly dependent on context, as the root node is dependent on its neighbouring nodes (equivalent to the Markov blanket of a Bayesian network). As a result of these additional dependences this component structure cannot be modelled by a simple MRF and the next level comprising of the root node, its siblings and parents also need to be considered in the mapping process. The choice of graphical structure that it will be represented by is determined by the type of relationship between the variables.

Option 1: Where both level 1 (i.e. root to children node relationships) and level 2 (the root node, its sibling nodes and parent nodes) are of type *IS-A* or *wrapper* because of the additional level 2, a probabilistic graph that models a hierarchical form is required. The Tree Structured Conditional Random Field (i.e. TS-CRF) was chosen because it is a graph structure that allows us to incorporate into the probabilistic model the hierarchical form from the knowledge structure by mapping the structure into a tree like structure (see [12] for further details on TS-CRFs).

Option 2: For root to children node relationships of the type *IS-A* or *wrapper* and level 2 relationships of type *contribute-to* this maps to a causally linked MRF. This is because the relationship between the parent and children nodes is non causal and so this can be modelled as an MRF. However because the relationship between the parent node and its siblings are causal, this introduces the causal element into the graphical structure. In [14] the learning of a causally linked MRF structure is discussed.

Option 3: For root to children node relationships of type *contribute-to* and for level 2 relationships of type *contribute-to* this maps to a directed graph. Again this is because of the inherent causality that exists in this structure.

Option 4: Finally for root to children node relationships of type *contribute-to* and for level 2 relationships of type *IS-A* or *wrapper* the root to children nodes will map to a directed graph and the non causal root node and its sibling relations to their parents will map to a MRF.

Application of the appropriate options will result in an overall chain graph structure. Briefly, chain graphs model the relations between variables using both directed and undirected edges with the constraint that they do not have semi-directed cycles. An in-depth discussion on chain graphs can be seen in [15].

4 The Development of the GRiST Probability Structure

In this section we use the building block structures identified in the previous section to convert a subsection of the GSK to a graphical structure.

Summary of the steps for the development of the GRiST probability graph:

Step 1: Starting from level one (i.e. datum nodes level) identify for the parent concept on path 1, the relationship kinds between the parent concept and its children nodes. With respect to the GKS subsection depicted in figure 3, this parent concept is *hopelessness* and its children nodes *life not worth living* and *plans for the future*. The relationship between the children nodes and the parent node in this case is a *contribute-to* one.

Step 2: For the parent and children nodes identified in Step 1, use the rules outlined in sections 3.2 to 3.4, to determine the type of component structure they are. For the example in figure 3, as the parent node is of type *generic* and the children nodes non generic distinct, this structure is a *fixed generic component structure*.

Step 3: Next identify the probabilistic graphical structure that it maps to (using the rules outlined in section 3.3). For our example, as the relationship between the children nodes and parent node is of type *contribute-to*, the probabilistic graphical structure it maps to is a directed graph.

Step 4: Repeat the above for every parent-child node combination in the structure until all the various sections have been mapped into equivalent probabilistic graphical structures. In the example in figure 3, the next set of node relationships to consider will be the relationship of the parent concept *feelings and emotions* to its children nodes (including the *hopelessness* concept). Step 1 here results in an identification of relationship type *IS-A* between the children and parent nodes. Then using the rules from sections 3.2 to 3.4, as the parent node is generic distinct and it has a generic

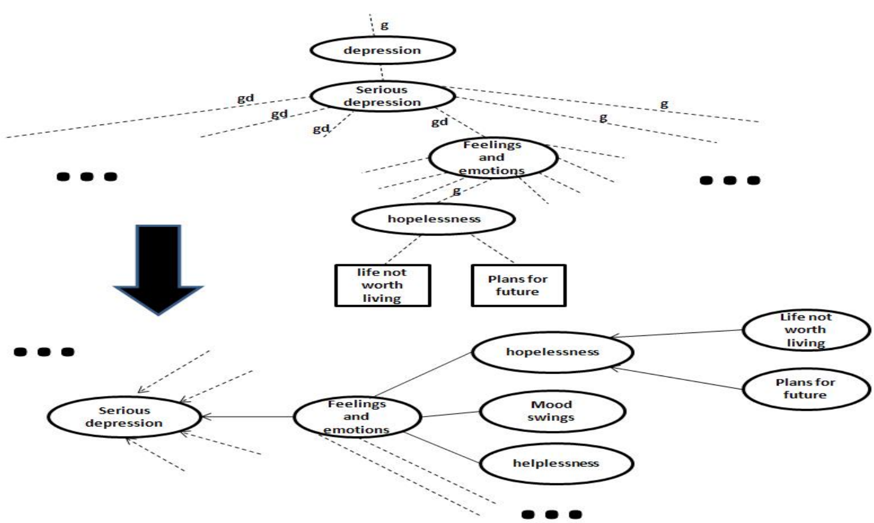


Fig. 3. Subsection of GRiST knowledge structure (top diagram) and its equivalent probabilistic graphical structure (bottom diagram)

ancestor *depression,* the component structure is identified as *generic distinct with a generic ancestor*. However to determine from the rules in section 3.4.1 which graphical structure to map to, we see that we need to go up a level and identify the relationship type between the parent node and its siblings and their parent concept node (i.e. *serious depression*). This is seen to be type *contribute-to* and finally from the rules, we see that this component maps to a causally linked MRF. Figure 3 depicts both the original GKS and the resulting probability graphical structure.

This resultant structure is, as expected, a mixed model consisting of both directed and undirected graphs and is in effect a chain graph. The importance of maintaining the integrity of the conditional independence relationships modelled by the structure was mentioned earlier. Informally, the conditional independence and dependence relations modelled in figure 3 can be read off the structure using the fact that two variables in the graphical structure are conditionally independent of each other given their prior and concurrent variables [16].

5 Conclusion and Future Directions

Trying to induce accurate probability graphs from a data set is notoriously difficult. The approach outlined in this paper shows that it can become more tractable by exploiting the structure induced from domain experts. We have shown how the formal specification of expert knowledge within the GRiST hierarchical structure can be used to direct the translation of the structure into appropriate graphs. The result is a chain graph that provides an alternative risk assessment based on probabilistic knowledge that can complement the structured clinical judgements modelled by GRiST. The methods discussed in this paper and the identified semantics and building block structures can be used for creating graphs for all the risks. Our ongoing research is into checking and validating the conversion process to ensure the semantics of GRiST expertise are appropriately represented by the probabilistic structures and that there are no clashes of constraints. We will then apply the chain graphs to real-world risk data and compare the probability estimations with the judgements from GRiST's psychological model as well as with outputs from alternative mathematical tools. Future work will focus on the defining and learning of parameters for the developed structure and how the methods may have applicability for other knowledge-based systems using hierarchical expertise.

References

1. Buckingham, C.D.: Psychological Cue Use and Implications for a Clinical Decision Support System. Medical Informatics and the Internet in Medicine 27(4), 237–251 (2002)
2. Department of Health: Safety First: Five-Year Report of National Confidential Inquiry into Suicide and Homicide by People with Mental Illness. Department of Health, London (2001)
3. Department of Health: Best Practice in Managing Risk. Department of Health, London (2007)
4. Shlonskya, A., Wagner, D.: The next step: Integrating Actuarial Risk Assessment and Clinical Judgment into an Evidence-based Practice Framework in CPS Case Management. Children and Youth Services Review 27(4), 409–427 (2005)

5. Buckingham, C.D., Adams, A.E., Mace, C.: Cues and knowledge structures used by mental-health professionals when making risk assessments. Journal of Mental Health 17(3), 299–314 (2008)
6. Buckingham, C.D., Ahmed, A., Adams, A.E.: Using XML and XSLT for flexible elicitation of mental-health risk knowledge. Medical Informatics and the Internet in Medicine 32(1), 65–81 (2007)
7. Jiang, C.A., Leong, T.Y., Poh, K.L.: PGMC: A Framework for Probabilistic Graphic Model Combination. In: Friedman, C.P., Ash, J., Tarczy-Hornoch, P. (eds.) Proceedings of the American Medical Informatics Association Annual Symposium (AMIA) 2005, pp. 370–374 (2005)
8. Kim, S.H.: Combining Decomposable Model Structures. Research Report 04-15, Division of Applied Mathematics, KAIST, Daejeon 305-701, S. Korea (2004)
9. Jensen, F.V., Nielsen, T.D.: Bayesian Networks and Decision Graphs, 2nd edn. Springer Publishing Company, Incorporated, Heidelberg (2007)
10. Bouman, C., Thompson, A.M., Brown, J.C., Kay, J.W.: Markov random fields and stochastic image models, Tutorial Presentation. In: IEEE International Conference on Image Processing, Washington D.C (1995)
11. Brachman, R.J.: What IS-A Is and Isn't: An Analysis of Taxonomic Links in Semantic Networks. Computer 16(10), 30–36 (1983)
12. Awasthi, P., Gagrani, A., Ravindran, B.: Image modeling using Tree Structured Conditional Random Fields. In: Sangal, R., Mehta, H., Bagga, R.K. (eds.) Proceedings of the 20th International Joint Conference on Artifical Intelligence, Hyderabad, India, January 06 - 12, pp. 2060–2065. Morgan Kaufmann Publishers, San Francisco (2007)
13. Gottard, A., Rampichini, C.: Chain Graphs for Multilevel Models. Statistics and Probability Letters 77(3), 312–318 (2007)
14. Hinton, G.E., Osindero, S., Bao, K.: Learning Causally Linked Markov Random Fields. In: Proceedings of the 10th International Workshop on Artificial Intelligence and Statistics. The Society for Artificial Intelligence and Statistics (2005)
15. Lauritzen, S.L., Wermuth, N.: Graphical Models for Associations between Variables, Some of which are Qualitative and Some Quantitative. Ann. Statist. 17, 31–57 (1989)
16. Blauth, A., Pigeot, I.: GraphFitI-A Computer Program for Graphical Chain Models. Techincal Report, Institute of Statistics, University of Munich, Germany (2000)

Controlling Security of Software Development with Multi-agent System

Esmiralda Moradian and Anne Håkansson

Department of Communication Systems, KTH,
The Royal Institute of Technology
Forum 100, 164 40 Kista, Stockholm, Sweden
moradian@kth.se, annehak@kth.se

Abstract. Software systems become distributed and complex. Distributed systems are crucial for organizations since they provide possibility to share data and information, resources and services. Nowadays, many software systems are not developed from scratch: system development involves reuse of already developed components. However, with the intrusion in the computer systems, it has become important that systems must fulfill security goals and requirements. Moreover, interdependencies of components create problems during integration phase. Therefore, security properties of components should be considered and evaluated earlier in the lifecycle. In this paper, we propose an agent-oriented process that supports verification of fulfillment of security goals and validation of security requirements during different phases of development lifecycle. Moreover, the system needs to support mapping of security requirements to threat list to determine if any of the attacks in the list is applicable to the system to be developed. This is performed by the meta-agents. These meta-agents automatically create a security checklist, as well as, provide control of actions taken by human agent.

Keywords: Multi-agent system, security engineering, risk management, security checklist, control system.

1 Introduction

Interdependencies between organizations in different sectors, such as business, financial, government, and medical are increasing. Hence, software systems become distributed and thereby more complex. However, distributed systems must be able to fulfill security goals and requirements and be able to operate securely. Unfortunately, security is often neglected during the development process. When considered, security is "relegated to the status of a few add-on fixes when all other design decisions have been frozen" [1]. Thus, security solutions, when implemented are isolated, and inadequate to business requirements. Lodderstedt *et al.* [2] point out three reasons for isolation, which are lack of understanding that security should be integrated in the development process, lack of tools supporting security engineering and lack of experience of the developers. Several security standards, among others

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 98–107, 2010.

Common Criteria (CC), and ISO 27000 series, make it possible to handle security requirements [3]. However, evaluation according to any of the standards is resource and time demanding [4]. Faults in software systems are unavoidable. To identify problems and find security risks that are not detected by human, automated processes are needed. We propose an agent-oriented process that can support security requirements engineering by performing automated search, provide checking for documents and intelligent matching of information. In the research in this paper, we present search, check and match processes that verify and validate security requirements against goals, risks and standards. Interdependencies of components create problems during integration phase. Hence, combined properties of components should be evaluated earlier in the lifecycle [1]. The work here explores the possibilities to automatically inspect combination of components during design phase and identify and analyze components that are necessary to achieve goals and fulfill requirements.

2 Related Work

Mouratidis *et al.* [5] present an approach that integrates security and systems engineering process. The work is done within the context of the Tropos methodology where security requirements are considered as an integral part of the whole development process. Authors [5] introduce security-oriented paradigm to the software engineering process. The process is iterative, which allows redefinition of security requirements in different levels. In our research, we present a semi-autonomous approach of controlling development of new systems or modification of existing systems. We focus on verification and validation of security requirements throughout the development process.

Brændeland and Stølen [6] propose an integrated process for component-based system development and security risk analysis. The authors [6] concentrate on identifying stakeholders and assets at the component level and identify and estimate the risks of component interaction. Moreover, they describe how the basic components can fit together into a composite component that refines the original requirements. In our work, we propose a multi-agent system to verify and validate security requirements and perform analysis of security properties of identified components. Moreover, the agent system inspects interdependencies of components. The output is a security checklist presented to the human agent.

Security checklist is not a new paradigm. Gilliam *et al.* [7] focuses their research on the development of a Software Security Checklist (SSC) for the life cycle. The authors [7] consider requirements engineering and specification, design and code issues, maintenance and decommissioning of software systems. Gilliam *et al.* [7] point at SSC as an instrument to guide organizations and system engineers in integrating security into the software life cycle. [7] Many checklist approaches were developed. However, listing pinpoints, without understanding and testing how they fit together during different phases of the system development, do not help in creating security architecture [1]. Thus, a multi-agent system is necessary in order to provide a holistic view throughout lifecycle. In our work, we propose the checklist that is automatically created by meta-agent as a result of performed analyses.

3 Agents in Multi-agent System

Nowadays, agent technologies are used in information and communication systems in order to provide management, search, monitoring, etc. In this paper, we present the multi-agent system that enforces control of security requirements and secure design. Moreover, the multi-agent system enables verification and validation of requirements throughout software development process. To handle this quality assurance, the multi-agent system consists of software agents and meta-agents where agents are working as a team.

The agents are communicative, mobile, cooperative, goal-oriented, autonomous, adaptive, and reactive [8]. Agents are communicative in the sense that they can communicate with users, other agents in the system, and other systems. Software agents communicate with the user in order to receive commands. From these commands, the agents search for information. The agents move between different locations over the networks while searching for available components and are therefore mobile.

Cooperation with the meta-level agents consists of passing significant information. Meta-agents use the information for analyses, control and matching. The meta-agents communicate with each other to united solve the problems [9]. They have ability to cooperate, i.e. work together with other agents to solve tasks. Goal oriented agents work towards a specific goal(s) [10]. The tasks have to be performed by the agents without any external guidance, and, thus, must be autonomous. Autonomous agents can operate without human intervention.

Software agents execute the user input. These agents are used to provide search for documents according to received information. Moreover, software agents can observe and monitor user actions through Interface module. To minimize search time in our system, software agents perform search in parallel in multiple databases.

Additionally, meta-agents are used as management, coordination, matching and checking agents. The meta-agents, in our system, can compare, analyse and combine information received from software agents. The meta-agents are also autonomous, i.e. meta-agents are able to act autonomously, reason and take decision in order to satisfy their objectives.

The meta-agents interact with software agents and other meta-agents to satisfy goals [9]. The communication with the software agents consists of collecting information from the software agents and giving commands that the software agents have to execute [11]. Meta-agents are reactive and adaptive, since they can learn, respond and adapt to changed environments.

The agents, in our multi-agent system, work in a deterministic environment, which means that next state of the environment is determined by the current state and the action that is being executed by an agent [11]. Environment is episodic, since software agents perform one task at the time while searching for information in databases. The task is complex and, therefore, divided into many simple tasks. Thus, we use multiple agents due to task complexity. To increase efficiency and shorten search time, the team of software agents execute in parallel. The environment, agents work in, is accessible, since the agents have access to information needed to accomplish tasks and satisfy goals.

Dynamic environment refers to the environment that can change. In our work, the agents need to monitor changes in the environment during execution because checklist is continuously updated by the human agent and requirements can change during the development process.

4 Security Requirements

Requirement describes the activities of the system to fulfil the purpose of the system and specify a system's behaviour [12]. Security is about protecting business goals and assets [1]. Analysis of business requirements enable to determine security goals of the system. Security policy expresses clearly and concisely what the protection mechanisms are to achieve and contains security goals agreed by the management [13]. Security goals help to identify security requirements, i.e. security constraints. Thus, security requirements must be identified and analysed according to business requirements and goals (business and security). The security policy specifies security properties of the system (i.e., identify secure and non-secure states, describe the secure way of information management, conditions of altering data, and identify level of trust), and exists to guide decisions in order to achieve desired goals [14]. However, the security policy does not state how security is to be delivered [1].

Building secure system requires that risk management is a part of development process [15]. The purpose of risk management is to produce knowledge about relevant security characteristics of the system(s). Risk management ensures that information security is managed appropriately during all phases in a system lifecycle. The first step is to assess criticalities of the system to be developed through identification of risks. Criticality is a measure of the extent to which a deficiency of the system in its environment may affect security [16]. Next step is to define or update the baseline of security requirements [17]. The ongoing risk analysis leads to the definition of security requirements and reduction or acceptance of remaining risk. Risk expresses probability (P) that undesirable event will occur and the consequence (C) of it. Security engineers have to understand that risks and threats should be put in content, and made realistic assessments, of what might go wrong [13]. Requirements engineering often suffers from problems such as: requirements are specified without any analysis or analysis restricted to functional end-user requirements; requirements specification is incomplete, inconsistent, and incapable of being validated [18]. Several standards that deal with security requirements exist. However, many researches [4, 19, 3, 20, 21] point out issues such as: lack of guidance on how to fulfill requirements during the development process [21], lack of methodological support [3], evaluation is time consuming [19, 21], and existence of gaps and overlaps [20]. Hence, analysis of security requirements and mapping those to security standards, as well as, control of implemented actions, that supports management, is needed. Therefore, a multi-agent system where meta-agents provide verification and validation of security requirements against goals, standards, risk analysis, as well as, control of performed actions, is highly needed.

5 Secure Architectural Design

This section presents secure architectural design of the enterprise system. When building distributed systems, architects need to considerate that resources and systems from different organisations must work together. This involves Web servers, databases, security systems, business logic components etc [22]. It is essential to describe components and analyse their properties, interconnection, and interdependencies, in order to, specify communication and control mechanisms, as well as, control behaviour of the system as a whole. Security architecture should encompass an enterprise-wide view, i.e., from goals and requirements to operation and maintenance.

Secure design embraces logical, physical and component security architecture [1]. The logical security architecture identifies the relationships and interdependencies between various elements of the system. Logical security architecture concerns with specifying logical security services (i.e., confidentiality and integrity protection); entities (i.e., users, administrators); domains; and security processing cycle (i.e., registration, login, session management) [1]. Physical security architecture describes physical devices that deliver logical services. It concerns with specifying data model and structures (i.e., tables, messages, signatures), rules (i.e., conditions, procedures, actions), security mechanisms (i.e., access control, encryption, virus scanning), security technology infrastructure, and time dependency (i.e., events, time intervals) [1].

Each component is an element of the whole system and its integration into the system affects security of the overall system, since components are selected from different vendors. Thus, component selection and component building is an important step, since it may be the case that some components are not available or not suitable for the specific system. During design phase should be determined what components to build or to buy [1]. Sherwood [1] points out that the components with compatible standardized interfaces should be used in order to achieve integration. Security analyses of the components, as well as, security evaluation of each identified component and all components together are necessary to achieve goals and fulfill requirements. As soon as, components are identified, which are necessary for goal fulfilment, the relationships between components are analysed. Thus, all security standards that are required, descriptions and specifications of all selected strategic technologies, products and tools along with guidance on how, why, where and when they should be used, are described during design phase. Moreover, roles, access rights, all processes, procedures and activities and how they relate to each other should also be specified.

Size and complexity of the systems are growing. Therefore, architectural risk management should be performed to refine identified assets and risks. Risk management is the ongoing process throughout development lifecycle. It involves threat and vulnerability analysis and identification, and analysis of design flaws. Vulnerability analysis is performed to understand the internal dependencies as well as the impact of external software dependencies [23]. Threats are mapped to the potential vulnerabilities and the security controls. Thus, architectural risk analysis identifies risks in the architecture, which improves the built-in security of a system by determining the security risks of the system. Misuse cases, i.e., negative scenarios or use cases are designed in order to determine behaviour of the malicious user.

We propose using multi-agent system, where the meta-agent analyse requirements and map requirements and attack patterns to each other in order to determine if any of attacks in the list are applicable to the system to be developed. Moreover, from knowledge repository the meta-agents select those attack patterns that are relevant to the system to be developed.

6 Multi-agent Control System

Many systems are built, operated, and maintained by different groups or organizations. Security plays an essential role in the development lifecycle. Security goals and priorities define security requirements of the overall system [17]. Policies, goals, risks, and requirements are analysed in order to identify conflicting, redundant and missing goals and detect insufficient, inconsistent, incorrect, unrealistic and untraceable requirements. Software system architecture should conform to the goals and requirements that are to be fulfilled, which include performance, security, and interoperability, integration, and supportability. Vulnerabilities embedded in the software and system components affect security of the system [17].

An agent based security requirements engineering consists of a multi-agent system with meta-agents and software agents. The multi-agent system is depicted in Figure 1.

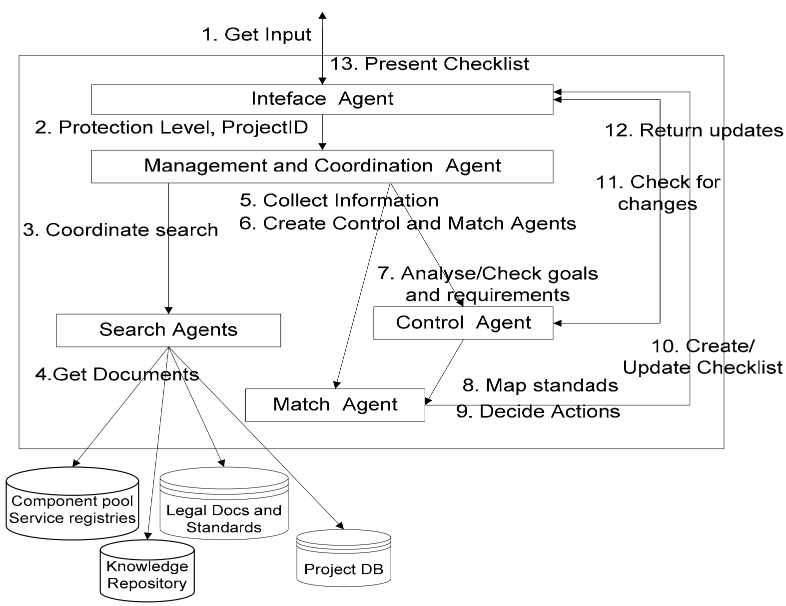


Fig. 1. Multi-agent system

The multi-agent system, in this research, is semi-autonomous. The agent system interacts with the external human agent in order to execute commands and capture changes. The process is depicted in Figure 1. The role of human agent is assigned to relevant actor throughout the lifecycle, which means that human agent can be a business and security manager, requirements engineer and risk manager as well as developer, architect, security specialist, and test specialist. [17] The interaction is going through the Interface Agent (IA). The Human agent (legitimate user) makes an input, by providing protection level and ProjectID. Documents that belong to the specific project should be tagged with its ProjectID. The IA gets this information (1:1) and passes it over to the Management and Coordination Agent (1:2). The Management and Coordination Agent is a meta-level agent that acts as coordination agent at this point. The meta-level agent is responsible for allocating search task to search agents, which are software agents and coordinating them. Depending on task complexity, agents can create new agents by cloning themselves. The replication allows performing tasks in parallel manner. The Management and Coordination agent assigns search tasks to search agents, i.e. gives a command to get documents according to input that are protection level and ProjectID (1:3, 1:4).

Search agents perform parallel search for documents in databases (1:4). For this pattern search is used. The software agents start to execute and a meta-agent is following along to the next node. Meta-agents provide control of software agents by keeping track on them. For the amount of databases to be searched the numbers of meta-agent and software agent are expanded. For the documents, a copy of the meta-agent and software agents are created and expanded with the amount of document. The procedure continues until agents reach every document that is relevant according to the input information. Notice, since there is a database with standards we need to consider the lifecycle phase, i.e. during requirements phase only standards that provide guidance for management security requirements will be valid. The meta-agents collect all information from search (1:5) and collaborate with each other and make decisions on how to proceed. Decision can concern standard(s) that should be used in the current phase, knowledge that is applicable for specific requirement, etc. The Management and Coordination Agent clones itself to Control Agent and Match Agent (1:6). The Control Agent is responsible for analyses and checks of documents (1:7), which contain important information that is necessary to perform the task. The documents such as business goals, system goals, requirements specifications (including security requirements) as well as risk analyses documents are analysed. Analyses include identification of conflicting, redundant and missing goals and insufficient or incorrect requirements. If such goals and requirements are identified, the system development is halted [17]. The agent system interacts with human agent in order to inform and receive new commands. The Match agent is responsible for performing mapping of standards, requirements, goals, security standards, domain expertise (1:8). Moreover, the Match Agent decides actions that should be implemented during development phases in order to satisfy goals (1:9) and create and update the checklist (1:10). The checklist is presented to the user by the Interface Agent (1:13). Goals and requirements can change during the development process. If that is the case then the checklist must be updated. The Control Agent needs iteratively perform checks of changes through Interface agent in order to detect deviations and capture new information (1:11). We assume that every single change is documented. If preconditions (goals and requirements) differ from initially defined conditions, the development of the system is halted until new search and analysis is performed (steps1:3 to 1:9). Moreover, the Control Agent is also responsible for control of the performed actions and validation of the requirements satisfaction. This meta-agent inspects the processes and procedures that have been applied. Check for changes (1:11, 1:12) are performed at each stage of the development process. All performed actions are documented in the checklist. The process is iterative and terminates when defined goals and requirements are satisfied [17].

Defective design accounts for 50% of security problems [15]. During the design phase software agents search (according to request from human agent) for available components and/or services (both internally within organisation and externally) in the component pool or in the service registry. See Figure 1:4. The meta-agent analyse component security properties and presents available components and analyses to the human agent. Software agents also perform search for documents such as, requirements, risks, list of threats and vulnerabilities. A Meta-agent compares and analyzes requirements and use cases and maps those to the attack patterns in order to determine which attacks can target the system. A Meta-agent collects and interprets the available information and maps the requirements to the threats in order to determine if any of attacks in the list are applicable to the system to be developed. The process is depicted in Figure 2.

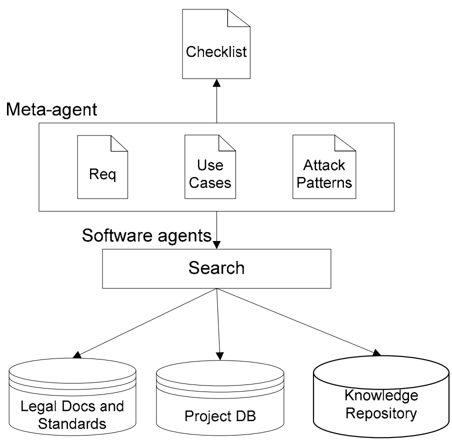


Fig. 2. The process of determining attacks

The meta-agent constructs a checklist with possible attacks that may target the system and communicate it to the human agent through the Interface agent, which is depicted in Figure 1. The checklist, created by the meta agent, can contain (but not limited to) following checkpoints: check authentication rules, implement random password generation, identify flaws, set up access control list, perform threat and attack analysis, design abuse cases, perform security testing for system integration, perform penetration tests, analyse network, implement and verify secure connections [17]. The Label "Yes" indicate fulfilled goals and requirements while "No" point to the unfulfilled goals and requirements. The checklist is extended and modified as soon as the change takes place. The resulting checklist is examined continuously both by the human agent and the meta-agent in order to detect changes, determine which of the security relevant activities took place, when (time, which phase of development lifecycle) and who (which human agent) is responsible for the specific activity.

7 Conclusion

In this work we presented a multi-agent system that can be used as a semi-automatic control system for verification and validation of security requirements throughout development lifecycle. Moreover, the proposed system can identify essential points in security requirements, such as conflicting, incomplete and inconsistent requirements. Furthermore, the system can compare and analyze requirements and use cases and map those to attack patterns in order to determine which attacks can target the system. The multi-agent system monitors documentation that includes requirements specifications, architecture overviews, component structure, interface specifications, analysis and results in order to keep track on changes in environment and have updated information. The proposed system can also control actions of the human agent during all the phases of system development.

References

1. Sherwood, J., Clark, A., Lynas, D.: Enterprise Security Architecture A Business-Driven Approach. CMP Books (2005), ISBN 1-57820318-X
2. Lodderstedt, T., Basin, D., Doser, J.: A UML-Based Modeling Language for Model-Driven Security*. In: Jézéquel, J.-M., Hussmann, H., Cook, S. (eds.) UML 2002. LNCS, vol. 2460, pp. 426–441. Springer, Heidelberg (2002)
3. Mellado, D., Fernández-Medina, E., Piattini, M.: A common criteria based security requirements engineering process for the development of secure information systems, vol. 29(2), pp. 244–253. Elsevier Science Publishers B. V, Amsterdam (February 2007), ISSN:0920-5489
4. Abbas, H., Yngström, L., Hemani, A.: Option Based Evaluation: Security Evaluation of IT Products Based on Options Theory. In: First IEEE Eastern European Conference on the Engineering of Computer Based Systems, pp. 134–141 (2009)
5. Mouratidis, H., Giorgini, P., Manson, G.: Integrating Security and System Engineering: Towards the Modelling of Secure Information Systems. In: Eder, J., Missikoff, M. (eds.) CAiSE 2003. LNCS, vol. 2681, pp. 63–78. Springer, Heidelberg (2003)
6. Brændeland, G., Stølen, K.: Using Model-based Security Analysis in Component-oriented System Development QoP 2006, Alexandria, Virginia, USA. Copyright 2006, ACM 1-59593-553-3/06/0010 (October 30, 2006)
7. Gilliam, D.P., Wolf, T.L., Sherif, J.S., Bishop, M.: Software Security Checklist for the Software Life Cycle. In: Proceedings of the Twelfth IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE 2003), pp. 243–248 (June 2003), ISBN: 0-7695-1963-6
8. Phillips-Wren, G.: Assisting Human Decision Making with Intelligent Technologies. In: Lovrek, I., Howlett, R.J., Jain, L.C. (eds.) KES 2008, Part I. LNCS (LNAI), vol. 5177, pp. 1–10. Springer, Heidelberg (2008)
9. Håkansson, A., Hartung, R.: Calculating optimal decision using Meta-level agents for Multi-Agents in Networks. In: Apolloni, B., Howlett, R.J., Jain, L. (eds.) KES 2007, Part I. LNCS (LNAI), vol. 4692, pp. 180–188. Springer, Heidelberg (2007)
10. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice-Hall, Englewood Cliffs (1995), ISBN: 0-13-103805-2

11. Moradian, E., Håkansson, A.: Approach to Solving Security Problems Using Meta-Agents in Multi Agent System. In: Nguyen, N.T., Jo, G.-S., Howlett, R.J., Jain, L.C. (eds.) KES-AMSTA 2008. LNCS (LNAI), vol. 4953, pp. 122–131. Springer, Heidelberg (2008)
12. Pfleeger, S.L.: Software Engineering Theory an Practice, 2nd edn. Prentice-Hall, Inc., Englewood Cliffs (2001), ISBN 0-13-029049-1
13. Andersson, R.: Security Engineering A guide to building. Dependable Distributed Systems, 2nd edn. Wiley, Chichester (2008), ISBN 978-0-470-06852-6
14. Bishop, M.: Introduction to Computer Security. Pearson Education, Inc., London (2005), ISBN 0-321-24744-2
15. McGraw, G.: Software Security Building Security. Addison-Wesley Pearson Ed., Reading (2006), ISBN 0-321-35670-5
16. Lindström, C., Násström, S.: Handbook for Software in Safety-Critical Applications. Swedish Armed Forces (2005)
17. Moradian, E., Håkansson, A., Andersson, J.-O.: Multi-Agent System Supporting Security Requirements Engineering. Accepted in The 9th International Conference of Software Engineering Research and Practice, SERP 2010 (2010)
18. Allen, J.H., Barnum, S., Ellisson, R.J., McGraw, G., Mead, N.: Software Security Engineering A Guide for Project Managers. Addison-Wesley, Reading (2008), ISBN: 0-321-50917X
19. Haley, C.B., Moffett, J.D., Laney, R., Nuseibeh, B.: A Framework to security requirements engineering. In: SESS 2006, Shanghai, China. Copyright, May 20-21, ACM 1-59593- 085-X/06/0005 (2006)
20. Magnusson, C.: Corporate Governance, Internal Control and Compliance (September 2007), `http://www.svensktnaringsliv.se/material/rapporter/article35898.ece`
21. Vetterling and Wimmel Secure Systems Development Based on the Common Criteria: The PalME Project SIGSOFT 2002/FSE10, Charleston, SC, USA, November 18-22. ACM 1581135149/02/0011 (2002)
22. Papazoglou, M.: Web Services: Principles and Technology. Pearson Education, Essex (2008), ISBN: 978-0-321-15555-0
23. Swiderski, F., Snyder, W.: Threat Modelling. Microsoft Press(2004), ISBN 0-7356-1991-3

On Łukasiewicz' Infinie-Valued Logic and Fuzzy$_{\mathrm{L}}$

Jorma K. Mattila

Lappeenranta University of Technology
Department of Mathematics and Physics,
P.O. Box 20, 53851 Lappeenranta, Finland
jorma.mattila@lut.fi

Abstract. An algebra of Łukasiewicz' logic $\text{Ł}_{\aleph}$ is considered without the extra operations $P \odot Q = \max\{0, P+Q-1\}$, the *restricted difference*, and $P \oplus Q = \min\{1, P+Q\}$, the *restricted sum*, later named by "Łukasiewicz t-norm" and "Łukasiewicz t-conorm", respectively. The logic $\text{Ł}_{\aleph}$ is compared with the fuzzy logic Fuzzy$_{\mathrm{L}}$ with the set of truth-values $[0,1]$. It appears that Fuzzy$_{\mathrm{L}}$ is a Łukasiewicz-like many-valued logic with a dual pair of modifiers (or fuzzifiers). Otherwise, the only reasonable version of Fuzzy$_{\mathrm{L}}$ is $\text{Ł}_{\aleph}$, without the additional connectives corresponding to the "Łukasiewicz norms" which are not original operations of Łukasiewicz' logics.

Keywords: Lukasiewicz Many-valued Logic, Fuzzy Logic.

1 Introduction

The author would like to respect Łukasiewicz's work in many-valued logic by giving a comment on the straight connection between Łukasiewicz implication and original disjunction. The comment says that there is a natural and easy formal connection between Łukasiewicz disjunction max and implication. In this paper, it is done by using an algebraic approach being a counter currency case, rather than the well known mainstream proceeding mainly developed for studying formal fuzzy logic on the base of residuated lattices, BL-algebras, MV-algebras, etc. (*cf.* a comprehensive presentation about these tools in Hájek's book [2]). In the mainstream proceeding, Łukasiewicz implication is developed by the t-conorm $\oplus$ to be an S-implication. Hence, this is a circuitous route from a *different disjunction* operation to Łukasiewicz implication. The circuitous route is completed by the case that the *different disjunction* is $\oplus$, and min operation can be expressed by means of $\oplus$ and $\odot$ (*cf.* Bergmann's deduction (4) below), and the dual of min is max. Hence, max operation can be expressed by $\oplus$ and $\odot$ by means of duality of min and max (*cf.* (6) below).

Some confusion or miss-understanding about the role of *restricted difference* $\odot$ and *restricted sum* $\oplus$ have appeared in some papers. The difference between the connectives disjunction and conjunction of Łukasiewicz many-valued logic and these additional connectives has not been very clear. This is possible if we do not draw a strict border between the connectives of original Łukasiewicz logic

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 108–115, 2010.

and the additional connectives. In late 80's the author has seen some few papers considering fuzzy control where introducing Łukasiewicz many-valued logic, authors have asserted that $\odot$ and $\oplus$ are the original Łukasiewicz conjunction and disjunction, respectively. Maybe, the reason for this is that the norms $\odot$ and $\oplus$ were renamed as Łukasiewicz's t-norm and Łukasiewicz's t-conorm, respectively. Actually, the right thing would be that min and max operations are the original Łukasiewicz' t-norm and t-conorm. In Łukasiewicz' logic, the operations min and max already have had a similar role before as in fuzzy logic nowadays.

Anyway, the system $([0,1], \odot, \oplus, ', 0, 1)$ is a many-valued logic with the negation $p' \equiv 1 - p$ for all $p \in [0,1]$, and where the implication operation can be created to be an S-implication of the system, but this system is not a lattice if we do not create supremum and infimum operations. We consider this more in Conclusion.

1.1 A Starting Point to Łukasiewicz' Many-Valued Logic

We begin with Łukasiewicz' many-valued logic $Ł_{\aleph_1}$ having the closed unit interval $[0,1]$ as the set of truth values.[1]

As we know, Łukasiewicz chose the connectives of *negation* and *implication* as primitives. Let v be any valuation of $Ł_{\aleph}$, then the truth value evaluation rules for negation and implication are

$$v(\neg A) = 1 - v(A) \tag{Neg.}$$

$$v(A \to B) = \min\{1, 1 - v(A) + v(B)\} \tag{Impl.}$$

By means of these connectives, Łukasiewicz defined the other connectives by the rules

$$A \vee B \overset{\text{def}}{\Longleftrightarrow} (A \to B) \to B \tag{Disj.}$$

$$A \wedge B \overset{\text{def}}{\Longleftrightarrow} \neg(\neg A \vee \neg B) \tag{Conj.}$$

$$A \leftrightarrow B \overset{\text{def}}{\Longleftrightarrow} (A \to B) \wedge (B \to A) \tag{Eq.}$$

The truth value evaluation rules for these derived connectives are

$$\max\{v(A), v(B)\} \qquad \text{for } A \vee B, \tag{1}$$

$$\min\{v(A), v(B)\} \qquad \text{for } A \wedge B, \tag{2}$$

$$1 - |v(A) - v(B)| \qquad \text{for } A \leftrightarrow B \tag{3}$$

for any valuation v of $Ł_{\aleph_1}$.

1.2 A Starting Point of Fuzzy Many-Valued Logic $\mathrm{Fuzzy}_{\mathbf{L}}$

The truth evaluation rules and the idea for $\mathrm{Fuzzy}_{\mathrm{L}}$ is taken from M. Bergmann [1]. She says:

[1] *Cf.* Rescher [9], p.36, and 337.

> "To specify a full fuzzy propositional logic, we begin with an assignment v of fuzzy truth values, between 0 and 1 inclusive, to the atomic formulas of the language. We call the set of real numbers between 0 and 1 inclusive the unit interval lattice $(\mathbb{I}, \leq)$. So we may say that for each atomic formula P, $v(P)$ is a member of $\mathbb{I}$, or in more concise notation, $v(P) \in [0, 1]$. We can then use the same numeric clauses that we presented for Łukasiewicz 3-valued logic to obtain the Łukasiewicz fuzzy system Fuzzy$_L$:
>
> 1. $v(\neg P) = 1 - v(P)$
> 2. $v(P \wedge Q) = \min\{v(P), v(Q)\}$
> 3. $v(P \vee Q) = \max\{v(P), v(Q)\}$
> 4. $v(P \rightarrow Q) = \min\{1, 1 - v(P) + v(Q)\}$
> 5. $v(P \leftrightarrow Q) = \min\{1, 1 - v(P) + v(Q), 1 - v(Q) + v(P)\} = 1 - |v(P) - v(Q)|$
> 6. $v(P \vartriangle Q) = \max\{0, v(P) + v(Q) - 1\}$
> 7. $v(P \triangledown Q) = \min\{1, v(P) + v(Q)\}$
>
> Some of these connectives can be defined using the other connectives, as we did in Ł$_3$. For example, ... we noted that $P \wedge Q$ is definable as $P \vartriangle (\neg P \triangledown Q)$ in Ł$_3$, and
>
> $$\begin{aligned} v(P \vartriangle (\neg P \triangledown Q)) &= \max\{0, v(P) + \min(1, 1 - v(P) + v(Q)) - 1\} \\ &= \max\{0, \min(v(P) + 1 - 1, v(P) + 1 - v(P) + \\ &\quad v(Q) - 1)\} \\ &= \max\{0, \min(v(P), v(Q)\} = \min\{v(P), v(Q)\}." \end{aligned} \quad (4)$$

By means of duality applied to Bergmann's result,

$$\min\{v(P), v(Q)\} = v(P \vartriangle (\neg P \triangledown Q)), \quad (5)$$

max operation can be expressed by means of $\vartriangle$ and $\triangledown$, too, i.e.,

$$\max\{v(P), v(Q)\} = 1 - [(1 - v(P)) \vartriangle (v(P) \triangledown (1 - v(Q)))]. \quad (6)$$

This can be verified by the definitions of $\vartriangle$ and $\triangledown$ given above.

Usually, the symbolization of $\vartriangle$ and $\triangledown$ used in the literature in the truth evaluation rules 6 and 7 are $\odot$ and $\oplus$, respectively. We now return to use this notation again.

If we have a fuzzy many-valued propositional logic, say, Fuzzy$_\mathrm{L}$, where the truth evaluation rules are those from 1 to 7, how we must interpret the connectives $\vartriangle$ and $\triangledown$, or correspondingly, $\odot$ and $\oplus$, *bounded difference* and *bounded sum*, especially when we consider inferences (or arguments) in natural language? A t-norm refers to conjunction and a t-conorm to disjunction. But we have already conjunction and disjunction in Fuzzy$_\mathrm{L}$, namely min and max. And, in general, we can manage with the connectives whose evaluation rules are 1 - 5 in the same way as in Ł$_{\aleph_1}$. In practice, the formula (4) is not so big case that it would necessarily take sides with the use of $\odot$ and $\oplus$ in the set of connectives in such logics like Fuzzy$_{\mathrm{L}}$.

It seems that the additional connective $\oplus$ cannot be interpreted as *exclusive or*, i.e., as $P \underline{\vee} Q$ which can be expressed by $\vee$, $\wedge$, and $\neg$ as $(P \wedge \neg Q) \vee (\neg P \wedge Q)$. 'Exclusive or' is more crisp than 'inclusive or'. On the other hand, $\oplus$ is more "weakening" in its interpretation because, as t-conorms, the value $v(P \oplus Q)$ is equal to or bigger than that of $v(P \vee Q)$ for any valuation v in Fuzzy$_{\mathrm{L}}$. This means that $\oplus$ can be used as a basis for constructing a certain *weakening* modifier. Similarly, as t-conorms, $v(P \odot Q) \leq v(P \wedge Q)$. Hence, $\odot$ can be used as a basis for constructing a certain *substantiating* modifier (*cf.*, for example, Mattila [5]). Hence, the interpretations in natural language for these connectives may be something like *definitely* P *and* Q for $P \odot Q$ and *something like* P *or* Q for $P \oplus Q$, because *definitely* can be understood as substantiating and *something like* a weakening expression. Of course, there may be alternative ways to translate $P \odot Q$ and $P \oplus Q$ into natural language, but the idea is that the translations are substantiating and weakening, respectively.

These kind of modifying expressions can be used as *fuzzifiers* (*cf.* Mattila [3], [4], [7]), and hence Fuzzy$_{\mathrm{L}}$ can be considered as a many-valued fuzzy logic with one dual pair of fuzzifiers.

2 An Alternative Algebraic Approach to Ł$_{\aleph}$

The attribute *alternative* means here that the approach we consider is alternative to the mainstream approach mentioned above. And here we do not need any additional tools like different t-conorms. A set of essential algebras involved in many-valued logics are considered e.g. in Rasiowa's book [8].

If we want to consider logics based on Łukasiewicz logic without residuated lattices, BL- and MV-algebras, etc., we may proceed as follows.

Consider an algebra

$$L_{\mathbb{I}} = ([0,1], \wedge, \vee, ', 0, 1) \tag{7}$$

Suppose that the algebra $L_{\mathbb{I}}$ has at least the following properties:

- $L_{\mathbb{I}}$ is a DeMorgan algebra.
- The binary operations *meet* $\wedge$ and *join* $\vee$ are commutative and associative.
- The operations $\wedge$ and $\vee$ are distributive to each others.
- The unary operation $'$ is a complementarity operation with the property of involution.

We may need a short analysis of the universe of discourse $[0,1]$ of the algebra $L_{\mathbb{I}}$. Because $[0,1]$ is a subset of the set of real numbers $\mathbb{R}$, $[0,1]$ has all the arithmetical properties as $\mathbb{R}$ has. Hence, $[0,1]$ is a metric space with the natural metric *distance* between any two points of $[0,1]$, i.e., the distance is

$$d(x,y) = |x - y| , \quad x, y \in [0,1]. \tag{8}$$

This formula (8) satisfies the general definition of the concept *metric*. We need it in the following consideration where we manipulate expressions involving maxima and minima.

In manipulating maxima and minima, the consideration can sometimes be done easier by using the following expressions for max and min operations:

$$\max\{x, y\} = \frac{x + y + |x - y|}{2}, \qquad \min\{x, y\} = \frac{x + y - |x - y|}{2} \tag{9}$$

First, consider the case where the operations of the algebra are min, max, and $1 - \cdot$.

Proposition 1. *If the operations are chosen such that for any* $x, y \in [0, 1]$, $x \wedge y = \min\{x, y\}$, $x \vee y = \max\{x, y\}$ *and* $x' = 1 - x$*, then the resulting algebra*

$$([0, 1], \min, \max, 1 - \cdot, 0, 1) \tag{10}$$

satisfies the conditions of the algebra (7).

Proof. Consider the operation $x \wedge y = \min(x, y)$ and $x \vee y = \max(x, y)$, where x and y are variables taking their values from the interval $[0, 1]$. Using the arithmetical formula for min operation, we have

$$\begin{aligned}
\min\{x, y\} &= \frac{x + y - \mid x - y \mid}{2} \\
&= \frac{2 - 1 - 1 + x + y - \mid 1 - 1 + x - y \mid}{2} \\
&= 1 - \frac{1 + 1 - x - y + \mid 1 - 1 + x - y \mid}{2} \\
&= 1 - \frac{(1 - x) + (1 - y) + \mid (1 - y) - (1 - x) \mid}{2} \\
&= 1 - \max\{\, (1 - x), (1 - y) \,\}.
\end{aligned} \tag{11}$$

From the formula (11), by replacing x by $1 - x$ and y by $1 - y$, and then solving $\max(x, y)$, the following formula follows:

$$\max\{x, y\} = 1 - \min\{\, 1 - x, 1 - y\}. \tag{12}$$

The formulas (11) and (12) show that max and min are dual of each other which is already a well known fact. Hence, DeMorgan laws hold in the algebra (10). Hence, (10) is a DeMorgan algebra.

It is well-known that the operations max and min are commutative and associative, and they are distributive to each others, and $x' = 1 - x$ for all $x \in [0, 1]$ is a complementarity operation with the property of involution. This completes the proof.

Łukasiewicz knew that the operations max and min are dual of each other. Actually, this property is easily found in the classical special case, i.e. using characteristic functions in presenting crisp sets. But the general proof for this is easily done by using the expressions (9) for max and min in such cases where in the universe of discourse a distance mertic is defined. This always holds at least for real numbers.

Second, consider the connection between max and Łukasiewicz implication (similar considerations are done in Mattila [6], but the following proposition 2 is not proved).

Proposition 2. *For all* $x, y \in [0, 1]$,

$$\max(x, y) = (x \underset{L}{\rightarrow} y) \underset{L}{\rightarrow} y, \tag{13}$$

where $x \underset{L}{\rightarrow} y$ *is Łukasiewicz implication.*

Proof. Consider disjunction operation $x \vee y = \max(x, y)$. Because $0 \leq x, y \leq 1$, using the arithmetical formula (9) for max, we have

$$\begin{aligned}
\max(x, y) &= \min\{\, 1, \max(x, y)\,\} \\
&= \min\{\, 1, \frac{x + y + \mid x - y \mid}{2}\,\} \\
&= \min\left\{\, 1, \frac{2 - 1 - 1 + x + 2y - y + \mid 1 - 1 + x - y \mid}{2}\,\right\} \\
&= \min\left\{\, 1, 1 - \frac{1 + (1 - x + y) - \mid 1 - (1 - x + y) \mid}{2} + y\,\right\} \\
&= \min\{\, 1, 1 - \min(1, 1 - x + y) + y\,\}
\end{aligned}$$

By (Disj.) we know that

$$\max(x, y) = (x \underset{L}{\rightarrow} y) \underset{L}{\rightarrow} y \tag{14}$$

We find two similar min-structures in the formula

$$\min\{\, 1, 1 - \min(1, 1 - x + y) + y\,\}$$

where one of the min-stuctures is inside part of the whole formula. If we denote the inner min-structure $\min(1, 1 - x + y)$ by z then the outer min-structure is $\min(1, 1 - z + y)$, i.e., the min-structures are really the same. The implication operations in (14) are situated in the same way. Hence, $\min(1, 1 - x + y)$ must be $x \underset{L}{\rightarrow} y$, by (Disj.). And the expression $\min(1, 1-x+y)$ is similar to the truth value evaluation rule of Łukasiewicz implication (Impl.).

Of course, Łukasiewicz must have known the connection between maximum operation and his truth evaluation formula (Impl.) of the implication because without any knowledge about this, he would have not been sure that everything fits well together in his logic. But how he has inferred this is not known.

The result of the proof of the formula (13) shows that from the join operation max of our algebra we deduce a formula that expresses the rule of Łukasiewicz' implication, and this formula is the truth value evaluation rule in $Ł_{\aleph_1}$. Hence, we have shown that from our algebra (10) it is possible to derive similar rules as the truth value evaluation rules in $Ł_{\aleph_1}$.

Hence, the author's comment can be given in a more formal way: If the cases

1. $x' = 1 - x$;
2. $x \vee y = \max(x, y)$;

hold, then the other cases

3. $x \wedge y = \min(x, y)$;
4. $x \rightarrow y = \min(1, 1 - x + y)$;
5. $x \leftrightarrow y = (x \rightarrow y) \wedge (y \rightarrow x) = 1- \mid x - y \mid$,

can be derived in the DeMorgan algebra $L_{\mathbb{I}} = ([0,1], \wedge, \vee, ', 0, 1)$.

The case 3 follows from the case 2 by duality and the case 4 follows from the case 2 by Prop. 2. The case 5 is deduced as follows:

$$\begin{aligned}
x \leftrightarrow y &= \min\{(x \rightarrow y), (y \rightarrow x)\} \\
&= \frac{(x \rightarrow y) + (y \rightarrow x) - |(x \rightarrow y) - (y \rightarrow x)|}{2} \\
&= \frac{\min(1, 1 - x + y) + \min(1, 1 - y + x)}{2} \\
&\quad - \frac{\min(1, 1 - x + y) - \min(1, 1 - y + x)}{2} \\
&= \frac{4 - 2\,|x - y| - |-2x + 2y - |x - y| + |x - y||}{4} \\
&= \frac{4 - 4\,|x - y|}{4} = 1 - |x - y|\,.
\end{aligned}$$

These cases are similar to the truth value evaluation rules for connected formulas in $Ł_{\aleph_1}$. Hence, if we want to use algebraic approach for $Ł_{\aleph_1}$ we need not necessarily to follow the mainstream using additional operations for studying the connections between the connectives in $Ł_{\aleph_1}$.

3 Conclusion

We return back to the logic Fuzzy_{L}. The primitive connectives are min as conjunction, max as disjunction, and $1 - \cdot$ as negation. If we want to avoid interpretational confusions about the role of the additional connectives $\odot$ and $\oplus$ we may reject them. Hence, the logics Fuzzy_{L} and $Ł_{\aleph_1}$ would be identical. Hence, one may have a feeling that the dual operation pair $\odot$ and $\oplus$ are in the system only for getting Łukasiewicz implication very easily to the system as an S-implication, by means of $\oplus$.

On the other hand, consider the system $([0,1], \odot, \oplus, ', 0, 1)$ where $\odot$ and $\oplus$ are the primitive binary connectives, even though they are not lattice operations on $[0,1]$. Hence, we can apply the formulas (4) or (6) in order to construct the join operation $\wedge$ and meet operation $\vee$, which are min and max, respectively. Now we have the case where $\odot$ and $\oplus$ are conjunction and disjunction, respectively. Now we have an interpretation problem with min and max. They are only additional operations in order to make the system to be a lattice $([0,1], \odot, \oplus, \min, \max,', 0, 1)$ where the infimum operation $\wedge$ is min and the supremum operation $\vee$ is max. In any case, the implication operation can be constructed by means of the *primitive* connective $\oplus$ as an S-implication.

In both cases, it immediately seems that we have the same logic

$$([0,1], \odot, \oplus, \min, \max,', 0, 1) \quad \text{or} \quad ([0,1], \min, \max, \odot, \oplus,', 0, 1), \tag{15}$$

but the binary primitive connectives are different, and further, they have the same interpretation in natural language. If we want to have a many-valued fuzzy logic with the set of truth values $[0,1]$ being as close to $\text{Ł}_{\aleph_1}$ as possible, we may choose this Łukasiewicz logic $\text{Ł}_{\aleph_1}$ instead of $\mathrm{Fuzzy}_{\mathrm{L}}$. Hence, we have no interpretation problems.

Another alternative case would be such that in the algebraic approach, the operations max and min belog to the primitives (because they have no modifying effects), and $\oplus$ and $\odot$ are some dual modifier operations, because the t-norm $\odot$ has some substantiating properties and t-conorm $\oplus$ some weakening properties. So, we would have a *Łukasiewicz-like many-valued modifier logic*. On t-norm-based modifiers, see Mattila [5] and [6].

In general, a totally different problem is whether $\text{Ł}_{\aleph_1}$ can be considered to be a fuzzy logic or not. Or, is some fuzzy-valued logic better?

References

1. Bergmann, M.: An Introduction to Many-Valued and Fuzzy Logic. Cambridge University Press, Cambridge (2008)
2. Hájek, P.: Metamathematics of Fuzzy Logic. Kluwer Academic Publishers, Dordrecht (1998)
3. Mattila, J.K.: Modelling fuzziness by Kripke Structures. In: Terano, T., Sugeno, M., Mukaidono, M., Shigemasu, K. (eds.) Fuzzy Emgineering toward Human Friendly Systems, Yokohama, Japan, vol. 2, pp. 889–900 (1991)
4. Mattila, J.K.: Modifier Logics Based on Graded Modalities. Journal of Advanced Computational Intelligence and Intelligent Informatics 7, 71–78 (2003)
5. Mattila, J.K.: Modifiers based on some t-norms in fuzzy logic. In: Soft Computing, vol. 8, pp. 663–667. Springer, Heidelberg (2004), doi:10.1007/S00500-003-0323-X (Published online: 2003)
6. Mattila, J.K.: On Łukasiewicz Modifier Logic. Journal of Advanced Computational Intelligence and Intelligent Informatics 9, 506–510 (2005)
7. Mattila, J.K.: Many-valuation, modality, and fuzziness. In: Seising, R. (ed.) Views on Fuzzy Sets and Systems from Different Perspective. Philosophy, Logic, Criticisms and Applications, Berlin. Studies in Fuzziness and Soft Computing, pp. 271–300 (2009)
8. Rasiowa, H.: An Algebraic Approach to Non-Classical Logics. North-Holland, Amsterdam (1974)
9. Rescher, N.: Many-valued Logics. McGraw-Hill, New York (1969)

A Meta-level Approach to Approximate Probability

Vesa A. Niskanen

University of Helsinki, Dept. of Economics and Management, P.O. Box 27,
00014 Helsinki, Finland
vesa.a.niskanen@helsinki

Abstract. Probability is examined from the standpoint of the philosophy of science. First, various interpretations on probability are considered. Second, Lotfi Zadeh's novel extended logic for approximate reasoning is presented. Third, by using this logic, application examples on approximate probability are provided.

Keywords: Approximate reasoning, approximate probability.

1 Introduction

Lotfi Zadeh has recently formalized his studies on approximate reasoning by constructing the extended fuzzy logic, *FLe* [15]. The author has already applied his novel ideas to theory formation, scientific explanation and hypothesis verification [4] [5] [6] [7] [8]. This paper, in turn, examines approximate probability by applying the *FLe*.

We consider approximate probability from the standpoint of the philosophy of science, and thus we only provide some meta-level results. Section 2 performs a brief concept analysis on probability. Section 3 presents Zadeh's logic *FLe*. Section 4 considers the role of approximate probability in the conduct of inquiry. Section 5 concludes our examination.

2 Interpretations on Probability

The conceptions on probability stem from the problems of uncertainty, and uncertainty, in turn, is an issue of epistemology. The formal theories on probability have their origin in the mathematical models for gambling in the Renaissance period (Pascal, Fermat, Bernoulli, Bayes, Laplace etc.), and they mainly dealt with the ratio of successful events to all possible events. Later Kolmogorov, in particular, provided a basis for modern mathematical calculus of probability theory, and this approach, which supports the metrization of probability, is adopted in the mainstream studies. However, there is also another, more "philosophical" approach, and in this context we examine probability from the epistemological standpoint by using linguistic expressions, whereas the mathematical approaches regard probability as being an abstract primitive term with no particular meaning [2].

The mathematical tradition usually presupposes that probability is independent of the human knowledge and thus only dependent upon physical properties assigned to

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 116–123, 2010.

the occurrences. This so-called physicalistic or objective approach is closely related to modality because the role of possible events have been considered in this context [2].

The physicalistic tradition comprises frequency and propensity interpretations. The former stems from Venn's, von Mises's and Reichenbach's ideas that probability is the limit value of event's relative frequency in a large number of trials, and this interpretation is widely used in the modern probability theory and statistics.

Since we are unable to apply the frequency approach to unique or non-recurrent events as well as to the probabilities of hypotheses, theories or inductive reasoning, another type of probability, the propensity approach (Popper, Hacking, Gieren), is also suggested within physicalistic theories. The frequency probability is presupposed to be attributed to events, whereas in the propensity theories this physical attribute is related to the "chance-set-up" as such. In other words, probability means such numerical "disposition" which yields certain relative frequencies or measures for unique events [2] [9].

On the other hand, the epistemic approaches, such as the logical and subjective interpretations, presuppose that probability is dependent upon our knowledge and ignorance. In the logical case we assume that probability is the relationship between a hypothesis and its evidence (Keynes, Carnap), and thus it is related to inductive reasoning. Logical probabilities are thus rational degrees of belief of hypotheses according to the evidence [1] [10].

The subjective theories also apply the idea on the degree of belief but now these degrees are mainly based on human rational assessments (de Finetti, Suppes). Thus, these degrees may vary among the evaluators but they are assumed to converge upon each other the more evidence is acquired. On some occasions subjective theories are combined with the utility theory (von Neumann) and Bayesian decision theories (Savage).

The epistemic approaches allow us to consider the probabilities of the hypotheses and theories. Then one method is to apply truth as an epistemic utility, i.e., the hypothesis with the highest truth value should always be accepted (in which case, however, tautologies would be best candidates).

Another approach, which combines truth and (epistemic) probability stems from the ambiguity of the term "probability" in the Ancient Greek and Latin. In that period of time it referred to both uncertainty and truth, and thus we still have such semantic problem in certain languages that the word for probability actually refers to truth (e.g. in German the word for probability is Wahrscheinlichkeit, i.e., "truth-looking"). The distinction, however, should be clear because epistemic probability and truth are issues of epistemology and semantics, respectively [3].

The conceptions on imprecision or fuzziness and uncertainty are also sometimes confused for similar reasons although the former expression usually belongs to semantics and the latter to epistemology. Possibility, in turn, as used within fuzzy systems, is related both to various approaches to probability and to modality.

Hence, there are various conceptions on probability and this state of affairs has aroused several confusions when this subject matter has been examined.

3 Zadeh's Extended Logic and Approximate Probability

Recently, Zadeh has established the principles of the extended fuzzy logic, *FLe*, which is a combination of "traditional" provable and "precisiated" fuzzy logic, *FLp*, as well as a novel meta-level "unprecisiated" fuzzy logic, *Flu* [15]. He states that in the *FLp* the objects of discourse and analysis can be imprecise, uncertain, unreliable, incomplete or partially true, whereas the results of reasoning, deduction and computation are expected to be provably valid. In the *Flu*, in turn, membership functions and generalized constraints are not specified, and they are a matter of perception rather than measurement. In addition, in the *FLp* we use precise theorems, classical deducibility and formal logic, whereas the *FLu* operates with informal and approximate reasoning.

Hence, precisiation means the conversion of unprecisiated (raw) proposition, p, expressed in a natural language into a computation-ready proposition, p*. In addition, p* may be viewed as a computational model of p [11] [13] [14].

Zadeh's ideas mean that we can apply both traditional bivalent-based and novel approximate validity, definitions, axioms, theories and explanations, inter alia.

The central concept in Zadeh's f-validity is the notion of truth. Unlike in the bivalent case, we can now operate with the degrees of truth, and we may apply the meta-rule that the closer a given statement is to its true counterpart, the higher is its degree of truth. For example, "John's age is 23" is fairly true, provided that John is 20. This principle can also be applied to hypothesis verification, theory formation, model construction and explanation, inter alia [4] [6] [8].

The fuzzy modus ponens, in turn, is the well-known application of this principle in semantic validity, i.e., we can replace the traditional bivalent syllogism "from *A* and *If A, then B*, we infer *B"* with its approximate counterpart "from *A'*, and *If A, then B*, we infer *B'"* in which A is approximately A' and B is approximately B' [12]. Hence, in the latter case we can also draw approximate conclusions. In other words, the approximate conclusion is presupposed to be in the proximity to its true counterpart. This type of f-validity is the corner stone in the *FLe.*

Zadeh's *FLe* provides us a basis for examining approximate concepts, theories, models and explanations. Below we consider examples on these.

4 Approximate Probability in the Conduct of Inquiry

4.1 Basic Tools

According to the *FLe*, the idea on the fuzzy probability generally means approximate probability variables and approximate values of these variables, i.e., statements of the type "the probability that John's age is 20 is very high". Naturally, other variables can also be approximate in which case the foregoing statement would be in the form "the probability that John is young is very high".

This principle, which has already been studied quite much in the literature, seems to be usable to both physicalistic and epistemic probability. In the former case we can apply such approximate probability distributions as approximate normal distributions, in particular in the frequency case. Thus, even the precise inputs can yield approximate outputs, i.e., even precise values of random variables yield approximate probabilities (Fig. 1).

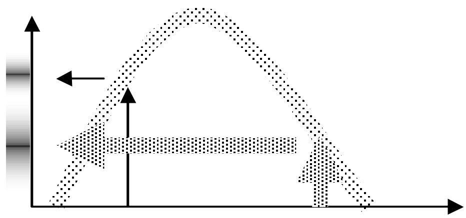


Fig. 1. An approximate probability distribution yields approximate values from both crisp and fuzzy values

Statistical tests apply much physicalistic probability and random distributions when we consider the acceptance of the null and alternative hypotheses according to the tests of significance. We accept the null hypothesis if the value of our test variable does not deviate too much from the “usual” case. Hence, in fact we consider the well-known type 1 error, i.e., we attempt to find justified reasons to reject the null hypotheses. In practice this error is today estimated with the p-value (level of significance) in which case we consider the rejection of the null hypotheses if the p-value is sufficiently small (e.g., $p < .05$). In other words, the p-value is our risk to make an erroneous conclusion if we reject the null hypothesis.

In practice, we have three alternatives, viz. (i) the null hypothesis is true but our data set is exceptional, (ii) the null hypothesis is false and our data set is typical and (iii) the null hypothesis is false and our data set is exceptional.

In the bivalent case the statistical hypotheses are mutually exclusive, and thus the rejection of one automatically means the acceptance of the other. Hence, our reasoning boils down to the metarules (i) If the p-value greater than or equal to _, accept the null hypothesis. (ii) If the p-value is less than _, reject the null hypothesis (and accept the alternative hypothesis instead). The usual threshold values for p are .05 or .01 (5 % and 1 % levels of significance, respectively).

A typical example is the statistical reasoning when the t-test is applied to two independent samples. We thus have the null hypothesis that the there is no difference in the means between the groups, whereas the alternative hypothesis asserts that this difference prevails (the two-tailed case).

Even in the bivalent case we nevertheless take into account the borderline cases when the acceptance of the null hypothesis is considered. If, for example, we establish the threshold value of .05 for its rejection, we in fact pay special attention to the p-values which are in its close neighborhood in order to avoid erroneous conclusions.

Statistical decision making is thus in this sense based on approximate reasoning and probability. Zadeh’s *FLe* can make this reasoning more formal and informative if we operate with the degrees of acceptance and rejection in this context. Then, we could apply such fuzzy metarule as the smaller the p-value, the lower the risk of error for rejecting the null hypothesis (and the higher the degree of rejection for the null hypothesis). Simultaneously, then it also holds that the higher the degree of accepting the alternative hypothesis (Fig. 2).

Fig. 2. Example of a fuzzy relation for assessing the acceptance of the null hypotheses

Hence, within the *FLe* we can take into account better the borderline cases of the p-values, i.e., the area in which we hesitate over our conclusion. This would mean such fuzzy rules for a fuzzy inference engine as

1. If the p-value is not small, then accept the null hypothesis.
2. If the p-value is small, then reject the null hypothesis.

However, more concrete tools for considering this problem are still expected within the *FLe*.

As was stated above, another approach for the verification of hypotheses was adopted by the epistemic approaches. In particular Carnap and such Bayesians as de Finetti and Keynes considered this problem in the light of inductive reasoning and they assumed that the probability for accepting a hypothesis is based on our evidence. Thus, this examination is closely related to decision making under uncertainty.

If we also adopt the cognitivistic approach, i.e., that we should aim at true hypotheses and theories, we can consider the inductive support and degree of confirmation for our hypotheses. In the former case we apply such epistemic utilities as the truth, simplicity or information contents, whereas in the latter case the amount of their true logical consequences play an essential role [2] [9]. In both cases our examination stems from the problems of truth and reasoning, and the *FLe* enables us to also use approximate reasoning and the idea on the degree of truth.

The simplicity of hypotheses is widely examined in the context of curve-fitting problems, and there are already a lot of examples in the literature on fuzzy systems how the application of the *FLe* can yield various simpler models than the corresponding conventional ones.

The degree of confirmation, in turn, is often related to the hypothetico-deductive method and thus to the above-mentioned degree of belief in epistemic probability. In the traditional case the hypothetico-deductive reasoning is essentially based on the bivalent modus tollens syllogism, i.e.,

1. If hypothesis h, then consequence c, and no consequence c
2. Thus, hypothesis is false (reject it)

Hence, given the hypothesis *h* and its observable or testable logical consequence *c,* the justifiability of *c* is assessed according to our knowledge, experiments and observations. If these facts are inconsistent with *c*, the modus tollens reasoning yields the conclusion that our hypothesis is false.

The problem with the classical modus tollens is that if our experiments correspond with *c*, it will not provide us with any resolution. Hence, in the latter case we have to replace deduction with induction, and Popper, for example, applies the idea on the degree of confirmation in this case [9]. Sufficient degree of confirmation, in turn, will lead to the truth or acceptance of the hypothesis.

Popper's approach was bivalent by nature, and thus the approximate version of the modus tollens seems more versatile if we apply the metarule that the more convincing the evidence for the hypothesis, the higher the degree of truth of our hypothesis. We can also use the concepts "degree of confirmation" or "degree of acceptance / rejection" in this context if necessary. Then, for example [8],

1. If hypothesis h, then consequence c, and almost the consequence c (fairly true)
2. Thus, hypothesis can even be fairly true (even fairly high degree of confirmation)

In this manner we can also apply the hypothetico-deductive method when the second premise is non-false.

In the previous examples the implication in the modus tollens was true, but it can also be non-true within the *FLe*. For example, if this implication is only fairly true, we can establish that our hypothesis is not necessarily false even though the evidence is fully inconsistent with the consequence. Equally, only partially supporting evidence for the consequence may already lead to the conclusion that the truth value of our hypothesis is anything between false and true. We may thus assume that the non-true implications in the modus tollens cause more dispersion, granulation or imprecision to our conclusions, and loose reasoning links of this type are typical particularly in the human sciences in which we usually operate with noisy data and the complex interrelationships between the variables.

The idea on the second-order probability under the *FLe* is another interesting issue. In this context we can examine such statements as "the probability that 'the probability of John being very young is fairly high' is high". However, this subject matter, which would extend a new frontier in the approximate probability theory, still awaits much further studies.

4.2 Prospects for Theory Formation and Scientific Explanation

At a more general level, approximate probabilities can play an active role in theory formation and scientific explanation.

If our theories comprise probabilistic laws or law-like statements, we can use such laws as "In Finland women live longer than men" or "The average age of all the cells in an adult person's body is about ten years". [2]. In particular, the *FLe* opens new prospects for the role of idealized laws in science. We know that many laws or models are formulated in an idealized or simplified form which means that they only approximately correspond with the real world. These laws or models are thus non-true, de facto.

By using approximate laws instead, we maintain that these laws are in the neighborhood in their true counterparts, and we can assess the degree of truth of our idealized laws by considering their similarity with their true counterparts in the real world. Now, there is not necessarily need for any additional concretizations of the idealized laws.

In the scientific explanations, in turn, we assume that in an approximate explanation model the explanans is in the neighborhood of its true counterpart, i.e., approximate explanations approximately explain the phenomenon under consideration [4].

The probabilistic and statistical explanations are based on probabilistic models and statistical results, respectively. In these explanation models we can also apply the *FLe*, and approximate probability variables, the values of these variables and approximate probability distributions are central once again.

For example, we can explain that most Swedish men are tall because their average height is approximately 180 cm. In particular in the human sciences these explanation models are relevant because the human behavior is usually indeterministic by nature.

5 Conclusions

We have applied Zadeh's extended logic to the probability theory. At a specific level, we examined how approximate probability functions and values of random variables can be used when various approaches to the conceptions of probability are adopted. At a more general level, we considered the role of approximate probability in hypothesis verification, theory formation and scientific explanation.

It seems that, first, approximate entities and expressions are actually used in the traditional probability theory but usually without any formal basis. Second, on some occasions, such as in the epistemic theories in which we operate with the degrees of beliefs and confirmations, only bivalent resolutions have been provided thus far. If we apply the *FLe*, we can formalize the use of approximate probabilities and resolve most of the preceding problems. The examination of approximate hypothesis verifications, theories and explanations also opens novel prospects in the conduct of inquiry. However, a lot of more concrete studies are still expected in these research areas in the future.

References

1. Carnap, R.: The Logical Foundations of Probability. The University of Chicago Press, Chicago (1962)
2. Niiniluoto, I.: Tieteellinen päättely ja selittäminen (Scientific Reasoning and Explanation). Otava, Keuruu (1983)
3. Niiniluoto, I.: Truthlikeness. Reidel, Dordrecht (1987)
4. Niskanen, V.A.: Application of Zadeh's Impossibility Principle to Approximate Explanation. In: Proceedings of the IFSA 2009 Conference, Lisbon, pp. 352–360 (2009)
5. Niskanen, V.A.: Fuzzy Systems and Scientific Method – Meta-Level Reflections and Prospects. In: Seising, R. (ed.) Fuzzy Set Theory - Philosophy, Logics, and Criticism, pp. 51–82. Springer, Heidelberg (2009)
6. Niskanen, V.A.: Meta-Level Prospects for Scientific Theory Formation by Adopting the Information Granulation Approach. In: Proceedings of the IMS 2008 Conference, Sakarya University, pp. 33–41 (2008)

7. Niskanen, V.A.: Soft Computing Methods in Human Sciences. Studies in Fuzziness and Soft Computing, vol. 134. Springer, Berlin (2004)
8. Niskanen, V.A.: The Role of Approximate Reasoning in Hypothesis Verification. In: FUZZYSS 2009, Ankara, pp. 126–134 (2009) (Full paper in progress)
9. Popper, K.: The Logic of Scientific Discovery. Hutchinson, London (1959)
10. Von Wright, G.H.: The Logical Foundations of Probability. Blackwell, Oxford (1957)
11. Zadeh, L.: From Computing with Numbers to Computing with Words - From Manipulation of Measurements to Manipulation of Perceptions. IEEE Transactions on Circuits and Systems 45, 105–119 (1999)
12. Zadeh, L.: Fuzzy Logic = Computing with Words. IEEE Transactions on Fuzzy Systems 2, 103–111 (1996)
13. Zadeh, L.: Toward a Theory of Fuzzy Information Granulation and Its Centrality in Human Reasoning and Fuzzy Logic. Fuzzy Sets and Systems 90/2, 111–127 (1997)
14. Zadeh, L.: Toward a Perception-Based Theory of Probabilistic Reasoning with Imprecise Probabilities. Journal of Statistical Planning and Inference 105/2, 233–264 (2002)
15. Zadeh, L.: Toward Extended Fuzzy Logic – A First Step. Fuzzy Sets and Systems 160, 3175–3181 (2009)

Comparing Ontologies Using Multi-agent System and Knowledge Base

Anne Håkansson[1], Ronald Hartung[2], Esmiralda Moradian[1], and Dan Wu[1]

[1] Department of Communication Systems
Royal Institute of Technology, KTH
Electrum 418, SE-164 40 Kista, Sweden
annehak@kth.se, esmirald@dsv.su.se, danwu@dsv.su.se
[2] Department of Computer Science
Franklin University
Columbus, Ohio, USA
Hartung@franklin.edu

Abstract. This paper presents an approach for handling several ontologies in a domain by integrating a knowledge base in a multi-agent system. For some online facilities, like e-business, several ontologies in different languages are needed to match the users' request. Finding the ontologies is one of the tasks, comparing and combining these ontologies is the other. The accomplishment of these tasks depends on the content in the ontologies like tags and structure but, in some cases, also language and matching techniques. Matching the contents is difficult due to differences between ontologies, often resulting from the lack of explicit and exact standards and development guidelines. This complication increases with the ontologies diverged languages. These problems are tackled by applying a multi-agent system wherein the agents, i.e., software agents and meta-agents, use the users' request to search for ontologies and the knowledge base to compare and combine the contents of the ontologies to create an overall solution. The software agents search for ontologies; the meta-agents keep track of the software agents, ontologies and the knowledge base that reasons with the contents of the ontologies.

Keyword: Multi-agent systems, Software Agents, Meta-agents, Ontologies, Knowledge bases, ´Reasoning.

1 Introduction

Many ontologies are developed for e-business, supporting e-commerce and e-services [3; 5]. These ontologies assist in online commerce and services, like supporting customers' purchasing products, travel trips and making restaurant reservations, online, as well as, facilitating companies buying products for their own manufacturing. The benefit is that these facilities give competitive advantages to the enterprises by supplying additional and extra-ordinary services and offers. The facilities can also help the enterprise to expand the number of customers, nationally and internationally.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 124–134, 2010.

Ontologies can be seen as templates for collecting information from individuals in order to complete tasks. In addition to collect data, the ontologies apply structures and build relationships between concepts. The notion of ontologies is an explicit specification, which can be used to model a domain with objects, concepts, properties and relations [3]. However, the ontologies lack of explicit and exact standards for designing, structuring and implementing their contents make them difficult to compare. A comparison facility can make a combination of ontologies applicable for e-business, especially, if the combination and interaction of the ontologies benefit the customers in purchases and, thereby, benefit selling companies. Additionally, a system can produce significant business value by providing opportunities to the companies or travellers, which can be new or additional information about offers that otherwise might be unknown. For example, companies can order sets of parts from the same enterprise and bargain for quality, price and delivery dates; an e-tourist can make reservations for flights, hotels, trips within the destination country, and events that take place at the time for travelling.

To support e-business customer, with commerce and services providing additional and targeted information, we propose a multi-agent system (MAS) that utilises both ontologies and knowledge bases to compare and combine the ontologies. The MAS uses software agents to search for ontologies that match the user's request and meta-agents to handle the contents of the different ontologies. From the user's request and the ontologies, the meta-agents compare similar parts with knowledge from the knowledge bases. With this knowledge, the meta-agents can support reasoning with the ontologies and combine several different ontologies.

2 Related Work

Multi-agent systems have been used for e-business systems. These systems include agent approaches to e-business with negotiation and user preference [13; 7; 15]. The applications range from buying and selling products, including information products, to optimizing traveller's arrangements, where agents have all kinds of roles [16, 2, 1].

An earlier attempt using MAS for business applications in multilingual ontologies [11] is a system that receives user requests and searches for the ontologies in the language of the destination country [9]. The MAS locates the ontologies that correspond to the request and returns the links and paths to the ontologies. The system upholds communication with the users by keeping track of the information that needs to be supplied. Moreover, the system performs the mapping between ontologies and language translation and can be used to reason with the content [6, 9]. The research in this paper is an extension of this earlier work. It extends the MAS with a better tool to reason with the ontologies, i.e., a knowledge base and interpreter. By extension, the agents cannot only reason with the contents of each ontology and make sure that they correspond to the users' request, the knowledge base can also support comparing the different ontologies and support filling in missing information into web sites, linked from the ontologies.

Another similar work is ontology-based context modeling and reasoning using OWL [18]. The ontology is in the context of knowledge management, where the ontology is referred to as a shared understanding of some domains. The reasoning is a logic-based context-reasoning scheme that reasons with the environment based on the content of the ontology [18]. The ontology is used to reason with the context, whereas we use a knowledge base as a knowledge source when reasoning with the ontologies, which has the context association for the task. The context is not derived from the ontologies – instead the agents' tasks handle the context and matching.

Knowledge is generally used to guide the agent's autonomous local decision-making processes, supporting agents' problem solving by providing expertise or search in the agents' database [12]. This is an interesting approach, but we use the knowledge and an interpreter to reason about the contents of the ontologies and support the users supplying information. Our approach guarantees that the meta-agents and interpreter use the same knowledge. The meta-agents are autonomous and perform simple matching by using the knowledge base to compare the contents of the ontologies; the interpreter perform advanced matching by interpret the knowledge in the knowledge based against the contents in the ontologies.

3 The Ontologies

Ontologies can form a formal and explicit specification [3] of a set of concepts and relationships in a domain. This specification is critical to the systems providing automatic search of the ontologies and reasoning with their contents. However, often the specifications differ, since there are no agreements upon standards for designing, structuring and implementing the ontologies. Hence, most of this work i.e., search and reasoning, is performed manually by users providing intelligence to accomplish the tasks [4], but the amount work requires a tool to perform the work, automatically.

The ontologies can be seen as a knowledge representation of parts of the world, captured entities, ideas, and events, with their properties and relations. They are modelled to gain understanding of existing things and their relationships in a language close to natural language [10]. Among different usages, the ontologies are created for business processes to construct models of services. Often the business process definitions and related artefacts (e.g. business documents, business objects) are represented as knowledge based on ontologies [10].

When using ontology in business processes, there are two aspects. First is the classic view, where the ontology describes the object of commerce, or the objects of the process space. The objects of commerce are the things bought or sold, plus properties of those objects. We can extend the properties to include aspects like warranties, conditions of sale, payment terms, as well as, physical properties of the objects like color, size, and functionality. The second is the business process, which is a sequence of steps and may include conditionals, branching (alternatives) and loops. The steps must be concrete and well defined. The advantage of ontology, or other, formal representation is that a reasoning system can enact the process.

A more complex issue is being able to understand the process of a correspondent system, that is an external system to interact with for a business purpose. Here, the ontology, acting as a description of the process, lends itself to automating the transactions and even applying negotiation to the business process. These business processes comprise a known set of various purchase processes and other processes to support searching and presenting options for purchase services that are expressed in ontologies

[9]. If only one company is used for the e-business task, only a single ontology might be needed for handling the task, like combined facilities offered by travel agencies. However, if several ontologies are needed in conjunction, this becomes complex. A tool for handling several ontologies, automatically, can simplify the process of purchasing services from several distinct companies.

4 The Multi-agent System

The multi-agent system is developed to support users finding acceptable solutions to tasks in e-business. The tasks are sometimes complex and may require several ontologies to accomplish the assignments. These ontologies can be used as a collected set that enables the handling of a complex endeavour rather than one task at the time.

Working with several different external ontologies is difficult. These ontologies can have different structures, use a variety of names on the tags and require a variety of data and information since the design is due to the developers' structure choices and application of languages' standards. These differences must be taken care.

The purpose of the system is to analyse the content and combine the ontologies to create more complete solutions, as well as, support filling-in the missing information. Thus, the system must check the ontologies to find the web sites in which parts are to be filled in and communicate with the users to collect the information.

The MAS has connection with the web. The ontologies are fetched from the web, but the system compares the contents of the ontologies locally, by working with copies of the ontologies. The reason is that the web is expanding rapidly, which requires locally handling services on the computer rather than on the web. Nonetheless, the MAS needs to communicate with the original ontologies since these ontologies contain links to the web sites at which the missing information needs to be filled-in.

The system is a MAS with a knowledge base. The system is not purely an agent-based system since it uses a knowledge base to reason with the known knowledge components of the ontologies and also incorporate new parts. Neither is the system a knowledge-based system since the agents deal with the contents of the knowledge base, not the reasoning strategies. In the system, the reasoning strategies are not steering the process of providing the system with parts that it needs to draw conclusions. Neither are the users are providing the system with facts to operate on, the agents are. This works by splitting the matching and combining functions between the meta-agents and the knowledge base. For simple matching cases, the meta-agents can handle the match. The rules of a knowledge base provide a more powerful mechanism for matching and combining ontologies. The other significant feature of the rules is the ability to modify and add rule-based match reasoning. The modification and adding are performed as the system is being refined.

The multi-agent system contains software agents and meta-agents, a container with copies of the ontologies, a knowledge base with knowledge about the content of the ontologies and an interpreter that interpret the contents for architecture see Figure 1.

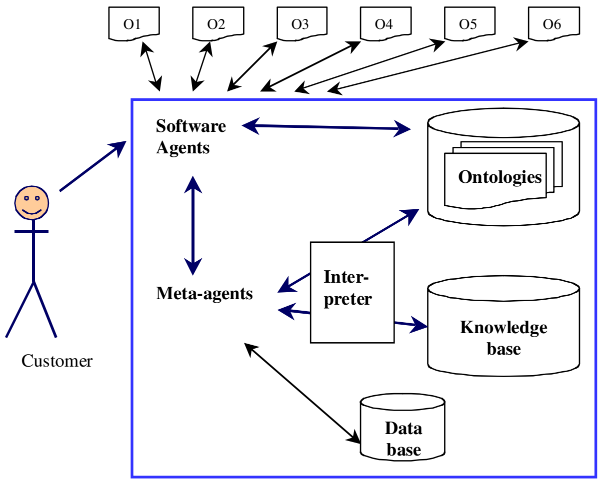


Fig. 1. The architecture of the multi-agent system

Briefly, the system procedure is: the customer launches the system with a user request. The software agents receive the request and look for the ontologies that match the request. The found ontologies, presented as O1-O6 up in the figure, are copied to an "ontologies-container", towards which the meta-agents work. The meta-agents use the knowledge base to compare the ontologies and the interpreter to interpret the knowledge in the knowledge based against the contents in the ontologies. Thus, the meta-agents make the simple matching by comparing different ontologies against the knowledge base. If there is a direct match the meta-agents can continue to work with the next part of the ontologies. However, not all the information, tags or names are represented in the knowledge base and, therefore, some interpretations of the rules are needed. The interpreter matches the content by inferring conclusions that can match the ontologies. If not, additional facts have to be provided by the users. The results of the interpretation and the user-given information are temporary stored in a database.

4.1 The Agents in the Multi-agent System

The multi-agent system uses software agents and meta-agents, to search for the ontologies and compare and combine the contents of the ontologies. The software agents are like search agents finding the ontologies that match the user requests. The meta-agents are managers that control the software agents and handle the ontologies. Although we use software agents, it is possible to use intelligent agents that learn from the environment. However the web is a changing environment and, for the current work in this paper, it is better to search for new information, every time the user poses a new request.

The system is launched by a user request. From this request, the software agents start searching for ontologies, independently of each other. The search is performed without any authorised control or user involvement, which is a benefit with agents [14]. Thus, the agents must be capable of autonomous actions, situated in different environments [17] where the agents independently search for the ontologies and return with locations for each ontology (the URL of the ontologies).

The MAS works internally on host computer and externally on the web, relieving the web by reducing computation. The internal part includes most of the MAS, handling the agents, communicating with the web, comparing the content, whereas searching is made externally. The combination of internal and external parts makes the agents' working environment interesting. The internal environment is fully observable, deterministic, episodic, discrete and static whereas the external environment is partially observable, stochastic, episodic, continuous, and dynamic.

In the internal system, the meta-agents can observe the environment and they know the next event, and are deterministic. Moreover, the task is divided into atoms, performed one atom at the time, i.e., episodic, and each agent has a finite number of states and actions it can take, i.e., discrete, in a static environment.

In the external environment, the agents, the software agents and the meta-agents, only partly know the environment. These agents know the closest nodes but not the nodes, at a distance. Although, the agents can learn the environment, it changes often enough to stay partly observable. The agent cannot predict the behavior of the environment since the states, which will result from performing an action, are unclear.

In the environment there will be stochastic elements, i.e., web sites and ontologies that randomly appear. The software agents task is searching ontologies but the contents and locations of the ontologies can vary with time as well as, finding ontologies, depending on when actions these are performed. Hence, it difficult to reuse all of results of early searches since the environment is highly unpredictable.

The tasks are episodic and the software agents work with one task at the time. Hence, the software agent has one assignment, which is following a connection to its node. If this node has several connections, there will be equally many software agents. When the agents reach a destination, node, which is not an end node, the agents will launch other agents [8], one for each new connection. To keep track of the software agents, we use meta-agents. These meta-agents are built on the software agents and can be used to influence the further work of the agents.

The environment is dynamic and changes on a daily bases. Even though the connections are not changed, the content of the web sites and ontologies can change. Moreover, the user requests change either with refinement of old requests or a totally new request. Hence, for each new request, the software agents need to restart the search. Only the closest nodes are reused and, further away from the starting point, new unexplored parts, most certainly, will be exercised.

The meta-agents become a chain of software agents, where the first agent is the first one launched from the user request and the last agent is the one that reached a web site, successfully. The chain is a path that successfully leads to destination of an ontology that corresponds to user request. These ontologies are tagged for further comparisons and combinations.

One problem with software agents is the continuous work on the web. There might be uncountable number of states and actions, which arise from the continuous time problem. The solution we use is execution suspension that controls the agents' performance to prevent them to continue the search, endlessly.

When the agents finish the search, the agents bring back a number of ontologies to the MAS. The meta-agents start analyzing these ontologies to compare their contents.

The comparing of the ontologies depends on the content in the ontologies where the structure is one of the major obstacles and the language of the ontologies is another. The first step is to use a matching technique that compares the ontologies with the user's request, which is more sophisticated than the software agents' search. The second step is to establish the equivalence between the ontologies.

The meta-agents need different functions: search meta-agent that contains the paths to the original site of the ontology and reasoning meta-agent that supports comparing the ontologies to the user's request, as well as, to the other ontologies and the knowledge in the knowledge base. These meta-agents are then used to combine the ontologies to give total solution of a request and support filling-in data in the ontologies. There are many parts to keep track of: the user request, the ontologies, the parts that have been analysed and are successfully compared to the knowledge base, the parts that are not found in the knowledge base but in the other ontology, and parts in the ontology that are found in the knowledge base, but not in the other ontology.

The knowledge can be stored centrally as a central unit or locally in the agents [12]. If the knowledge is central, it can be easily accessed be all the parts in the system. If the knowledge is local, the different agents need to be contacted to retrieve the knowledge. Since the accessibility is important, as well as, support comparison in parallel and frequently updating the knowledge, a central part is more effective than have agents carrying around knowledge, hence a separate knowledge base.

4.2 The Knowledge Base in the Multi-agent System

When the MAS performs complex tasks, it can benefit from a supporting tool, like a knowledge base. The knowledge base can provide necessary knowledge and guidance to the MAS, to perform more efficiently.

The agents search for the ontology by a search string in, e.g., the URL, annotations, definitions, and classes or properties. When the user requests ontologies, the request is transformed into a term, which is then specified in the search string, for example, if user request includes "travel", the term can be *url:travel* or *def:travel*. These are terms are implemented in the knowledge base, as well as, other well-known terms.

The knowledge base includes parts that can be found in the ontologies. Some commonly used fields, which is also used to constraint user requests, are basic metadata, document metadata and RDF metadata. The basic metadata includes *url:, desc:*, *def:*, *ref:*, *pop:*, and *ns:*, The document metadata consists of *hasEncoding:*, *hasLenght:*, *hasMd5sum:*, *hasFiletype:*, *hasDateLastmodified:*, and *hasDataCache*. The RDF metadata has *hasGrammar:*, *hasCntTriple:*, *hasOntoRatio:*, *hasCntSwtDef:* and *hasCntInstance*. For the description term, there are other terms associated to it. For example, *log:forSome*, *title*, *s:label*, and *s:subClassOf*. To the term "Class", there are terms like *s:comment*, *s:label*, *ont:UniqueProperty*, *s:domain*, and *s:range*. All these terms have to be represented as knowledge to be able to compare the content of the ontologies.

In the knowledge base, there are facts and different levels of complex rules. Some terms can be matched directly to the facts whereas other terms must use rules to reach a conclusion that matches the contents of the ontologies. Thus, these are matching terms in the rules of the knowledge base. The rules are either compound terms or can have different names leading to the same conclusion about the term. These rules are used to lead the same term in the ontologies to be compared.

When a term in the ontology is not directly represented in the knowledge base, it can be represented as a rule with several different alternatives. An example of several alternatives is synonyms, which can be represented as rules in the knowledge base.

The terms that are not covered by the knowledge base are stored temporarily. These are stored while finding out if they are significant and if there are any correspondences to other ontologies. If so, the new term can be permanent in the knowledge base, either as facts or as rules. It becomes a fact if it is direct correspondence; a rule if the other rules are used to reach conclusion about the term. Letting the knowledge base upgrade itself can be dangerous if it learns wrong terms but important for further communication. Therefore, expanding the knowledge base should be made under supervision of knowledgeable users.

4.3 The Interpretation

The MAS can either be ruled by the knowledge, or use the knowledge to perform tasks. Then, the interpretation mechanism can either be the engine that steers the system and its actions or the system can use the interpretation for partial solutions and use the results for further operation. We use the latter.

To reason with ontologies, and find correspondences between the ontologies, the system must inspect the ontologies and compare their contents to the knowledge base and to other ontologies. This is performed by the meta-agents and the interpretation.

The knowledge is accessible whenever the agents need it. By the knowledge in the knowledge base, it is possible to derive the different tags and match those to knowledge about the tags and text. If common ontology languages are used, tags are commonalities in the ontologies. This is not always the case and to get more accurate information about the tag equality, text is used. Even though the tags are not the target for the comparing the ontologies, they still need to be compared because they give additional information that can be decisive importance for the of the text comparison.

The interpreter is an inductive interpreter that works through the ontology with the knowledge from the knowledge base. It checks if the tags are found in the knowledge base and if these tags need to be explored further. The exploration occurs if the ontology links to web sites. All web sites must be examined to find out if information is missing. If the tags that are not present in the knowledge base is temporary stored in a database. These are used for the comparison of two ontologies.

All end tags and comment tags are left out since the knowledge base is not a compiler or an automatic documentation tool. Moreover, these tags do not add any additional information to the interpreter or the users. Nonetheless, if they would be significant, it is easy to include them in the knowledge base.

Then the text, associated with the tags is checked. Commonly the text is a string that starts with # , http:// or a word like Resource. These are important when it comes to comparing ontologies.

For the comparison, a second ontology is worked through. For each tag in the first ontology, the whole second ontology is checked until the tag is found. The tags are compared by interpret them with the knowledge in the knowledge base. This because it might not be a direct matching between the tags as when comparing the words presented as facts in the knowledge base. For example, the word professor can be the same as senior researcher, and senior lecture the same as teacher.

It can also be by using rules, where an *#Activity* can be *#BunjeeJumping*, *#Safari*, *#Sunbathing*, or *#Yoga*. This is then presented in a rule. To illustrate the comparison, we present real examples of the content in ontologies found on the web, in Figure 2.

```
<rdf:RDF
xml:base="http://www.mindswap.org/2004/multipleOnt/F
actoredOntologies/factoredTravel/FactoredActivity.owl">
<owl:Ontology rdf:about=""/>
<owl:Class rdf:about="#Adventure">
<rdfs:subClassOf>
<owl:Class rdf:about="#Activity"/>
</rdfs:subClassOf>
</owl:Class>
…
<owl:Class rdf:about="#Adventure">
<owl:disjointWith>
<owl:Class rdf:about="#Sightseeing">
  </owl:Class>
</owl:disjointWith>
</owl:Class>
…
<owl:ForeignClass
rdf:about="http://www.mindswap.org/2004/multipleOnt/F
actoredOntologies/factoredTravel/FactoredContact.owl#C
ontact">
<owl:foreignOntology
rdf:resource="http://www.mindswap.org/2004/multipleOn
t/FactoredOntologies/factoredTravel/FactoredContact.owl
"/>
</owl:ForeignClass>
<owl:LinkProperty rdf:about="#hasContact">
<owl:foreignOntology
rdf:resource="http://www.mindswap.org/2004/multipleOn
t/FactoredOntologies/factoredTravel/FactoredContact.owl
"/>
```

```
<rdf:RDF>
<owl:Ontology rdf:about="">
<rdfs:comment>Travel Itinerary</rdfs:comment>
<owl:versionInfo>
$Id: itinerary-ont.n3,v 1.4 2003/10/03 20:05:44 mdean
Exp $
</owl:versionInfo>
</owl:Ontology>
<Aircraft
rdf:about="http://www.daml.org/2001/06/itinerary/itin
erary-ont#A300">
   </Aircraft>
…
<owl:Class
rdf:about="http://www.daml.org/2001/06/itinerary/itin
erary-ont#Aircraft">
<owl:oneOf rdf:parseType="Resource">
<rdf:first
rdf:resource="http://www.daml.org/2001/06/itinerary/i
tinerary-ont#First"/>
<rdf:rest rdf:parseType="Resource">
…

<owl:oneOf rdf:parseType="Resource">
<rdf:first
rdf:resource="http://www.daml.org/2001/06/itinerary/i
tinerary-ont#First"/>
<rdf:rest rdf:parseType="Resource">
```

Fig. 2. Ontologies being compared and combined

In the figure, the first ontology is presented to the left and the second ontology to the right. The first is checked against the knowledge base, with tags, such as, *RDF*, *owl:Ontology*, *owl:Class*, *rdfs:subClassOf*, *owl:disjoint*, *owl:ForeignClass*. These are all found in the knowledge base. The interpreter turns to the second ontology and checks if the first tag in the first ontology also is present in the second, thus *<rdf:RDF>*, and then saves *xml:base=http://…*. Then the interpreter checks *owl:Ontology*, which is also present in the second ontology and moves to *owl:Class*. The first line in the second ontology is the comment, *versionInfo*, followed by end tags, which are all skipped until *owl:Class* is found. In the first ontology *#Adventure* is found and in the second a link to itinerary and aircraft. These are not compared but stored for further investigation later on. The adventure has some categories in the ontology that needs to be handled and a link to a web site, which is fetched and checked for parameters or requirements for user given information. When a statement is not known, we assume it is true until proven, hence, open world assumption.

For combining the ontologies, there must be parts that can be matched or related. Consider the ontologies in Figure 2, again, it is possible to combine those into a third ontology (O) or a document with ontological structure. The entities that are in the first ontology (A) and the second ontology (B) and matches the knowledge in the knowledge base, can be built into the third ontology, hence, A + B -> O. This means that every entity A1, of ontology A and entity of B1 of ontology B, is inserted in ontology O in such way that:

```
If the entity A1 (and B1) is found both A and B then insert
the entities in O and give the correspondence of equal,
subset partial or coverage.

If the entity A1 is found in either A or B, then insert the
entity A1 into O.

If the entity B1 is found in either A or B, then insert the
entity B1 into O.
```

The process proceeds until significant entities are found and matched and become a combined ontological structure with links to both ontologies A and B. As mentioned above, in some cases the users must be involved in this process.

5 Conclusion and Further Work

This paper presents an approach to handle ontologies using a multi-agent system with a knowledge base, which support online facilities, like e-tourism and e-business. The system searches for relevant ontologies, as well as, comparing and combining the ontologies by analyzing the structure of the ontologies. If good match is found, the multi-agent system can support filling in missing data in the ontologies.

In the multi-agent system, software agents and meta-agents are used to search for ontologies from the the users' request and work with the contents of the ontologies The meta-agents keep track of the software agents, ontologies and knowledge in the knowledge base and reason with contents of the ontologies.

The knowledge base is at a novel level handling RDF ontologies and, thus, needs to be expanded with more knowledge to handle many different kinds of ontologies with different structures. Also testing is needed to make sure that a high level of correspondences is met. A small number of discrepancies can become extremely costly. The combination is an interesting technique since, if tuned and thoroughly tested, it can be used to developed ontologies automatically. This can minimize problems with mistakes when using two or several different ontologies.

Acknowledgement

This work is supported by Sweden-Korea Research Cooperation Programme funded by STINT, The Swedish Foundation for International Cooperation in Research and Higher Education. http://www.stint.se/

References

1. Apelkrans, M., Håkansson, A.: Information coordination using Meta-Agents in Information Logistics Processes. In: Lovrek, I., Howlett, R.J., Jain, L.C. (eds.) KES 2008, Part III. LNCS (LNAI), vol. 5179, pp. 788–798. Springer, Heidelberg (2008)
2. Bhavsari, V., Boley, H., Yang, L.: A Weighted-tree similarity algorithm for Multi-agent systems in E-Business environment (2004), http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.89.8260&rep=rep1&type=pdf

3. De Bruijin, J., Fensel, D., Keller, U., Lara, R.: Using the Web Service Modeling Ontology to enable Semantic eBusiness. Commnications of the ACM 48(12) (2005)
4. Flett, A., Brown, M.: Enterprise-standard ontology environments for intelligent e-business (2001), http://www.csd.abdn.ac.uk/~apreece/ebiweb/papers/flett.pdf
5. Gordijn, J., Akkermans, H.: Designing and Evaluating E-Business Models. IEEE Intelligent Systems Archive 16(4), 11–17 (2001)
6. Hartung, R.L., Håkansson, A.: Using Meta-agents to Reason with Multiple Ontologies. In: Nguyen, N.T., Jo, G.-S., Howlett, R.J., Jain, L.C. (eds.) KES-AMSTA 2008. LNCS (LNAI), vol. 4953, pp. 261–270. Springer, Heidelberg (2008)
7. Helal, S., Wang, M., Jagatheesan, A.: Service-centric brokering in dynamic e-business agent communities. JECR, Special Issue in Intelligent Agents in E-Commerce (2001)
8. Håkansson, A.: An Approach to Event-Driven Algorithm for Intelligent Agents in Multi-agent Systems. In: Nguyen, N.T., Jo, G.-S., Howlett, R.J., Jain, L.C. (eds.) KES-AMSTA 2008. LNCS (LNAI), vol. 4953, pp. 411–420. Springer, Heidelberg (2008)
9. Håkansson, A., Hartung, R.L., Jung, J.J.: Using Multi-Agent System to handle information in Multilingual Ontologies. Accepted in Smart Innovation, Systems and Technologies 2010 (2010)
10. Jenz, D.: Business Process Ontologies: Speeding up Business Process Implementation (2003), http://ip-68-178-224-58.ip.secureserver.net/publications/07-03%20WP%20BP%20Ontologies%20Jenz.pdf
11. Jung, J., Håkansson, A., Hartung, R.L.: Indirect Alignment between Multilingual Ontologies: a Case Study of Korean and Swedish Ontologies. In: Håkansson, A., Nguyen, N.T., Hartung, R.L., Howlett, R.J., Jain, L.C. (eds.) KES-AMSTA 2009. LNCS, vol. 5559, pp. 233–241. Springer, Heidelberg (2009)
12. Marik, V., Pechoucek, M., Stepankova, O.: Social Knowledge in Multi-agent Systems. In: Luck, M., Mařík, V., Štěpánková, O., Trappl, R. (eds.) ACAI 2001 and EASSS 2001. LNCS (LNAI), vol. 2086, p. 211. Springer, Heidelberg (2001)
13. Purvis, M., Nowostawski, M., Oliveira, M., Cranefield, S.: Multi-agent interaction protocols of e-business. In: IAT IEEE/WIC, October 13-16, pp. 318–324 (2003)
14. Russell, S.J., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice Hall, Upper Saddle River (2003), ISBN 0-13-790395-2
15. Shin, H., Bo-Yeon Shim, B.: e-Business Agent Oriented Component Based Development for Business Intelligence. In: Gabrys, B., Howlett, R.J., Jain, L.C. (eds.) KES 2006. LNCS (LNAI), vol. 4252, pp. 803–811. Springer, Heidelberg (2006)
16. Wickramasinghe, L.K., Amarasiri, R., Alahakoon, L.D.: A Hybrid Intelligent Multiagent System for E-Business. Computational Intelligence 20(4), 603–623 (2004)
17. Wooldridge, M.: An Introduction to MultiAgent Systems. John Wiley & Sons Ltd., Chichester (2002), ISBN 0-471-49691-X
18. Wang, X.H., Gu, T., Zhang, D.Q., Pung, H.K.: Ontology Based Context Modeling and Reasoning using OWL. In: IEEE International Conference on Pervasive Computing and Communication (PerCom 2004), Orlando Florida, March 14-17 (2004)

ACTL Local Model Update with Constraints

Michael Kelly[1], Fei Pu[2], Yan Zhang[1], and Yi Zhou[1]

[1] Intelligent Systems Laboratory
School of Computing & Mathematics
University of Western Sydney, Australia
{mkelly,yan,yzhou}@scm.uws.edu.au
[2] School of Computer and Information Engineering
Zhejiang Gongshang University, China
pufei@zjgsu.edu.cn

Abstract. The recent development of model update aims to enhance model checking functions and provides computer aided modifications in system development. On the other hand, constraints have been playing an essential role in describing rational system behaviours. In previous model update approaches, constraints are usually not considered in the update process. In this paper, we present an ACTL - a widely used fragment of Computation Tree Logic (CTL), local model update approach where constraints have been explicitly taken into account. This approach handles constraints effectively by integrating constraint automata into the underlying model update. We demonstrate the effectiveness of our approach through the case study of the correction of the well known mutual exclusion program.

1 Introduction

In current approaches of system verification and modifications, an important idea is to integrate AI techniques, such as belief revision and model update, into model checking to develop effective system repairing tools, e.g. [3,7,11]. However, a major obstacle restricting this idea to apply to large scale problem domains, is that the update prototype has to take the entire system model (e.g. a complete Kripke model) as the input. This, obviously, is not generally feasible for large scale domains.

On the other hand, it is well understood that counterexamples generated from a model checking procedure plays an essential role in system repairing, because a counterexample usually localizes certain information that reveals how the system fails the specification property, e.g. [6]. Therefore, a natural way to overcome the difficulty of model update approach mentioned above is that we should develop a local model update that only applies to counterexamples and effectively generates candidate modifications for the original system. Furthermore, since constraints have been used in specifying system's rational behaviours, the underlying update approach should also carefully address this issue during an update process.

In this paper, we present an ACTL local model update approach where constraints have been explicitly taken into account. This approach handles constraints effectively by integrating constraint automata into the underlying model update. We demonstrate the effectiveness of our approach through the case study of the correction of the well known mutual exclusion program.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 135–144, 2010.

2 Preliminaries

ACTL is a special fragment of Computation Tree Logic (CTL) and has attracted considerable studies from researchers, e.g. [2,4]. Besides Boolean connectives, ACTL provides both linear time operators X, F, G and U and the branching time operator A. The linear time operators allow one to express properties of a particular evaluation of the systems given by a series of events in time, and the branching time operator takes into account the multiple possible future scenarios starting from a given state at certain time.

ACTL has the following syntax given in Backus Naur form:

$\phi ::= \top \mid \perp \mid p \mid \neg p \mid \phi \wedge \psi \mid \phi \vee \psi \mid \mathrm{AX}\phi \mid \mathrm{AG}\phi \mid \mathrm{AF}\phi \mid \mathrm{A}[\phi \mathrm{U} \psi]$, where p is any propositional atom (variable). Let AP be a set of propositional variables. A *Kripke structure* M over AP is a triple $M = (S, R, L)$, where S is a finite set of states, $R \subseteq S \times S$ is a binary relation representing state transitions, and $L : S \to 2^{AP}$ is a labeling function that assigns each state with a set of propositional variables.

An ACTL formula is evaluated over a Kripke structure. A *path* in a Kripke structure from a state is a(n) (infinite) sequence of states. Note that for a given path, the same state may occur an infinite number of times in the path (i.e., the path contains a loop). To simplify our following discussions, we may identify states in a path with different position subscripts, although states occurring in different positions in the path may be the same. In this way, we can say that one state precedes another in a path without much confusion. Now we can present useful notions in a formal way.

Let $M = (S, R, L)$ be a Kripke structure and $s_0 \in S$. A *path* in M starting from s_0 is denoted as $\pi = [s_0, \cdots, s_i, s_{i+1}, \cdots]$, where $(s_i, s_{i+1}) \in R$ holds for all $i \geq 0$. If $\pi = [s_0, s_1, \cdots, s_i, \cdots, s_j, \cdots]$ and $i < j$, we denote $s_i < s_j$. For any $s \in S$, the satisfaction relation between (M, s) and an ACTL formula ϕ, denoted by $(M, s) \models \phi$, is defined in a standard way as described in [8].

Now we introduce the concept of tree-like Kripke structures [4]. Let G be a directed graph. A *strongly connected component* (SCC) C in G is a maximal subgraph of G such that every node in C is reachable from every other node in C. C is *nontrivial* iff either it has more than one node or it contains one node with a self-loop. The *component graph* $c(G)$ of G is the graph where the vertices are given by the SCCs of G, and where two vertices of $c(G)$ are connected by an edge if there exists an edge between vertices in the corresponding SCCs. Then we say a graph G is *tree-like* if (1) all its SCCs are cycles; and (2) $c(G)$ is a directed tree. We should note that condition (1) is necessary because some SCCs may not be cycles. For instance, in a graph $G = (V, E)$, where $V = \{s_1, s_2, s_3\}$ and $E = \{(s_1, s_2), (s_2, s_3), (s_3, s_3), (s_3, s_2)\}$, the subgraph $G' = (\{s_2, s_3\}, \{(s_2, s_3), (s_3, s_3), (s_3, s_2)\})$ is a SCC, but it is not a cycle because edge (s_3, s_3) also forms a self-loop.

Consider a Kripke model (M, s_0), where $M = (S, R, L)$ and $s_0 \in S$. We say that (M, s_0) is a *tree-like Kripke model* if its corresponding graph $G(M) = (S, R)$ is tree-like. In this case, we call the initial state s_0 the *root* of this tree-like model. Since a tree-like Kripke model may not be a strict tree (e.g. it may contain some cycles along a branch), we cannot follow the traditional notions of *child and parent* in a tree-like model. Instead, we define the following new concepts. We say state s is an *ancestor* of state s', if there is a path $\pi = [s_0, \cdots, s, \cdots, s', \cdots]$ such that $s < s'$ in π and for all $s^* \in \pi$ where $s^* < s$, $s^* \neq s'$. s is a *parent* of s' if s is an ancestor of s'

and $(s, s') \in R$. In this case, we also call s' is a *successor* of s. A state s is called *leaf* if it is not an ancestor of any other states. A tree-like model (M', s') is called a *submodel* of (M, s), if $M' = (S', R', L')$, $s' \in S'$, $S' \subseteq S$, $R' \subseteq R$, for all $s^* \in S'$, $L'(s^*) = L(s^*)$, and in M, s' is an ancestor of all other states in M'.

Fig. 1 from [4] shows an example of a tree-like counterexample for a specific ACTL formula.

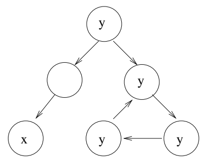


Fig. 1. A counterexample for AG$\neg x \vee$AF$\neg y$

It has been proved by Clarke et al that if an ACTL formula is not satisfied in a Kripke structure, then this Kripke structure must contain a tree-like counterexample with respect to this formula [4].

Theorem 1. *[4] ACTL has tree-like counterexamples.*

3 ACTL Local Model Update

Now we start to formalize the update on tree-like Kripke structures. For convenience, in the rest of the paper, we will call (M, s) a *tree-like Kripke model* without explicitly mentioning the corresponding tree-like Kripke structure $M = (S, R, L)$ where $s \in S$. We also define $\mathsf{Diff}(X, Y) = (X \setminus Y) \cup (Y \setminus X)$ where X, Y are two sets.

Definition 1. *Let (M, s) and (M_1, s_1) be two tree-like Kripke models. We say that a binary relation $H \subseteq S \times S_1$ is a* weak bisimulation *between (M, s) and (M_1, s_1) if:*

1. *$H(s, s_1)$;*
2. *given $v, v' \in S$ such that v is a parent of v', for all $v_1 \in S_1$ such that $H(v, v_1)$, the condition holds: (a) if v_1 is not a leaf, then there exists successor v_1' of v_1 such that $H(v', v_1')$, or (b) if v_1 is a leaf, then $H(v', v_1)$ (forth condition);*
3. *given $v_1, v_1' \in S_1$ such that v_1 is a parent of v_1', for all $v \in S$ such that $H(v, v_1)$, the condition holds: (a) if v is not a leaf, then there exists a successor v' of v such that $H(v', v_1')$, or (b) if v is a leaf, then $H(v, v_1')$ (back condition).*

Definition 2. *Let (M, s), (M_1, s_1) and (M_2, s_2) be three tree-like models, H_1 and H_2 be two weak bisimulations between (M, s) and (M_1, s_1) and between (M, s) and (M_2, s_2) respectively. We say that H_1 is* as bi-similar as *H_2, denoted by $H_1 \leq H_2$, if for all nodes $v \in S$, the following condition holds:*

1. *there exists an ancestor v' of v such that for all $v_1 \in S_1$ and $v_2 \in S_2$ satisfying $H_1(v', v_1)$ and $H_2(v', v_2)$, $\mathsf{Diff}(L(v'), L_1(v_1)) \subset \mathsf{Diff}(L(v'), L_2(v_2))$; or*
2. *for all $v_1 \in S_1$ and $v_2 \in S_2$ satisfying $H_1(v, v_1)$ and $H_2(v, v_2)$, $\mathsf{Diff}(L(v), L_1(v_1))$ $\subseteq \mathsf{Diff}(L(v), L_2(v_2))$.*

We write $H_1 < H_2$ iff $H_1 \leq H_2$ but $H_2 \not\leq H_1$.

Definition 2 specifies how we compare two weak bisimulations among three tree-like models. Intuitively, if H_1 and H_2 are two weak bisimulations between (M, s) and (M_1, s_1), and between (M, s) and (M_2, s_2) respectively, then $H_1 \leq H_2$ means that M_1 represents at least the same amount of *similar information* of M as M_2 does under H_1 and H_2 respectively.

Definition 3 (Tree-like model update). *Let ϕ be an ACTL formula and (M, s) a tree-like model such that $M \not\models \phi$. A tree-like model (M_1, s_1) is called a result of updating (M, s) with ϕ, if and only if*

1. $(M_1, s_1) \models \phi$;
2. *there is a weak bisimulation H_1 between (M, s) and (M_1, s_1) such that there does not exist another tree-like model (M_2, s_2) satisfying $(M_2, s_2) \models \phi$ and a weak bisimulation H_2 between (M, s) and (M_2, s_2) such that $H_2 < H_1$. In this case we say that (M_1, s_1) is an update result* under H_1.

Definition 3 is a declarative representation for the result from a tree-like model update. Condition 1 simply states that after the update, the resulting tree-like model should satisfy the updating formula. Condition 2 ensures that the resulting tree-like model is minimal from the original model under the weak bisimulation.

Example 1. Consider a tree-like model M as described in Fig. 2, which is a counterexample of AG$\neg x \vee$ AF$\neg y$. Then according to Definition 3, it is not hard to verify that (M_1, s_1) is a result of the update of (M, s) with AG$\neg x \vee$ AF$\neg y$, where (M_2, s_2) is not although it also satisfies AG$\neg x \vee$ AF$\neg y$.

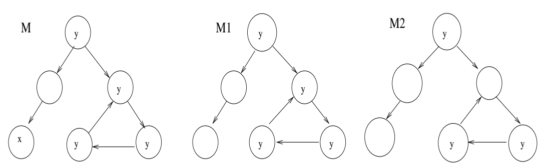


Fig. 2. Updating (M, s) with AG$\neg x \vee$ AF$\neg y$

4 Constraints and Constraint Automata

The minimal change principle for tree-like model update is purely defined based on Kripke structures, while no system constraints and other domain dependent information are considered in generating an update result. However, when we perform a model update, we may require this update not violate other specified system functions (e.g. breaking a deadlock should *not* violate a liveness in a concurrent program). Further, even if an updated model satisfies our minimal change criterion (i.e. Definition 3), it may not represent a valid result under the specific domain context. For instance, as showed in Example 1, (M_1, s_1) is a minimal updated model to satisfy formula AG$\neg x \vee$ AF$\neg y$. In practice, however, M_1 might not be acceptable if changing the variable x's value is not allowed in all states in model (M, s).

This motivates us to take relevant system constraints into account when we perform a model update. Besides logic based *system domain constraints*, which can usually be specified using ACTL (or CTL) formulas, there are some more complex constraints that are usually not expressible or difficult to be represented in the form of ACTL (or CTL) formulas. For instance, constraints related to system actions cannot be directly represented by ACTL (CTL) formulas. In the following, we study two such typical constraints: *variable constraints* and *action constraints* related to the underlying system behaviours.

Given a set of propositional variables V and a set of system actions A, we define a *Variable Constraint Automaton* constructed from V and A to be a finite deterministic automaton $\mathcal{VC}(V, A) = (S, \Sigma, \delta, q_0, F, v)$, where $S \subseteq 2^V \cup \{v\}$ is the set of states, $\Sigma = A$ is the input action symbols, $\delta : S \times \Sigma \rightarrow S$ is the total state transition function, $q_0 \in S$ is the *initial state*, $F \subseteq S$ is the set of final states, and $v \in S$ is the unique violation state.

A variable constraint automaton represents certain relations bound between a set of variables and a set of system actions. Consider two states $s_i, s_j \in S$, where s_i and s_j are not the violation state v, then a transition from s_i to s_j via action a: $\delta(s_i, a) = s_j$, indicates that by executing action a, variables' values represented by state s_i have to be changed to the corresponding variables' values represented by state s_j. Consider the variable constraint automaton depicted in Figure 3, action $i := t + 1$ ties two variables i and t, so that i's value must depend on t's value when this action is executed. On the other hand, action $i > 0?$ will not affect i and t's values, but execution of action $i < 0?$ will lead to violation state v ($*$ stands for any action symbols from Σ).

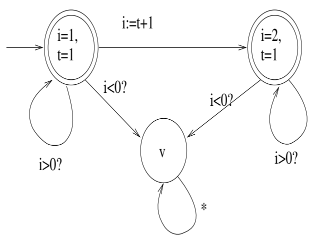


Fig. 3. A variable constraint automaton

Given a set of system actions A, we define an *Action Constraint Automaton* constructed from A to be a finite deterministic automaton $\mathcal{AC}(A) = (S, \Sigma, \delta, q_0, F, v)$, where $S \subseteq 2^A \cup \{v\}$ is the set of states, $\Sigma = \{\mathsf{preceded}, \mathsf{next}, \mathsf{exclusive}\}$ is the set of input action constraint symbols, $\delta : S \times \Sigma \rightarrow S$ is the total state transition function, q_0 is the initial state, F is the set of final states, and v is the unique violation state.

In an action constraint automaton, each state except the violation state is identified by a set of system actions. Then a transition between two states represents certain execution constraint between two specific sets of actions. For instance, if a_i and a_j are two system actions and states s_i and s_j are identified by actions $\{a_i\}$ and $\{a_j\}$ respectively, then $\delta(s_i, \mathsf{preceded}) = s_j$ means that action a_i should be executed *earlier* than action a_j, $\delta(s_i, \mathsf{next}) = s_j$ indicates that an execution of action a_i must *enforce* an immediate execution of action a_j, and $\delta(s_i, \mathsf{exclusive}) = s_j$ states that an execution of action a_i must exclude a following execution of action a_j.

As an example, Figure 4 shows that the philosopher has to pick up left fork first before he picks up right fork (similarly, here $*$ stands for any action constraint symbol from Σ). Although here we only consider three typical action constraints (i.e. preceded, next and exclusive), it is easy to extend this action constraint automaton to capture other action constraints.

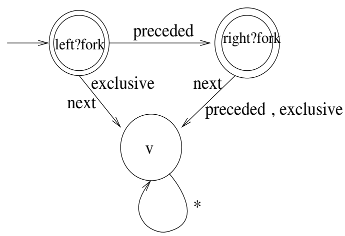


Fig. 4. An action constraint automaton

Generally from a given system, we may generate more than one variable and action constraint automata. Now we take the constraint automata into account when we perform model update. As discussed in the beginning of this section, in order to produce a more meaningful update result, our minimal change principle should be enhanced by integrating domain constraints, and system behaviour related variable and action constraints. Towards this aim, we first associate a set of system actions to a given Kripke structure. Recall that in the extent of model checking and model update, a Kripke structure actually represents the underlying system's behaviours where each state transition in the Kripke structure is caused by an execution of some system action. Therefore, for a given system, we can associate a set of system action A to the corresponding Kripke structure $M = (S, R, L)$ such that each state transition $(s, s') \in R$ is labeled by some action $a \in A$.

Definition 4. *Let $M = (S, R, L)$ be a Kripke structure, A a set of system actions associated to M, V a set of propositional variables, $\mathcal{VC}(V, A) = (S^{\mathcal{VC}}, \Sigma^{\mathcal{VC}}, \delta^{\mathcal{VC}}, q_0^{\mathcal{VC}}, F^{\mathcal{VC}}, v^{\mathcal{VC}})$ a variable constraint automaton, and $\mathcal{AC}(A) = (S^{\mathcal{AC}}, \Sigma^{\mathcal{AC}}, \delta^{\mathcal{AC}}, q_0^{\mathcal{AC}}, F^{\mathcal{AC}}, v^{\mathcal{AC}})$ an action constraint automaton. We say that M* complies to *$\mathcal{VC}$ and $\mathcal{AC}$, if the following conditions hold:*

1. *For each state transition $\delta^{\mathcal{VC}}(s_1, a) = s_2$ in $\mathcal{VC}(V, A)$ where s_1, s_2 are not the violation state, if there is a state $s \in S$ such that $s_1 \subseteq L(s)$ and $(s, s') \in R$ is labeled with a, then $s_2 \subseteq L(s')$ (i.e variable bindings through action a);*
2. *For each $\delta^{\mathcal{AC}}(s_1, \mathsf{preceded}) = s_2$ in $\mathcal{AC}(A)$, where s_1, s_2 are not the violation state, for each $a \in s_1$ and $a' \in s_2$, there exists a path $\pi = [s_0, \cdots, s_i, s_{i+1}, \cdots, s_j, s_{j+1}, \cdots]$ in M such that (s_i, s_{i+1}) is labeled with a and (s_j, s_{j+1}) is labeled with a' (i.e. a occurs earlier than a');*
3. *For each $\delta^{\mathcal{AC}}(s_1, \mathsf{next}) = s_2$ in $\mathcal{AC}(A)$, for each $a \in s_1$ and $a' \in s_2$, there exists a path $\pi = [s_0, \cdots, s_i, s_{i+1}, s_{i+2}, \cdots]$ in M such that (s_i, s_{i+1}) is labeled with a and (s_{i+1}, s_{2+2}) is labeled with a' (i.e. a' occurs next to a);*

4. *For each* $\delta^{\mathcal{AC}}(s_1, \textsf{exclusive}) = s_2$ *in* $\mathcal{AC}(A)$*, for each* $a \in s_1$ *and* $a' \in s_2$*, there does not exist a path* $\pi = [s_0, \cdots, s_i, s_{i+1}, \cdots, s_j, s_{j+1}, \cdots]$ *in* M *such that* (s_i, s_{i+1}) *is labeled with* a *and* (s_j, s_{j+1}) *is labeled with* a' *(i.e.* a*'s execution excludes* a'*'s execution).*

Given a set of domain constraints $\mathcal{C}$ (ACTL formulas) and a class of constraint automata $\Im$, we say that a tree-like model (M, s) *complies to* $\mathcal{C}$ and $\Im$ if $(M, s) \models \mathcal{C}$ and (M, s) complies to each constraint automaton in $\Im$. Now we can extend our previous tree-like model update with complying to domain constraints and constraint automata.

Definition 5 (Update complying to constraints). *Let* (M, s) *be a tree-like model,* $\mathcal{C}$ *a set of ACTL formulas specifying the domain constraints,* $\Im$ *a class of system constraint automata, and* ϕ *a satisfiable ACTL formula such that* $(M, s) \not\models \phi$*. A tree-like model* (M_1, s_1) *is called a* result of updating (M, s) with ϕ complying to $\{\mathcal{C}, \Im\}$*, iff*

1. $(M_1, s_1) \models \phi$ *and complies to* $\mathcal{C}$ *and* $\Im$*;*
2. *there is a weak bisimulation* H_1 *between* (M, s) *and* (M_1, s_1) *such that there does not exist another tree-like model* (M_2, s_2) *satisfying that* $(M_2, s_2) \models \phi$*,* (M_2, s_2) *complies to* $\mathcal{C}$ *and* $\Im$*, and a weak bisimulation* H_2 *between* (M, s) *and* (M_2, s_2) *such that* $H_2 < H_1$*.*

5 A Case Study: The Mutual Exclusion Program

Based on the approach described earlier, we have implemented a system prototype to perform the tree-like local model update. A detailed report about the algorithms and system prototype implementation is referred to our other two papers [10,12]. To test our approach, we have undertaken four major case studies on this prototype. In this section, we provide some information one particular case study - the well known mutual exclusion program, and show how our approach can help to find a proper system modification.

Consider the concurrent program encoded in SPIN in above table. The program consists of processes PA() and PB(), which share two common boolean variables x and y. To ensure mutual exclusion of the assignments to x and y, some control variables $flag$ and $turn$, are introduced. Then there are two critical sections in each process, one for the assignments to x (statement 13 and statements 45-54 in PB()), and another one for the assignments to y in PA() (statements 23 in PA() and 52 in PB(), respectively. Notice that in PB(), the critical section for y is nested into critical section for x. Each variable flagiV (i=1,2 and V=A,B) indicates the request of process PV() to enter critical section i, and turniB tells whether such a request by process PB() in case of simultaneous requests should be granted.

The specification is formalized in an ACTL formula: $\varphi =$AG$(\neg(ta \wedge (tb \vee tc))$, which describes that the program satisfies mutual exclusion for assignments to x and y, respectively. For example, PA() must not execute statement 13, if PB() executes statement 45 or 54 at the same time.

As mentioned in [3], this program contains about 10^5 states. We apply SPIN model checker to check whether this program satisfies property φ. With SPIN optimization,

Table 1. A mutual exclusion program - SPIN source code

<table>
<tr><td>1: bool flag1A, flaf2A;</td><td>38: proctype PB() {</td></tr>
<tr><td>2: bool turn1B,turn2B;</td><td>39: do</td></tr>
<tr><td>3: bool flag1B,flag2B;</td><td>40: :: flag1B=true;</td></tr>
<tr><td>4: bool x,y;</td><td>41: turn1B=false;</td></tr>
<tr><td>5: bool ta,tb,tc,td;</td><td>42: do</td></tr>
<tr><td>6: proctype PA() {</td><td>43: if</td></tr>
<tr><td>7: do</td><td>44: :: !flag1A|| turn1B →</td></tr>
<tr><td>8: :: flag1A=true;</td><td>45: atom{x=x && y; tc=true;}</td></tr>
<tr><td>9: turn1B=false;</td><td>46: tb=false;</td></tr>
<tr><td>10: do</td><td>47: flag2B=true;</td></tr>
<tr><td>11: :: if</td><td>48: turn2B=false;</td></tr>
<tr><td>12: :: !flag1B||!turn1B →</td><td>49: do</td></tr>
<tr><td>13: atom{x=x && y; ta=true;}</td><td>50: :: if</td></tr>
<tr><td>14: ta=false;</td><td>51: :: !flag2A||turn2B →</td></tr>
<tr><td>15: flag1A=false;</td><td>52: y=!y;</td></tr>
<tr><td>16: if</td><td>53: td=false;</td></tr>
<tr><td>17: :: turn1B →</td><td>54: atom{x=x||y; tb=true};</td></tr>
<tr><td>18: flag2A=true;</td><td>55: tb=false;</td></tr>
<tr><td>19: turn2B=true;</td><td>56: flag2B=false;</td></tr>
<tr><td>20: do</td><td>57: flag1B=false;</td></tr>
<tr><td>21: ::if</td><td>58: break;</td></tr>
<tr><td>22: ::!flag2B||!turn2B →</td><td>59: :: else;</td></tr>
<tr><td>23: y=false;</td><td>60: fi;</td></tr>
<tr><td>24: tc=false;</td><td>61: od;</td></tr>
<tr><td>25: flag2A=false;</td><td>62: break;</td></tr>
<tr><td>26: break;</td><td>63: :: else;</td></tr>
<tr><td>27: ::else;</td><td>64: fi:</td></tr>
<tr><td>28: fi;</td><td>65: od;</td></tr>
<tr><td>29: od;</td><td>66: od;</td></tr>
<tr><td>30: ::else;</td><td>67: }</td></tr>
<tr><td>31: fi;</td><td></td></tr>
<tr><td>32: break;</td><td>68: int{</td></tr>
<tr><td>33: ::else;</td><td>69: run PA();</td></tr>
<tr><td>34: fi;</td><td>70: run PB();</td></tr>
<tr><td>35: od;</td><td>71: }</td></tr>
<tr><td>36: od;</td><td></td></tr>
<tr><td>37: }</td><td></td></tr>
</table>

the program still contains 1800 states during the checking process. After SPIN model checking, it reports that property φ does not hold for this program and returns a tree-like counterexample (in fact it is linear), which only contains 22 states, as shown in Fig. 5.

Now we consider to update this counterexample by using our approach. First, we construct relevant constraint automata for this program, as discussed in section 4. In this case study, the particular variable constraint automaton shown in Fig. 6 will be directly embedded into the update process. In Fig. 6, "$*1$" indicates any statement in process PA() except statement $A : x = x\&\&y$, and "$*2$" indicates any statement in

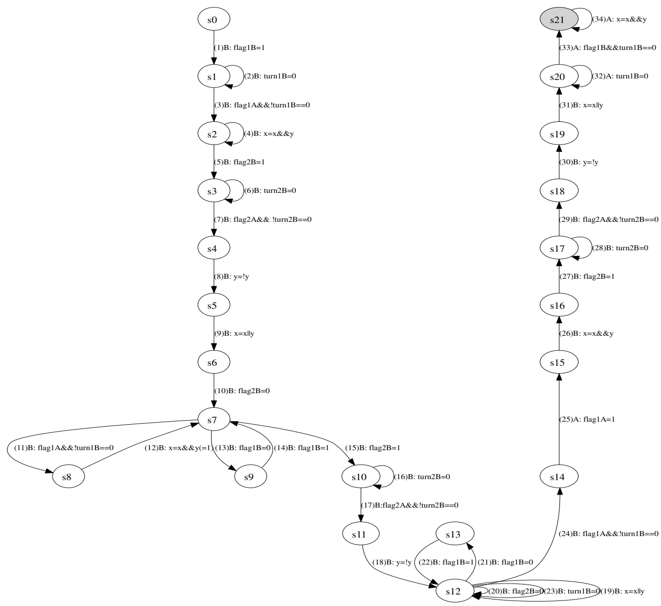


Fig. 5. A counterexample for AG$(\neg(ta \wedge (tb \vee tc))$

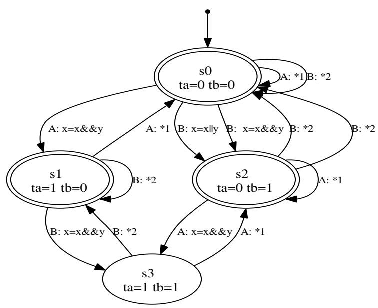


Fig. 6. The variable constraint automaton for ta and tb

process PB() except statements $B : x = x\&\&y$ and $B : x = x||y$. Intuitively, this variable constraint automaton represents the constraints between variables ta and tb with respect to various actions (statements) in the program.

Considering the counterexample shown in Fig. 5, we observe that state $s21$ violates property φ where $L(s21) = \{ta = 1, tb = 1, tc = 0, flag1A = 1, ...\}$. Note that the transition from $s20$ to $s21$ in Fig 5 corresponds to the automaton states $s2$ and $s3$ in Fig. 6 respectively. By applying our tree-like local model update with the associated constraint automaton (Definition 6), the counterexample will be minimally updated to satisfy φ: state $s21$ will be updated to either (1) $s21^{'}$: $L(s21^{'}) = \{ta = 0, tb = 1, tc = 0, flag1A = 1, ...\}$, or (2) $s21^{''}$: $L(s21^{''}) = \{ta = 1, tb = 0, tc = 0, flag1A = 1, ...\}$, while all other states in the local model will remain unchanged.

This update suggests that one possible modification for the original program (see Table 1) is to change statement 9 in PA() from “turn1B=false;” to “turn1B=true;”. A further SPIN model checking for the revised mutual exclusion program will confirm that this result from such local model update is a final correction to the original program.

6 Conclusion

In this paper we have developed an approach for ACTL tree-like local model update. In order to effectively generate the update result, we have proposed a minimal change principle based on weak bisimulation on tree-like structures, defined domain dependent constraint automata, and integrated them under a unified update formulation.

References

1. Berard, B., et al.: System and Software Verification: Model-Checking techniques and tools. Springer, Heidelberg (2001)
2. Buccafurri, F., Eiter, T., Gottlob, G., Leone, N.: On ACTL formulas having linear counterexamples. Journal of Computer and System Sciences 62, 463–515 (2001)
3. Buccafurri, F., Eiter, T., Gottlob, G., Leone, N.: Enhancing model checking in verification by AI techniques. Artificial Intelligence 112, 57–104 (1999)
4. Clarke Jr., E., Jha, S., Lu, Y., Veith, H.: Tree-like counterexamples in model checking. In: Proceedings of the 17th Annual IEEE Symposium on Logic in Computer Science (LICS 2002), pp. 19–29 (2002)
5. Clarke Jr., E., Grumberg, O., Jha, S., Liu, Y., Veith, H.: Counterexample-guided abstraction refinement for symbolic model checking. Journal of ACM 50, 752–794 (2003)
6. Fraser, G., Wotawa, F.: Nondeterministic testing with linear model-checking counterexamples. In: Proceedings of the 7th Intel. Conf on Quality Software, QSIC 2007 (2007)
7. Harris, H., Ryan, M.: Theoretical foundations of updating systems. In: Proceedings of the 18th IEEE International Conference on Automated Software Engineering, pp. 291–298 (2003)
8. Huth, M., Ryan, M.: Logic in Computer Science: Modelling and Reasoning about Systems, 2nd edn. Cambridge University Press, Cambridge (2004)
9. Jobstmann, B., Staber, S., Griesmayer, A., Bloem, R.: Finding and fixing faults. Journal of Computer and System Sciences (2010) (to appear)
10. Kelly, M., Zhang, Y., Zhou, Y.: Local model update with an application to sliding window protocol (2010) (under review)
11. Zhang, Y., Ding, Y.: CTL model update for system modifications. Journal of Artificial Intelligence Research 31, 113–155 (2008)
12. Zhang, Y., Kelly, M., Zhou, Y.: Foundations of tree-like local model updates (2010) (under review)

Bridging Multiple Motor-Skills in a Community Site

Kenji Matsuura[1], Naka Gotoda[2], Tetsushi Ueta[1], and Yoneo Yano[2]

[1] Center for Advanced Information Technology, The University of Tokushima, Minamijosanjima 2-1, Tokushima, 770-8506, Japan
{matsuura,tetsushi}@ait.tokushima-u.ac.jp
[2] Faculty of Engineering, The University of Tokushima, Minamijosanjima 2-1, Tokushima, 770-8506, Japan
{gotoda,yano}@is.tokushima-u.ac.jp

Abstract. This paper touches upon the design issue for multiple skill development of human beings in an online community environment. It focuses on motor skills such as rope-skipping, running, swimming and so forth in the physical world. The approach attempts to adopt social networking service for mutual contribution on that purpose as the base of online communication. In order to treat various sorts of skill, the web-based system allows community authors to handle several types of media for storing records of skill training of individuals and representing them in the system. This study developed such an environment and made trial use for the first stage.

Keywords: motor skill, skill development, online community, media type, SNS.

1 Introduction

With rapid innovation of ubiquitous technologies, the academic community of learning science or technology-enhanced learning is about to extend the research target from traditional intelligence and knowledge to physical intelligence and domain. Hollan et al. [1] surveyed the shift of working and researching style with computer-mediated interaction. Sharing knowledge is a core method to circulating best practices among members in a community of interest [2]. Fischer et al. [3] also reported the oretical framework of socio-cultural approach for learning in a context of collaborative settings. The theory on communities of practice (Cop) is one of the hottest topics in computer supported collaborative learning. Technology stewarding for communities is well organized and summarized in “Digital Habitats” by Wenger et al [4]. Many community-based learning styles are discussed in the context of technology oriented. We believe that verbal communication with other people on line is actually a key factor for collaborative learning but adopting other media for nonverbal things enriches learning and development of physical skills.

Today’s technologies dealing in learning science are not only networked PCs but also other devices such as sensor, touch-screen, cell-phones, and whatever available for learning and communication. They made us possible to monitor physical activities

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 145–152, 2010.

at real time or asynchronously. The concrete domain of this study is sports including multiple types of repeating actions such as jumping, running and so forth. We designed the supporting environment with various kinds of media in a social networking service for encouraging communication with other members. In this environment, learners can get opportunities to know skills of other learners beyond the different categories of sports. As a result, we expect good influence for users to accelerate the development of standing skill while they would get new opportunities of challenging unknown skills.

2 Sports Skill Development

2.1 Discussion on Sports Skill

Researches on motor skill have the long history as one of difficult domains [5]. According to well-known taxonomy for motor skill by Gentile [6], human skill could be divided into two-dimensional space by the axis of environmental-context and action function. There are still four sub spaces in each space by classifying regulatory conditions, intertribal variability, body orientation, and manipulation.

Lots of research approaches for this domain are mainly based on the analytical view. If successful or failure performance of a certain action is captured by sensors, movies and so forth, researchers try to analyze the reasons why they could or could not. The analysis is carried out quantitatively and qualitatively. However, the analyzed reasons in some cases would be difficult to suggest improvement for the next time. It is because either environmental conditions or physical conditions will change from those in the past.

2.2 Closed Skill and Open Skill

Another presumable discussion for motor skill arose by Allard and Starkes about the difference between "Open skill" and "Closed skill" [7]. Closed skill assumes predictable conditions in a stable environment. Once a performer achieved the skill, s/he can do it again in the same situation. On the other hand, open skill cannot assume such predictions. Regarding open skills, a player tries something in facing different objects, conditions, and environmental factors at every trial time. Since the environmental factors, i.e. weather conditions, always change, most of field sports or trainings are open skill. Therefore it is necessary to acquire an ability to accept environmental change or potential change even if the physical or mental conditions are different each time.

We believe that it is better to know or to train various kinds of skill than mastering only single skill. However, from the design perspective of supporting systems for skills, most systems are based on a domain dependent concept. For example, the system designer usually setup specialized function to control the skill development based on a certain model. Typical examples are seen in music and cooking. In this study, we try to design the supporting system that aims at community-based training for players of multiple sports. For instance, to deal with swimming and running as targets, some analyzing parameters in the system are almost the same, "distance", "speed" and "time". The difference for them is only the scale or range. Therefore, we believe we can handle the parameters in a same way in the same system.

2.3 Media Type and the System

Human activities in a training can be monitored and stored in a various way. For example, some skills are available to be monitored and represented in a video and/or sensors while others use text if they can be represent in a verbal way. We designed an authoring environment where each community author of system users can freely customize the training record space for a skill by way of combining some modules. An example is "rope skipping", which can integrate video and text media for inputting and representing.

If a potential user is interested in a community room, s/he can register oneself to this room. Then, as a community member, s/he stores her/his daily exercising records to this room. In such a situation, we have to take into account of input and output for each module. Though current policy about the input/output of a module from designing viewpoint is one by one, we will extend the variety of combining input and output for improving flexibility and scalability.

2.4 Process of a Skill Development

Fig.1 indicates a human process of a skill development that is assumed in this study. If learners want to acquire a skill listed in a community space, they can join the community to recognize their stage of current skill. The recognition is carried out in both ways; the self-recognition by capturing their action with a video, sensors or text as a meta-recognition, or the detection of a current status for the skill in a comparison among community articles. Then, they train themselves with recording something in a community space. During the training, they make trial and error about the skill including off-line activities. After its process, they can evaluate their article by own to decide the next action whether continuing training or completing it. If they satisfy their skill represented in an article, they can complete and proceed next skill or keep training according their will. However, if not, they should go back to the recognition process in this figure.

We designed supporting functions based on this assumed process. The first one is a description space to make members know the target skill, methodology in a practice, criterion for achievement, evaluation and so forth before enrolling to the community. The next one is a view space of an article in a comfortable way. The recognition and self-evaluation process need this function. The third one is flexible input method

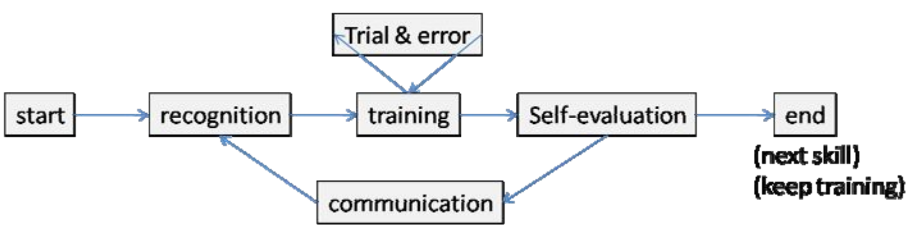


Fig. 1. General process of a motor-skill development

combining various media. During training, learners sometimes input their data of practice to the system. The forth one is communication that enables learners to interact each other by text or multimedia to know the better way for a skill if needed. At last, as an optional function, the automatic analysis function will support the skill-development activities with this system. However, it will be domain dependent technology is necessary. Therefore, we made a few functions for this purpose at this studying stage.

3 System Development

3.1 Framework

Based on the principal discussed in the former section, we developed the prototype that enabled users to create a new space for a community of a sport. The architecture of the system is illustrated in Fig.2. The prototype was developed on "OpenPNE" platform (http://www.openpne.jp/) that is an open source application of "symphony" PHP framework (http://www.symfony-project.org/). The description of symphony in the official website is "*Symfony is a full-stack framework, a library of cohesive classes written in PHP.*" The system is a kind of social-networking service, sometimes called "social software" on a web.

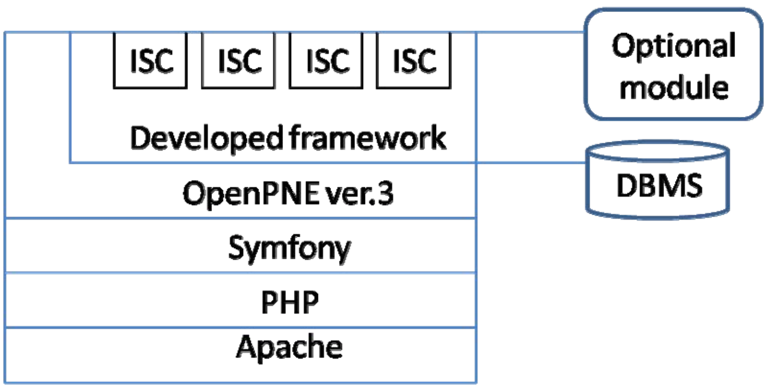


Fig. 2. System architecture

The prototype provides some templates for each module that is, for example, video module for inputting and showing in an embedded html. "ISC" in the figure is an "Implemented item of a Skill Community". If an author creates a community space in the system, it means a new ISC is created. The prototype uses database to store platform data for OpenPNE and the original data altogether. If an advanced module was developed for analysis of an article and so forth, it is called from outside. For instance, an image processing module for movie data is a typical module that should be separated from the server as it needs high performance. In this way, the prototype has several modules and the organization is flexible from a technical viewpoint.

3.2 Authoring Environment

The target user of the authoring environment is a community owner who intends to open the specific space for the community of interest [8]. The system provides an interactive interface for the user to select and combine some modules into a unified space. For example, Fig.3 is a snapshot of third phase of configuring the space. In this case, the user creates an environment that presents members “text area” for a title, “text area” for description of blog-body, “point” for the number of repeating actions.

The drag and drop operation by mouse is available at this phase. If a user finds requisite elements on the right of the interface, s/he can move them to the left. The order of embedding position for each module is configurable freely. However, members have to input something to each module assigned to a motor skill for the community theme. We have developed six types of module for the prototype so far; 1)Text, 2)Score, 3)Image, 4)Video, 5)HRM(Heart Rate Monitor), 6)GPS(Global Positioning System).

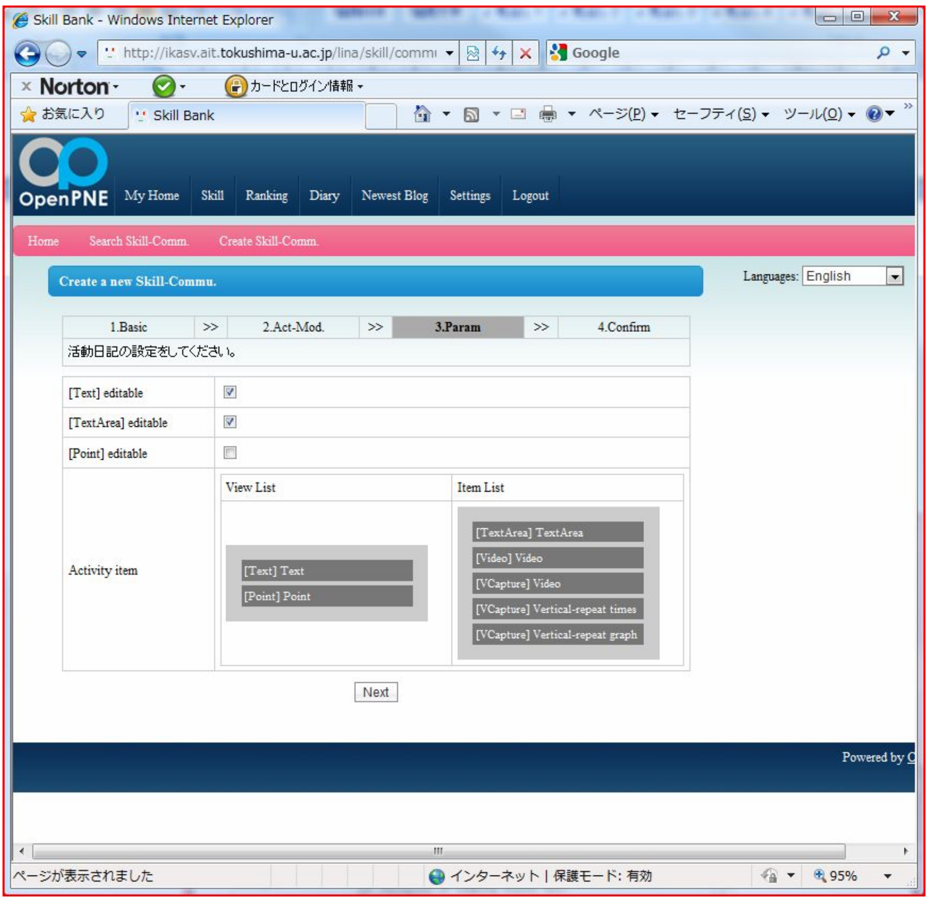


Fig. 3. Snapshot of authoring a community space

3.3 Displaying Environment

Fig.4 is a snapshot of a user interface for a community space that aims at sharing information for developing rope-skipping skill. It is one of several communities created on the same SNS platform. It combines several modules such as “text” and “video”. In addition, as a pluggable module for optional function, a video processing technique is installed in this case. The purpose of integration of image processing technology is to detect wave motion of gross motor skill. Wave motion is useful to count actions. It is also available to compare the stability between each trial time. Stability is originally defined by a calculation using variance of local-max and local-min. Learners in a state of plateau may need good examples of skill images in such an easy-to-understanding way.

Fig. 4. A snapshot of a rope-skipping community

In this example community, the displaying frame embedded the video that have been converted from any kinds of conceivable media to Flash-video because this media seems comfortable for many client types. In the Flash-video, the jumping wave form is over-wrapped onto the captured image. It was useful for community members to understand the stability and the counting adequacy by the system.

4 Trial Use at Current Stage

4.1 Organization

Twenty students of our university have used this system for four weeks. Since the system was a prototype, we attempted to investigate how many skill-communities the system allowed community members to create at first. We also wanted to survey the number of combination patterns of module they can create.

4.2 Some Findings

Subjects created 28 different communities for motor skills such as jogging, playing pencils, Hula-Hoop, football juggling. These communities were created by fifteen different owners that means 75 % of members could done individually. Only in a few weeks trial term, total 82 articles were created in these communities. Therefore, three articles were included in each community space in average.

In terms of module combination, there were twelve patterns we found through this trial. For instance, “text”, “text area”, “score” and “video” were combined at the community of Hula-Hoop. Some skill communities shared the same combination pattern of module even though the target skill was different. It implicates that there is a potential extension to make the linkage between these different skill-communities. For example, jogging community and swimming community are completely different each other on skill category perspective. However, the combination of media types and the parameters that are dealt with them, i.e. distance, time and speed, can be the same in the system. The community member in each space may mutually contribute in order to keep their motivation highly.

5 Summary

This paper described the design issue of an SNS system that promotes skill-development for various kinds of motor skill. From a technical viewpoint, multiple input and output methods are integrated into the unified SNS so as to combine them in a flexible way for a motor skill. This framework offers additional possibilities to support linking relevant skills due to the integrated media types, analyzing method, and human decision. The system helps learners to transfer skills among community-members, to understand a skill deeply, and to discover new skills or articles relevant to the history of their own training.

We conducted the trial use though the research project is in an ongoing stage. Subjects were pleased to create various kinds of skill community. We found some interesting implications for the future work. However, we have to go further investigations on motivation of users. We also believe the principal subject on detecting methodology for similar skill-community and sometimes contrasting ones will be interesting topic for another studying stream.

Acknowledgments

This work was partially supported by Grant-in-Aid for Young Scientists (B) No.20700641 and Scientific Research (A) No.19200055. In terms of system development, an alumnus of our laboratory, Mr. Toyoaki Nabeshima, made a profound contribution to this work.

References

1. Higgins, S.: Motor Skill Acquisition. Physical Therapy 71, 123–139 (1991)
2. Bansler, J.P., Havn, E.C.: Sharing best practices: An Empirical Study of IT-Support for Knowledge Sharing. In: Proceedings of The 9th European Conference on Information Systems, pp. 653–664 (2001)
3. Fischer, G., Rohde, M., Wulf, V.: Community-based learning: The core competency of residential, research-based universities. International Journal of Computer-Supported Collaborative Learning 2(1), 9–40 (2007)
4. Wenger, E., White, N., Smith, J.D.: Digital Habitats, CPsquare (2009)
5. Gentile, A.M.: Skill acquisition: Action, movement, and neuromotor processes. In: Carr, J.H., Shepherd, R.B. (eds.) Movement Science: Foundations for Physical Therapy in Rehabilitation, pp. 111–187 (2000)
6. Allard, F., Starkes, J.L.: Motor-skill experts in sports, dance, and other domains. In: Ericsson, K.A., Smith, J. (eds.) Toward a General Theory of Expertise, pp. 126–150. Cambridge University Press, Cambridge (1991)
7. Hollan, E., Hutchings, E., Kirsh, D.: Distributed cognition: toward a new foundation for human-computer interaction research. ACM Transactions on Computer-Human Interaction 7(2), 174–196 (2000)
8. Fischer, G.: Communities of Interest: Learning through the Interaction of Multiple Knowledge Systems. In: Proceedings of the 24th IRIS Conference, pp. 1–14 (2001)

Topic Visualization for Understanding Research Paper in Collaborative Discussion

Masato Aoki, Yuki Hayashi, Tomoko Kojiri, and Toyohide Watanabe

Graduate School of Information Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464-8603, Japan
{maoki,yhayashi,kojiri,watanabe}@watanabe.ss.is.nagoya-u.ac.jp

Abstract. Our objective is to construct a collaborative discussion support system for understanding the research paper in the research group by visualizing the contents of the discussion. When reading a research paper, it is essential not only to understand the contents of the paper but also to obtain the related knowledge. Our system evaluates topics in discussion from the viewpoints of similarity between a topic and the paper and that between topics. Then, the system visualizes topics as circles around chapter nodes whose distances from their target chapter nodes and other topic nodes represent the similarities. By observing the visualized topics, to derive topics from various perspectives is encouraged.

Keywords: understanding research paper collaborative discussion, discussion visualization.

1 Introduction

When we read the related research papers, it is essential not only to understand the contents of the paper but also to obtain the related knowledge. However, each of us has limited knowledge and may not be able to obtain enough knowledge by himself/herself. One solution for this problem is to discuss the contents of the paper in a group to obtain various knowledge. Through discussion with others, related knowledge that other participants have can be acquired. Moreover, knowledge of different perspectives may be derived through the discussion. There are many researches for supporting the discussion in a distributed environment. By these researches, participants are able to communicate with each other without limitations of time and location. Therefore, this research focuses on a collaborative discussion in a research group for obtaining related knowledge of the research paper in a distributed environment. In order to appropriately assess the paper, participants should discuss various research stages in the paper such as the background, objective and method. However, participants cannot always discuss the paper effectively because they sometimes focus on the topics from limited perspective. If they can notice the situation of discussion timely, they can grasp a diversity of discussed topics. That is, if they discuss the paper only from limited perspectives, they may be able to derive new topics from different

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 153–162, 2010.

perspectives. Our objective is to construct a collaborative discussion support system for promoting the effective discussion by visualizing the diversity of the discussed topics. In our research, topics are defined as groups of utterances whose target is a certain part of the paper, such as chapter, section and words.

For promoting the reflection of the finished discussion, there are many researches which visualize the structure of the discussion after the discussion has been finished. Conklin et al. constructed a discussion support system for detecting inconsistencies and less discussed topics by representing the relationship between utterances for problem-solving[1]. The system divides participants' utterances into four types(Issue, Position, Argument, Other) and represents them by nodes with different attributes. In addition, the relations between utterances are represented by labeled links, such as *generalizes*, *specializes* and *replaces*. In order to organize effective discussion, available relations are defined according to the type of the target node. When participants make utterances, they need to input the types of utterances and the relations to the target utterances. Therefore, participants cannot conduct free discussion since their utterances are controlled by the types of target utterances.

Kojiri et al. proposed a system which visualizes the structure of an ongoing discussion[2]. In order to extract and present important utterances in topics, the number of utterances of the topics and the uttered time are considered. This system can indicate the current important utterances, but cannot promote an effective discussion. Instead of extracting only specific utterances of the topics, it is necessary to evaluate the discussion by considering all the utterances.

Some researches try to activate the discussion by showing activeness of participants. Viegas et al. developed the system called Chat Circles which expresses the utterances of each participant in resizable circles[3]. Active participants are displayed as large circles since the sizes of the circles become smaller depending on the time to say nothing. In addition, the system has a function for browsing the chat log by arranged bars in chronological order whose lengths vary depending on the number of words of each utterance. Erickson et al. proposed the visualization method of representing activeness of discussion by positions of circles which correspond to participants[4]. Each circle is placed on a common circle which shows their workspaces and the center of the circle means high activity level. Xiong et al. constructed the system which represents the activeness of each participant by using the metaphor of a flower[6]. In order to distinguish individual participants and to comprehend the overall interaction, each participant is represented as a flower whose length of the stem expresses the length of his/her login time of each participant. Moreover, the petals of each participant represent proposed topics, and the responses from other participants are displayed as small circles at the distal end of the target topic. Tat et al. focused on the social interaction centered around one participant, and supported self exploration of one's own chat history[5]. They constructed the system to display each chat history of the participants as a line of circles(utterances) in the direction corresponding to the participants. For the purpose of understanding conversational tones, this system estimates the emotions of each participant by used emoticons and represents the emotions by colors of translucent planes in the directions corresponding to the participants. In addition, this system changes the color strengths of the circles according to the number of characters of utterances. The time filtering function allows participants to filter conversations by selecting a certain time and participants can understand the activeness of the discussion at the time. However, in these researches, participants cannot always grasp the goodness of the discussion which indicates varieties of perspectives derived in their discussion. The quality of the discussion is not represented only by active participants and active time. In order to facilitate effective discussion, the goodness of the discussion should be visualized in the interface.

In this research, topics are evaluated and visualized from the perspective of obtaining related knowledge for the purpose of promoting discussion. To obtain related knowledge is to acquire opinions from various perspectives for each phase of the research. In order to obtain additional related knowledge, it is desirable to discuss various topics from diverse perspectives. Thus, every topic in the discussion is evaluated from the viewpoint of similarity between the topic and the paper and that between topics. By organizing topics around a chapter, participants can understand the situation of the discussion intuitively and are encouraged to discuss diverse topics voluntarily.

2 Approach

The target paper of discussion is research paper in engineering. We assume that the participants of discussion are researchers who are interested in the target research field of the paper.

A discussion is generally classified into *creative discussion* and *problem-solving discussion*. In *creative discussion*, participants do not always have a clear goal, but intend to find various perspectives regarding the discussion theme. In *problem-solving discussion*, participants aim to reach a specific goal. The discussion for obtaining the related knowledge of the paper is *creative discussion* since a clear goal does not exist. Moreover, *creative discussion* is classified into *focused discussion* and *global discussion* depending on the location of the paper which participants need to acquire knowledge from. The purpose of *focused discussion* is the deep understanding of certain parts of the paper such as a technology and the assumed environment in the paper. In *global discussion*, participants desire to acquire the opinions about any part of the paper from various perspectives. Since each chapter of the paper corresponds to each stage of the research, to obtain comprehensive knowledge of the paper leads participants to consider various aspects of the paper such as background, objective, solution and evaluation. Therefore, in this research, we aim to support *creative* and *global discussion*.

Since each chapter of the paper is individually made up of the meaningful research aspects, it is important to discuss all the chapters. In addition, if there are many topics from the same perspective, participants evaluate the paper based on only limited viewpoint. Therefore, various topics should be discussed. In *creative discussion*, developed topics which are associated with the paper are required.

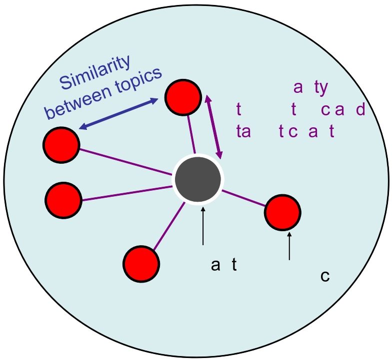


Fig. 1. Concept of visualization

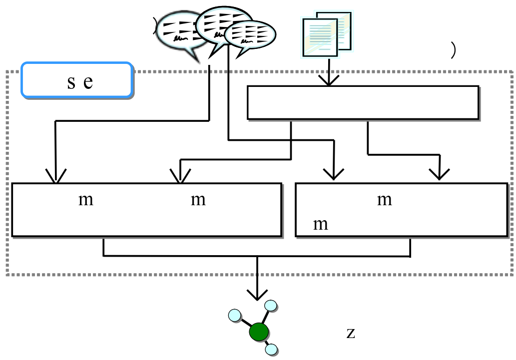


Fig. 2. Processing steps

Topics such as chitchats which do not directly relate to the contents of papers do not contain the information to evaluate the papers. Thus, in this research, every topic in the discussion is evaluated from viewpoint of the similarity between a topic and its target chapter and that between topics.

Figure 1 shows the concept of visualizing topics in the discussion. The visualization should make participants aware of the situation of the discussion. In our system, the topics which are collections of utterances are extracted and placed as “topic node” around “chapter node” which indicates the contents of the target chapter. The similarities between topics and chapters are calculated based on keywords of topics in the paper. The similarity between a topic and its target chapter is represented by the distance between a topic node and a chapter node. The similarity among topics is represented by the distance among topic nodes. Participants are able to be aware of the situation of their discussion by distribution of topic nodes around the chapter nodes. If many topic nodes exist near the chapter node or in the same direction, participants are urged to derive developed topics from other perspectives. In addition, the chapter which lacks the discussion is able to be grasped by the distribution of the topic nodes for each chapter.

Figure 2 shows the processing steps for visualizing the topics. Currently, we focus on the text-based discussion using the chat. The system gets the keywords of chapters in advance from the texts of each chapter. The similarity between a topic and its target chapter is calculated based on the keywords contained in the topic. The similarity between topics is regarded as the difference between the target locations in the chapter. Sentences represent the context of the paper. Similar or related sentences are written in near locations. Therefore, the similarity between topics is calculated by the distances between target locations. The system determines the locations of the topics and displays them based on the calculated degrees of the similarity between a topic and its target chapter and the target location of the topics.

3 Topic Visualization Method

3.1 Extraction of Keywords in Chapter

The chapter contents can be expressed as a set of keywords. In this research, keywords are defined as characteristic words(noun) in each chapter. Such words appear frequently in the chapter and do not appear in a whole paper. In order to acquire keywords for each chapter, important degrees of each word for the chapters are calculated by Equation 1. $value(c, a)$ represents the importance degree of word a in chapter c. $count(c, a)$ is the number of word a which appears in chapter c and $N(c)$ is the total number of words in chapter c. The important degree of the word becomes large if the word is used frequently in the chapter and becomes small if the word appears in the whole paper. Our system extracts words whose important degrees are larger than the threshold as keywords.

$$value(c, a) = \frac{count(c, a)}{N(c)} \times \log(\frac{\sum_i N(i)}{\sum_i count(i, a)}) \tag{1}$$

3.2 Expression of Similarity between Topic and Chapter

It is desirable that the topic which has strong connection to contents of the target chapter is placed close to the chapter node. The distance between a topic node and the chapter node is defined as Equation 2. $distance(c, t)$ represents the distance of a topic t from target chapter c. $\sum_i value(c, i)$ is the sum of important degrees of all keywords in chapter c. $relation(c, t)$ is the degree of the similarity between topic t and its target chapter c. According to the equation, the distance of the topic which has a large degree of the similarity becomes small as shown in Figure 3.

The similarity between a topic and its target chapter is expressed by the ratio of keywords included in a topic. Thus, $relation(c, t)$ is defined as Equation 3. $W(t)$ is the total number of words in topic t. $chaptIn(c, t)$ is the number of keywords of target chapter c contained in topic t. α is a constant number that coordinates the effect of $\sum_{i \in t \cap c} value(c, i)$ and the value takes from 0 to $\sum_i value(c, i)$. As α gets larger, the effect of $chaptIn(c, t)$ becomes greater. Based on the equation, the degree of the topic that includes large number of important keywords of the chapter becomes large.

$$distance(c, t) = \sum_i value(c, i) - relation(c, t) \tag{2}$$

$$relation(c, t) = \frac{chaptIn(c, t)}{W(t)} \times (\alpha + \sum_{i \in t \cap c} value(c, i)) \tag{3}$$

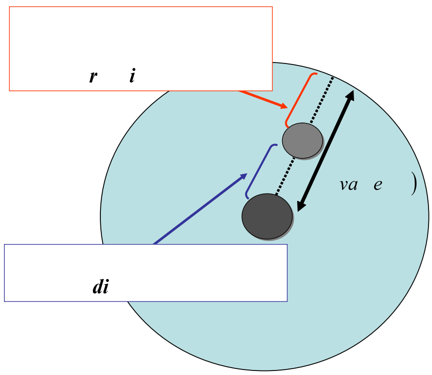


Fig. 3. Expression of similarity between chapter and topic

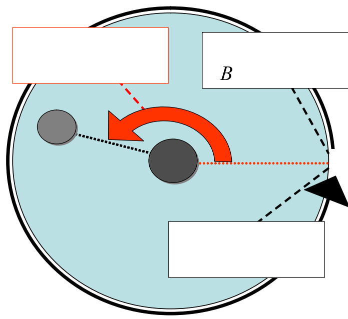


Fig. 4. Expression of target location

3.3 Expression of Similarity between Topics

The similarity between topics can be determined by the target locations in the chapter. We define the degree of the similarity between topics as the distance between the target locations of topics. The target location of a topic can be grasped by keywords of the chapter that it contains and is calculated by Equation 4. $position(c,t)$ represents the target location of topic t in target chapter c and its value takes from 0(beginning of the chapter) to 1(end of the chapter). $loc(c,i)$ indicates the appearance position of keyword i in chapter c. If keyword i appears in multiple locations, $loc(c,i)$ is set as the middle point of the appearance position of the keyword i. The target location of the topic is represented as the average of the appearance positions of all emerging keywords.

The angle of the topic node is determined according to the $position(c,t)$. The beginning of each chapter corresponds to 0° and the end of the chapter is 360° around the chapter node. The angle of the topic node is calculated by Equation 5. $angle(c,t)$ is an angle of topic t around chapter c. An angle is determined by multiplying 360° by a target location $position(c,t)$ as shown in Figure 4. Based on this expression, the beginning and the end of a chapter are placed near. It is often observed that main theme of the chapter is insisted in the first sentence and is summarized in the end. Thus, this expression is valid to some extent.

$$position(c,t) = \frac{\frac{1}{chaprIn(c,t)} \times \sum_{i \in t \cap c} loc(c,i)}{N(c)} \tag{4}$$

$$angle(c,t) = 360° \times position(c,t) \tag{5}$$

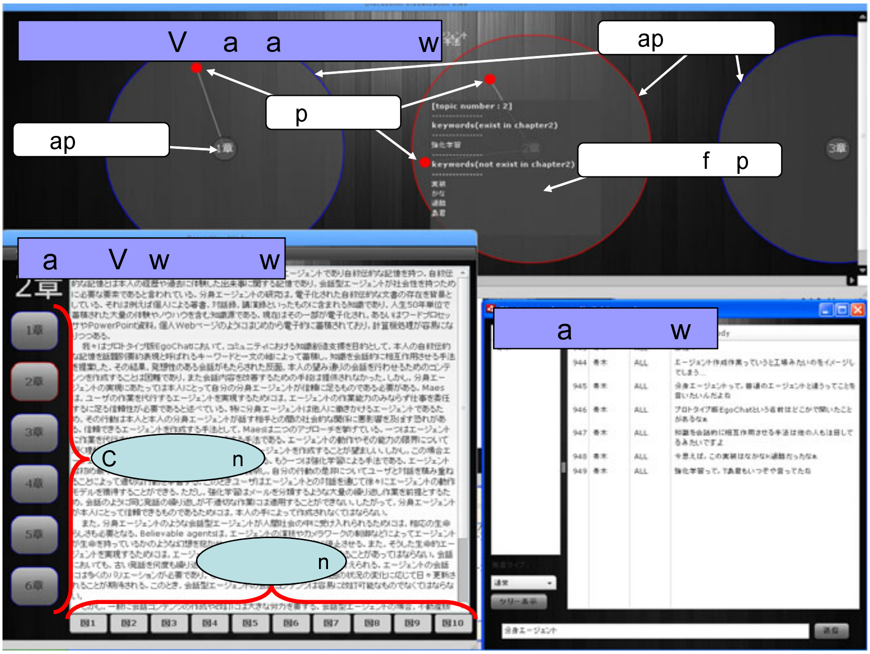


Fig. 5. Interface of prototype system

4 Prototype System

We have constructed a prototype system by embedding paper view mechanism and discussion visualization mechanism in the collaborative learning support system in our laboratory[7]. The interface of this system consists of three windows as shown in Figure 5. Participants make utterances using the chat window. Paper view window displays the contents of the paper to all participants. This window provides the same contents to all participants. By selecting the chapter button, the contents of the selected chapter appear in the windows. By selecting the figure button, the figure in the chapter is displayed in a separate window. When a chapter button is pushed, our system regards that the topic is changed and gets the utterances that compose the topic. Moreover, these utterances are analyzed by our visualization method and the result is sent to discussion visualization window.

In discussion visualization window, each chapter is viewed as a circle with its chapter node exists in the center of the circle. Topic nodes are represented by red circles around the chapter node. Information of words within the topic are shown by moving the mouse cursor over a topic node(Figure 6). In addition, the utterances of the topic are displayed by clicking a topic node(Figure 7). When a participant clicks the circumference of the chapter circles, the diverse keywords of the chapter are displayed whose angles correspond to the emerging location

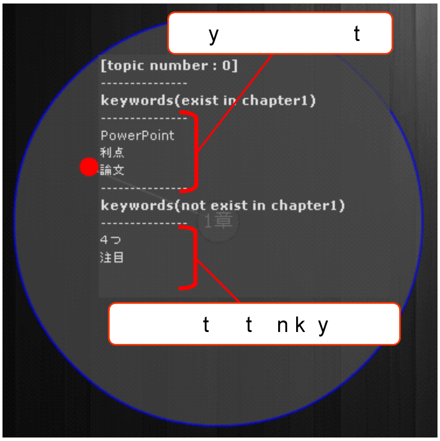


Fig. 6. Information of words in topic

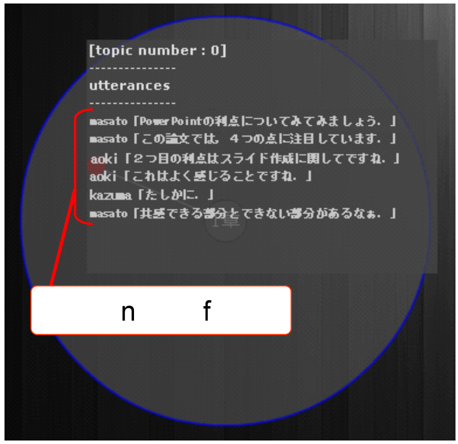


Fig. 7. Utterances of topic

in the chapter. Discussion of the specific keyword may be enforced by observing such keywords. By clicking the displayed keyword, the keyword is posted in the input area of the chat window.

5 Experiment

5.1 Experimental Setting

We evaluated the adequacy and effectiveness of the visualization method using the prototype system. In this experiment, maximum length of the distance between a chapter node and a topic node was normalized to 100, and α was set to 50.

Two groups(A and B) of four students in our laboratory were asked to discuss the research paper which consists of six chapters. The purpose of the discussion was to get the related knowledge of the paper from various perspectives. Each examinee was asked to read the paper and understand the contents in advance. If they wanted to check the contents during the discussion, they were asked to view their own papers and not to use the paper view window. The paper view window was used only for changing the target of the discussion topics. One examinee of each group was asked to determine the end of the discussion. In both groups, the discussion continued for more than one hour.

After the discussion, the examinees answered a questionnaire about the similarity between a topic and the chapter and that between topics. In addition, the examinees were asked to observe the discussion record of the other group that were organized for each topics. In order to evaluate the validity of calculated degrees of the similarity between a topic and its target chapter, the examinees were asked to divide the topics of the other group according to the relation to the chapter, slightly related ones and greatly related ones. For evaluating validity of calculated degrees of the similarity between topics, they were also asked to select the combination of similar topics of the other group with reasons.

5.2 Experimental Result

Table 1 shows the results of the average distances of greatly related or slightly related topics to the chapter calculated by the system. The average distance of all topics in group A is 73.14, and that in group B is 71.64. For both groups, the average distances of slightly related topics are greater than those of greatly related ones. Therefore, the system adequately expresses the similarity between a topic and the target chapter.

The average angles between similar topics are shown in Table 2. The average angle between all topics in group A is 90.62°, and that in group B is 70.59°. For both groups, the average angles between similar topics are smaller than that of all topics. Therefore, the system is able to place similar topics in the near location.

The result of the questionnaire about topic visualization method is shown in Table 3. For each question, 1 is the worst and 5 represents the best. In questions about the adequacy and effectiveness of visualizing topics*(a,b,c)*, answers were good. Therefore, the topic visualization method is appropriate for understanding the situation of the discussion. However, the results of questions about the effectiveness for triggering a new topic*(e,d)* indicate that the visualization did not lead to specific topics. Examinees commented that discussion topic changes according to the context, so it is difficult to change topics according to the keywords in the circumference of the circle. Therefore, the method for guiding a discussion topic needs to be considered so as to reflect the context of the discussion.

Table 1. Average distances between topic and target chapter

	Group A	Group B
Topics slightly related to chapter	83.40	77.05
Topics greatly related to chapter	66.32	69.54

Table 2. Average angles between topics

	Group A	Group B
Similar topics	56.97°	38.02°

Table 3. Average scores of each question

	Group A	Group B
a. Adequacy of distance between topic and chapter	3	3.5
b. Adequacy of distance between topics	3.75	4
c. Effectiveness for grasping situation of discussion	3.5	4
d. Effectiveness for selecting target chapter	1.75	2.5
e. Effectiveness for selecting target location in chapter	2.25	1.75

6 Conclusion

In this research, we proposed a system for supporting the collaborative discussion for obtaining the contents related to the research paper. The topics of the discussion are displayed for each chapter according to the similarity between a topic and the chapter and that between topics. The experimental results showed that the visualization of topics is appropriate for grasping the situation of the discussion, but does not contribute for leading a discussion for specific topics.

For our future work, we have to devise a method for leading effective discussion by showing the keywords according to the progress of discussion. Appropriate keywords for the next topic may relate to the previous topics. In order to select such keywords, the detection for keywords which are not discussed effectively in the previous topics and relate to the current topic needs to be developed.

Our collaborative discussion system focuses on a research activity of reading a research paper. The objective of reading the research paper is to make the position of own research clear by evaluating the capability of other researches. In order to help participants assess the paper after the discussion, it is desirable that the discussed topics are arranged from each participant's viewpoint. In our future research, we will help participants evaluate the paper using the results of the discussions.

References

1. Conklin, J., Begerman, M.: glBIS: A Hypertext Tool for Team Design Deliberation. In: Proc. of Hypertext 1987, pp. 247–251 (1987)
2. Kojiri, T., Yamaguchi, K., Watanabe, T.: Topic-tree Representation of Discussion Records in a Collaborative Learning Process. The Journal of Information and Systems in Education 5(1), 29–37 (2006)
3. Viegas, F.B., Donath, J.S.: Chat Circles. In: Proc. of SIGCHI 1999, pp. 9–16 (1999)
4. Erickson, T., Kellogg, W.A., Laff, M., Sussman, J., Wolf, T.V., Halverson, C.A., Edwards, D.: A Persistent Chat Space for Work Groups: The Design, Evaluation and Deployment of Loops. In: Proc. of the 6th ACM Conference on Designing Interactive Systems, pp. 331–340 (2006)
5. Tat, A., Carpendale, S.: CrystalChat: Visualizing Personal Chat History. In: Proc. of the 39th Annual Hawaii International Conference on System Sciences, vol. 3, pp. 58–68 (2006)
6. Xiong, R., Donath, J.: PeopleGarden: Creating Data Portraits for Users. In: Proc. of the 12th Annual ACM Symposium on User Interface Software and Technology, pp. 37–44 (1999)
7. Hayashi, Y., Kojiri, T., Watanabe, T.: Focus Support Interface Based on Actions for Collaborative Learning. International Journal of Neurocomputing 73, 669–675 (2010)

Building a Framework to Design and Evaluate Meta-learning Support Systems

Kazuhisa Seta, Minoru Fujiwara, Daijiro Noguchi,
Hiroshi Maeno, and Mitsuru Ikeda

[1] Osaka Prefecture University, Department of Mathematics and Information Sciences, 1-1, Gakuen-cho, Naka-ku, Skai, Osaka, 599-8531, Japan
[2] JAIST, 1-1, Asahi-dai, Nomi, Ishikawa, 923-1292, Japan
seta@mi.s.osakafu-u.ac.jp,
{fujiwara,noguchi}@kbs.cias.osakafu-u.ac.jp, ikeda@jaist.ac.jp

Abstract. It is difficult to generalize and accumulate experiences of system development as methodologies for building meta-learning support systems. Therefore, we need to build a framework that is useful to design and evaluate meta-learning support systems. Thus we propose a framework as a basis to design and evaluate meta-learning support systems. In this paper, we firstly describe our philosophy to solve the problem. Secondly, we propose a meta-learning process model as a basis to understand meta-learning task and what kinds of factors of difficulty exist in performing meta-learning activities. Thirdly, we explain our conceptualization as a basis to design support functions for prompting meta-learning processes. Then, we integrate a meta-learning process model and the conceptualizations, so that we can design and evaluate meta-learning systems. Finally, we illustrate the usefulness of the framework by taking our presentation based meta-learning system as an example.

Keywords: meta-learning, model-directed approach, meta-learning model.

1 Introduction

The meaning of the concept “meta-cognition” [1, 2] is quite vague, so that the contents of meta-cognitive activities cannot be identified clearly. Therefore, the contents of “meta-cognition support” implemented in learning systems indicate different kinds of supports without explicit analysis/ descriptions [3]. This problem also causes the effects for the evaluation of meta-cognition support systems: it is difficult to evaluate each function embedded into the system eliminates/ removes which factors of difficulties in performing meta-cognitive activities. Therefore, we cannot evaluate the usefulness of each function in detail, although we can show the effectiveness of the system by performing transfer exam. This also means that it is difficult to generalize experiences of system development as methodologies for building meta-learning support systems [4, 5, 6, 7]. Therefore, we need to build a framework that is useful to design and evaluate meta-learning support systems.

In this paper, we firstly describe philosophy of our research to understand our model-directed approach. Secondly, we propose a meta-learning process model as a

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 163–172, 2010.

basis to understand meta-learning task and what kinds of factors of difficulty exist in performing meta-learning activities. Thirdly, we explain our conceptualizations as a basis to design support functions for prompting meta-learning processes. Furthermore, we integrate a meta-learning process model and the conceptualizations, so that we can design and evaluate meta-learning systems based on the deep understanding of meta-leaning processes. Finally, we illustrate the usefulness of the framework by taking our presentation based meta-learning system as an example.

2 Underlying Philosophy of Our Research

Our research aims to build a meta-learning support system that facilitates learner's learning skill development through reflecting his/ her own learning processes. We call "learning of learning activities" (learning of learning methods) as meta-learning. It is well-known that providing meta-cognitively aware instruction is significant to facilitate meta-learning processes [8]. In learning history, for instance, the student might be asking himself as internal self-conversation, "who wrote this document, and how does that affect the interpretation of events," whereas in physics the student might be monitoring her understanding of the underlying physical principle at work.

In learning software development method, not only to memorize how to depict each diagram in UML but also to prompt internal self-conversation processes, e.g., inquire his /herself to answer the reusability and functional extendibility of a designed class structure, is quite important to deepen the learner's own understanding. Meta-cognitively aware instruction is to give learners domain-specific adequate inquiries from the teacher to deepen their understanding. It also facilitates their acquisition of these kinds of domain-specific learning strategy. In our system, we had realized a guidance function that provides meta-cognitively aware instruction based on learning skill ontologies [6].

This function is our original based on the valid knowledge in educational psychology field, however, we don't think it is not enough to accumulate sharable knowledge to develop meta-learning support systems. Thereby, we adopt model-directed approach to clarify the intention of the function, i.e., which factors of difficulties in performing meta-cognitive activities we intend to remove/ eliminate according to a model. This approach is meaningful for the development of human-centric systems in general.

Our meta-learning support system and a learner compose an interaction loop: the system gives stimulations according to the learner's behaviors and they encourage their own intellectual activities prompted by them.

Therefore, we must recognize the learner as a part of the system to achieve our goal of meta-.learning support.

On the other hand, we could not take it into account systematically since cognitive activities of human beings are quite various, latent and context dependent, so that a system developer of a human-centric system can design a sophisticated interaction loop between the system and the learner. Therefore, they tend to design support functions that seems to be subjectively valid based on their experiences without clarifying design intentions. Then, they investigate the validity by performing exams. It means relations among theories clarifying characteristics of cognitive activities in human

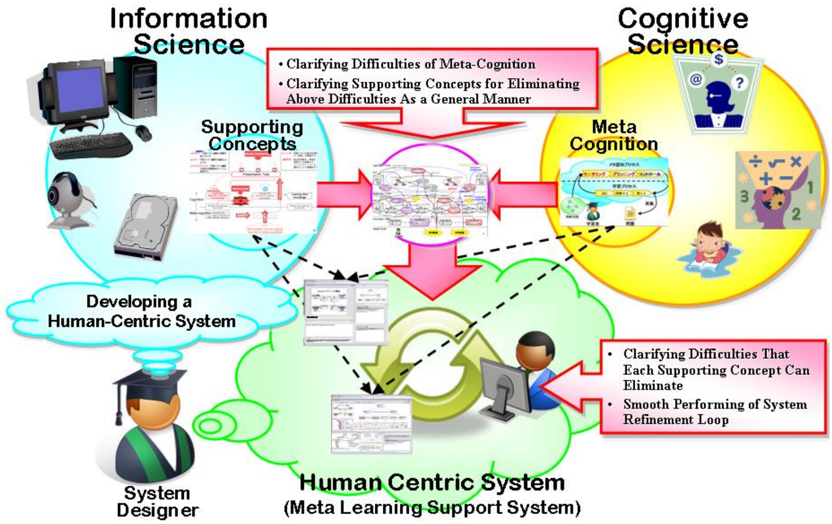


Fig. 1. Model Directed Approach for Building Human Centric Systems

mind and support functions tend to be weak, thereby experiences in developing a system cannot be used/ shared well.

Figure 1 outlines our model directed approach. A system developer, who intends to develop a human-centric system has to design adequate interaction loop to encourage learners' meta-learning activities.

In our framework, we clarify meta-learning model as a reference model to understand which difficulties developers intend to eliminate by extending Kayashima's computational model that is specified based on knowledge in cognitive psychology field (upper right circle). Furthermore, we specify information system support concept at the specific system independent level (upper left circle). Then, we integrate them as a foundation to design and evaluate support functions that eliminates the difficulties. Therefore, they can design support functions based on the understanding of what kinds of difficulties they have to eliminate for the learner and what kinds of support functions they have to realize.

One of the significant differences between ordinary and our approach is that we can clarify design rationales of each support function implemented into the system.

3 Foundation to Build the Framework

In this section, we overview the structure of meta-learning processes to understand the meta-learning task. Then, we build meta-learning process model corresponding to upper right circle in fig. 1, and explain our conceptualizations corresponding upper left circle in fig. 1.

3.1 Structure of the Meta-learning Task

Figure 2 represents cognitive activities in performing problem-solving processes (left side) and the ones in performing learning processes (right side), respectively. A problem-solver performs cognitive activities, e.g., reads, understands a given problem and solves it. At this time, he/ she also performs cognitive activities that monitor, re-plan and control them. These are called meta-cognitive activities in the sense of cognitive activities that handle cognitive activities.

Kayashima et. al. constructs a framework by which we can understand factors of difficulties in performing meta-cognitive activities in problem-solving. It clarifies factors of difficulties based on cognitive psychology knowledge, e.g., segmentation of process, invisibility, simultaneous processing with other activities, simultaneous processing with rehearsal, a two-layer working memory, acquisition of criteria for cognitive activity and so on.

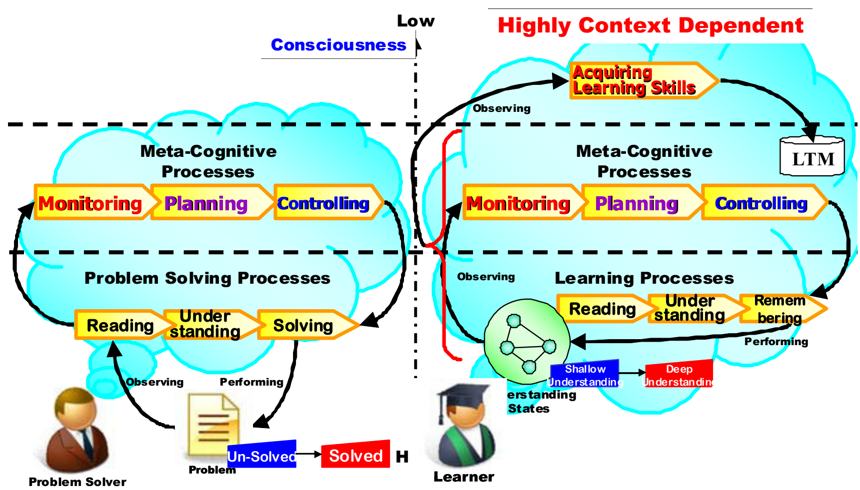


Fig. 2. Structure of Meta-Cognitive Task in Performing Problem-solving Processes (left side) and Learning Processes (right side)

On the other hand, performing meta-cognitive activities in learning (planning of learning processes) is more difficult than the one in problem-solving (planning of problem-solving processes), since problem-solving activities and their results are visible whereas those of learning activities and their results (learner's understanding states) are in-visible. Thus, the learners do not tend to be aware of necessity of monitoring and controlling their learning processes. Thereby, they do not tend to perform meta-cognitive activities spontaneously.

Furthermore, planning learning activities compels heavier cognitive loads to the learner since they require monitoring activities of one's own invisible understanding states, It is difficult for ordinary learners to perform even though they intend to do.

Furthermore, learning know-how for learning process planning (top layer in fig. 2(right)) by reflecting learning activities performed at the second layer is more latent and higher level activities than meta-cognitive ones performed at the second level in problem-solving processes.

3.2 Meta-learning Process Model

Here, we give more detailed model of meta-learning activities. Figure 3 shows a meta-learning process model by extending Kayashima's computational model capturing meta-cognitive activities in problem-solving processes (see [3] in detail). It captures meta-learning processes in a learner's head (working memory). It is separated by three layers. It represents changing processes of the learner's understanding states by performing learning activities at the lowest layer, ones of planning learning processes at the middle layer, and ones of reflecting activities for acquiring learning skills at the top layer, respectively. Each ellipse represents a product produced at each layer and "t*" represents the time. Therefore, the order of "t*" represents changes of a product.

The model represents that the learner had intended to make himself/ herself understand the feature of functional extendibility of the Iterator Pattern in software design pattern (at the middle layer). But he/ she could not understand well by itemizing these features (at the lowest layer). Then she got be aware of the lack of her understanding by monitoring own understanding states and re-plans her learning processes (at the middle layer).

If this meta-cognitive activities (learning process planning) adequately performed, e.g., make the learning plan to understand functional extendibility of the iterator patterns by considering a correspondence between functional extendibility and class structures, she can understand the topic deeply.

Learning-skill acquisition processes at the top layer require following cognitive activities: (i) reflecting and observing the learning processes (reflecting cognitive activities performed at the lowest), (ii) detecting meaningful domain-dependent learning operators that had deepened his/ her understanding states, and (iii) re-evaluate, generalize and store them into the long term memory.

This model plays a role of clarifying factors of difficulty in performing meta-learning process. It will be described in section 4.

3.3 Conceptualization to Design Support Functions for Meta-learning

By building the detail meta-learning process model, we can conceptualize general support concept in-dependent of concrete support functions.

Table 1 shows five concepts for supporting meta-learning that we specified from the viewpoint of information system development: SHIFT, LIFT, REIFICATION, OBJECTIVIZATION, TRANSLATE. This conceptualization corresponds to the upper left circle in fig. 1.

SHIFT means that stagger the time of developing learning skills after performing learning processes.

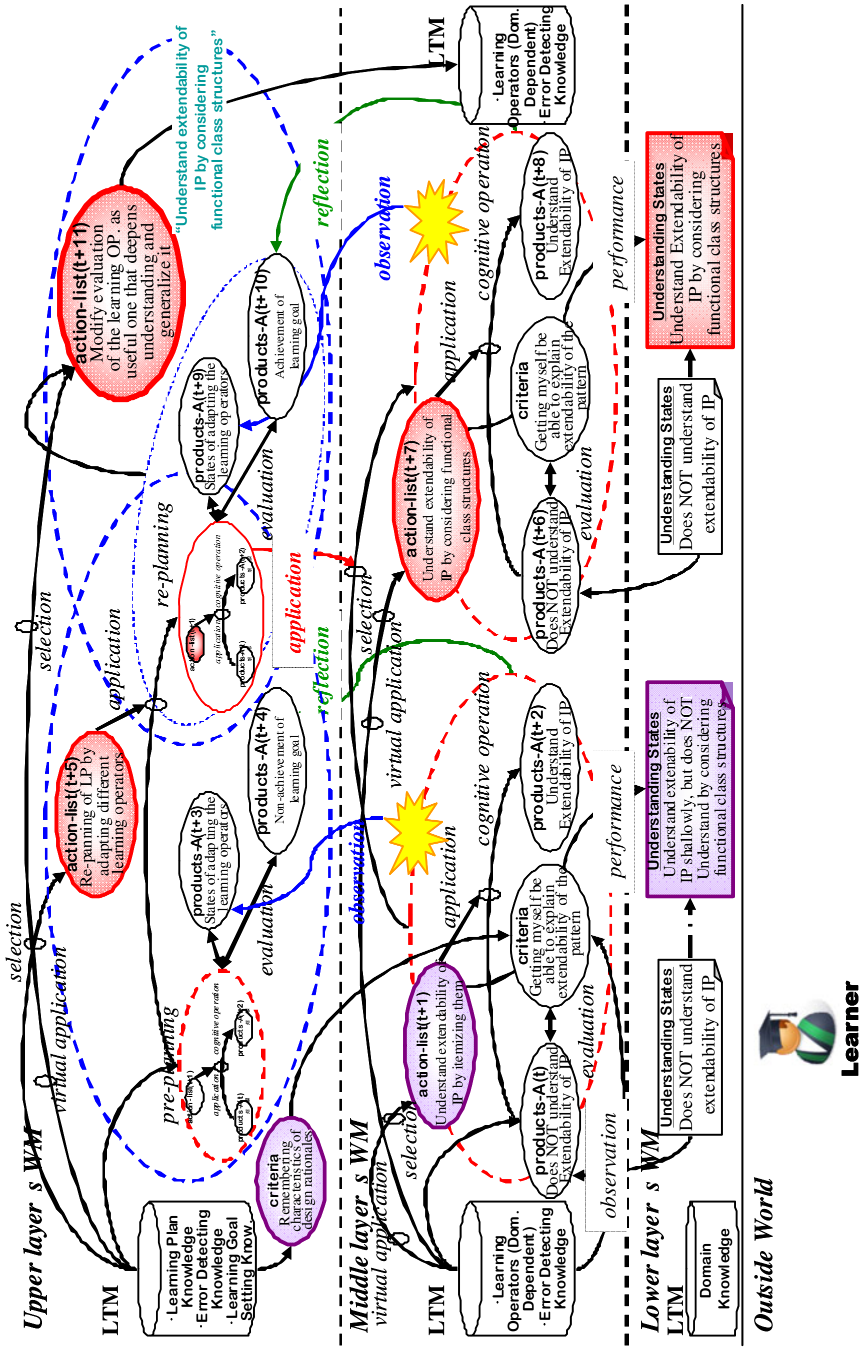


Fig. 3. Meta Learning Process Model

Table 1. Correspondence Among Supporting Concepts and Their Targets

Conceptualization	Meaning	Target to eliminate factors of difficulties	Learning Scheme Design
SHIFT	Stagger the time of developing learning skills after performing problem-solving processes	• Simultaneous processing with other activities • Planning • Influence on virtual application at a lower layer	Task Design (giving a presentation topic the learner had already learned)
LIFT	Make the learner be aware of learning skill acquisition	• Invisibility • Influence on virtual application at a lower layer (learning process • Planning	Visualization Environment
		• Reduce a three-layer WM to a two-layer WM • Evaluation of influence on virtual application at a lower layer (problem-solving)	Guidance Function
		• Acquisition of learning operators	Guidance Function
REIFICATION	Give appropriate language for his/her self-conversation to acquire learning skills	• Segmentation of process	Providing Domain Specific Terms of Learning Activities
TRANSLATE	Transfer the learning skill acquisition task (LSAT) to a problem-solving task that includes same task structure of LSAT.	• Reduce a three-layer WM to a two-layer WM • Influence on virtual application at a lower layer	Task Design (giving a presentation task to explain to other learners)
OBJECTIVIZATION	Objectify her/his self-conversation processes by externalizing them for learning communications with other learners	(Context Dependent)	CSCL Environment

Furthermore, it is needed to give appropriate stimulation to encourage their meta-cognition. This stimulation can be interpreted that it gets the meta-cognitive task as easy as cognitive task by changing internal self-conversation task to usual conversation task. Thus, we conceptualize LIFT as making the learner be aware of learning skill acquisition as a principle for the system development in this research.

We think that how we can realize the SHIFT and LIFT is the key issue for developing meta-cognitive skills.

REIFICATION means that giving appropriate language for the subject of meta-cognition. Of course, since we can not realize the LIFT if we do not give appropriate reification. Thus the concept of REIFICATION is included in the concept of LIFT. But we can not always realize appropriate LIFT even if we realize REIFICATION; we need to give suitable REIFICATION to prompt learners' meta-cognition. Thus, we separate the REIFICATION from LIFT concept since it is easy to discuss from the technical aspects.

By OBJECTIVIZATION, we intend making the internal self-conversation processes objective by discussing with others.

TRANSLATE means changing the learning skill acquisition task to a problem-solving task that includes same task structure of learning skill acquisition task.

The conceptualizations play a role of principle for our learning scheme design.

4 Integrating Meta-learning Process Model and Conceptualizations

In this section, we integrate the meta-learning process model and conceptualizations to build a framework to design and evaluate meta-learning support systems.

Meta-Learning process model clarifies the factors of difficulties in performing meta-learning processes. Third row in Table 1 shows them: simultaneous processing with other activities, planning, invisibility, three-layer WM, acquisition of learning operators and so on.

Table 1 represents correspondence among conceptualizations and their targets to eliminate/ remove factors of difficulties in performing meta-learning processes. For example, SHIFT removes factors of simultaneous processing with other activities and eliminates those of planning and influence on virtual application at a lower layer, while REIFICATION does factors of segmentation of process. Furthermore, TRANSLATE reduce a three-layer WM to a two-layer WM and removes difficulty of influence on virtual application at a lower layer by translating learning skill acquisition task to the problem-solving task.

Consequently, we can understand which factors of difficulties we should eliminate and how we should realize.

The right row in the table illustrates concrete supports implemented in our presentation based meta-learning scheme. For example, based on SHIFT principle, we set presentation task whereby the learner makes a presentation material about already learned topic. Therefore, developers can refer/ compare what kinds of support functions are implemented to realize each conceptualization in designing support functions for their systems.

5 Using Framework to Design and Evaluate Our System

By building the framework, we can build a presentation based meta-learning support system with explicit clarification of design rationale of it. In this section, we illustrate the usefulness of our framework by taking our system as an example. Thereby, we describe our system briefly. See detail explanation in [6]: it is out of the subject in this paper.

5.1 Building Presentation Based Meta-learning Support System

Design principle of each support can be clarified based on the table, e.g., providing domain specific terms of learning activities intends to eliminate the difficulty of segmentation of process.

SHIFT and TRANSLATE are realized as our task design, that is, providing the task where the learner has to explain pre-learned knowledge. LIFT is realized as a function that provides the learner with guidance information for checking the validity of designed learning processes. We had embedded two kinds of guidance function: the one realizes meta-cognitively aware instruction and the other does navigation function to support meta-learning communication among learners. REIFICATION is realized as providing terms for representing learning processes and visualization environment. OBJECTIVIZATION is realized as embedding CSCL environment in the system.

In our system, more concretely, we presuppose a learner who has already learned a specific topic, UML and software design patterns [9] We give the learner the task of producing readily comprehensible presentation material for other learners whose academic ability is similar to that of the presenter. This task setting is important for the learner to focus on meta-cognitive learning: if the learner must perform both learning and making presentations, the learner cannot allocate sufficient cognitive capacity to perform the meta-cognitive activities. This task setting corresponds to the SHIFT. It staggers the time of performing monitoring and generalizing processes after performing learning. In preparing presentation materials, the learner monitors the previous own learning processes and gives queries to herself for validating them. This stimulation corresponds to the LIFT It lifts monitoring and generalizing processes to the cognitive level. Then, she discusses with others whether the presentation material is easy to understand or not. This corresponds to OBJECTIVIZATION. REIFICATION provides terms for representing learning processes, for instance, providing terms of “make the learner understand the functional extendibility of the DP by analyzing the class structure,” and plays an important role to realize appropriate LIFT and OBJECTIVIZATION.

5.2 Specifying Items for Evaluating Meta-learning Support Systems

Making questionnaire items for evaluating the usefulness of individual support functions is quite hard, since it is difficult to clarify design principle of them. By clarifying the correspondence among support concepts, factors of difficulty that each support concept removes/ eliminates and concrete supports embedded into each learning scheme, we can make questionnaire items based on it.

Actually, we specified and used 31 evaluation items for constructing a questionnaire for a pilot evaluation of our system. We list some of them as follows:

- Did you feel insufficiency of your understanding in making a presentation material even you thought you had already understood the topic? [Regarding SHIFT and TRANSLATE]
- Do you think making a presentation material makes it easier to reflect your own learning methods? [Regarding SHFT]
- Do you think reading vocabulary prompts your reflection on your own learning processes? [Regarding REIFICATION an LIFT]
- Do you think your presentation structure reflects your learning processes? [Regarding LIFT]

We can compare evaluation results of learning systems even though evaluation items are different, since they are specified based on general-level support concepts.

Because of the small pilot evaluation, we could not conclude the usefulness of each questionnaire items, however, we verify that all 7 learners can understand and answer each questionnaire item even though they are not familiar with cognitive science; most items ask about their cognitive activities that require them to be aware of their cognitive activities.

6 Concluding Remarks

In this paper, we described philosophy of our research to understand our model-directed approach. Then, we proposed a meta-learning process model and our conceptualization. Furthermore, we integrate a meta-learning process model and the conceptualizations, so that we can design and evaluate meta-learning systems based on the deep understanding of meta-leaning processes. It plays an important role to accumulate and share experiences of individual learning system development. Finally, we illustrate the usefulness of the framework by taking our presentation based meta-learning system as an example.

According to the framework we could evaluate usefulness of each function more detail and some questionnaire results suggested they worked well due to our design rationales. We will carefully address this matter in another paper because it is out of the scope of this paper.

References

1. Flavell, J.H.: Metacognitive aspects of problem solving. In: Resnick, L. (ed.) The Nature of Intelligence, pp. 231–235. Lawrence Erlbaum Associates, Hillsdale (1976)
2. Brown, A.L., Bransford, J.D., Ferrara, R.A., Campione, J.C.: Learning, Remembering, and Understanding. In: Markman, E.M., Flavell, J.H. (eds.) Handbook of Child Psychology, 4th edn. Cognitive Development, vol. 3, pp. 515–529. Wiley, New York (1983)
3. Kayashima, M., Inaba, A., Mizoguchi, R.: What Do You Mean by to Help Learning of Metacognition? In: Proc. of the 12th Artificial Intelligence in Education (AIED 2005), Amsterdam, The Netherlands, 18-22, pp. 346–353 (2005)
4. Kashihara, A., Taira, K., Shinya, M., Sawazaki, K.: Cognitive Apprenticeship Approach to Developing Meta-Cognitive Skill with Cognitive Tool for Web-based Navigational Learning. In: Proc. of the IASTED International Conference on Web-Based Education (WBE 2008), Innsbruck, Austria, March 17-19, pp. 351–356 (2008)
5. Kojima, K., Miwa, K.: A Case Retrieval System for Mathematical Learning from Analogical Instances. In: Proc. of the International Conference on Computers in Education (ICCE), pp. 1124–1128 (2003)
6. Maeno, H., Seta, K.: Guidance Generation for Facilitating Meta-Cognitive Learning Through Presentation Task*. In: Proc. of the 2009 International Conference on Multimedia, Information Technology and its Applications, Osaka, Japan, pp. 5–8. IEEE Press, Los Alamitos (2009), ISSN 1975-4736
7. Nakano, A., Hirashima, T., Takeuchi, A.: Developing and evaluation of a computer-based problem posing in the case of arithmetical word problems. In: The Fourth International Conference on Computer Applications, ICCA 2006 (2006)
8. John, B., Brown, A., Cocking, R. (eds.): Brain, Mind, Experience, and School., in How People Learn. National Academy Press, Washington (2000)
9. Gamma, E., Helm, R., Johnson, R., Vlissides, J.M.: Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley Professional, Reading (1994) (illustrated edn.)

Promoting Learning Attention with Gaze Tracking Integrated e-Learning Contents

Kai Li[1] and Yurie Iribe[2]

[1] Research Center for Agrotechnology and Biotechnology,
Toyohashi University of Technology, Japan
kaili@recab.tut.ac.jp
[2] Information and Media Center, Toyohashi University of Technology, Japan
iribe@imc.tut.ac.jp

Abstract. No learning could take place without input attention. In order to promote students' attention in video and slide synchronized e-learning content, this study focuses on gaze tracking data as a kind of enhanced input. An eye mark is showed at the gaze position to indicate where the teacher is explaining. As a result, gaze tracking could used to promote students' learning attention in e-learning. Further studies are needed to compare gaze input and other input in terms of efficiency and accuracy.

Keywords: e-learning, gaze tracking, multimodal interface, learning attention.

1 Introduction

1.1 Gaze Tracking

Gaze tracking has been well researched in the field of human-computer interaction [1][2]. Gaze-based computer interfaces so far have often focused on using gaze as the input channel, for instance for eye-typing or for moving the mouse cursor for disabled persons [3]. Prior studies show that multimodal interaction using click activation increases the efficiency of gaze interaction with respect to gaze-based selection such as dwell [4]. Moreover, recent studies have shown that the performance of a low-cost gaze tracker compares well to the commercial gaze systems in target-acquisition tasks [5]. However, the accuracy of the measured point of gaze is problematic. Due to drifting, the practical accuracy is often worse. This means that selection of small items is hard by gaze.

In real world, we often use pointing devices to indicate the important point on the contents, for example, we cloud use a mouse or laser pointer to catch students' attention in classroom lectures. However most of e-learning contents are produced in a professional studio room. The teacher just has a lecture before a camera without students. Most of the time, there is no blackboard and laser point in the studio, since the low resolution by video camera. Instead, we use a touch panel to display and to indicate the important point on the contents. In e-learning environment, students have no

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 173–179, 2010.

face-to-face communication with teachers; therefore they could miss some important information by just watching teacher's video. Even though the teacher could use touch panel to add handwriting memos in e-learning contents, we found little handwriting memos in our e-learning practice. As we mentioned gaze tracking could be used in target-acquisition tasks, we consider of using gaze tracking to promote students' attention in e-learning.

1.2 Attention

Studies in the fields of cognitive science [6] and cognitive psychology [7] demonstrated that no learning could take place without input 'attention'. Researchers [8][9] have been trying to find methods to attract learners' attention. The methods such as textual enhancement, input flooding and processing instruction were employed by some researchers as a means of guiding students' attention toward the learning contents. Textual enhancement, for instance, highlights the forms to be noticed by using bold or uppercase letters, underlining, shadowing and different fonts. However, the advantages of textual enhancement over unenhanced input have not been clearly demonstrated [10]; some studies have reported no positive benefits whereas others have [11]. More studies are, therefore, needed to clarify whether enhanced input can trigger processes of giving attention. The current study is an attempt to contribute to both this line of research and multimodal input studies. In this study, gaze tracking integrated e-learning content is conducted. In addition, it seeks to determine whether learners find this kind of enhanced input effective in attracting attention and improving accuracy than other methods.

The aim of this paper is to conduct a gaze tracking integrated e-learning content with accurate commercial gaze system (NAC Eye Mark Recorder EMR-8). Since our eyes represent a natural substitute for the mouse, an eye-mark is showed at the gaze position on teaching materials to indicate where the teacher is teaching.

To our knowledge, gaze tracking integrated video and slide synchronized e-learning contents have not been researched. This pilot study is the first step towards promoting attention by gaze tracking. We believe this area offers a rich set of opportunities for future research and development.

2 Previous Research

In order to verify enhanced input effective on promoting attention in video and slide synchronized e-learning, prior to the study, a questionnaire (see Table 1) was conducted on students' attention in handwriting memo integrated e-learning content. 16 participants were asked to learn a 20 minutes video and slide synchronized e-learning content which was integrated with teacher's handwriting memo such as underline, circle, and arrow (see Fig.1). Hand-writing memo can represent varieties of information, for example, to indicate the important point, the range of the contents or the position of where he is explaining.

Fig. 1. Handwritting memo integrated e-learning content

The results of the questionnaire (5-scale question, 1 means not so much, 5 means so much) show that students think teacher's handwriting is important in learning(M=3.75, SD=1.29), and they became to concentrate on the positions of teacher's handwriting memo(M=3.62, SD=1.2). It is also showed that handwriting memo cannot result in a positive effect on the key points of the contents (M=2.69, SD=1.25). In other studies, handwriting has also reported a positive effect on memory and teacher's presence [12]. It is also suggested that teachers should use different feedback strategies and offer appropriate types of feedback according to the level of students' proficiency in order to enhance their skills and learning motivations [12].

It seems that handwriting memo as a kind of enhanced input in video and slides synchronized e-learning content has positive effect on learning attention, but as a matter of fact, teachers have little handwriting memos on e-learning contents, since most of teaching contents are involved in e-learning slides, they do not need to write it again and some teachers said that handwriting sometime interrupt their teaching. Therefore, we consider using gaze tracking as an enhanced input to promote students' attention.

Table 1. Questionnaire on impression of handwriting memo

Question items	M	SD
I think handwriting memo in e-learning content is important.	3.75	1.29
I became to concentrate on the positions of teacher's handwriting memo.	3.62	1.2
I am satisfied with teacher's handwriting memo.	2.94	1.06
I prefer teacher's handwriting memo in e-learning content.	2.87	1.2
I have a deep impression on teacher's handwriting memo	2.69	1.08
I became concentrate on the learning content	2.69	1.25
I became to know the key points of the content	2.69	1.25

3 Task and Procedure

The task of the study is to record teacher's gaze tracking data, and integrate the data in e-learning contents to indicate where the teacher is teaching.

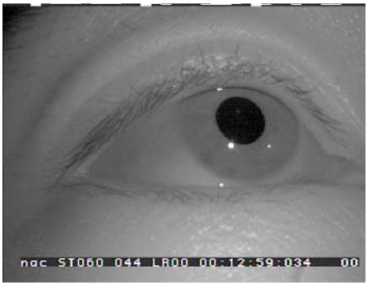


Fig. 2. Pupil detection

;00000096,0,999,640,480,512,091,290
;00000097,0,999,640,480,512,092,293
;00000098,0,999,640,480,512,091,297
;00000099,0,999,640,480,514,091,295
;00000100,0,999,640,480,514,091,294
;00000101,0,999,640,480,516,091,295
;00000102,0,999,640,480,516,092,296

Fig. 3. Sample data of gaze tracking 2-9 digits mean time code, 29-35 digits mean right-eye movement coordinate

3.1 Eye-Movement Measuring

First, teacher's eye-movements were recorded using a video-based eye tracker (nac:EMR-8). The teaching material (MS. PowerPoint) was displayed on a touch panel connected with a laptop computer positioned 60 cm from the teacher. The teacher could add handwriting memos on the slides with touch panel. Teacher's video, teaching materials and handwriting memos were synchronized with e-learning contents producing tool (CyberLink: StreamAuthor V4). The adjustment for pupil detection and calibration were conducted at the beginning (see Fig. 2). Eye-movement was tracked on a 1024 by 768 pixel screen at 60 Hz. The accuracy of the spatial resolution of this equipment is noted in the manufacturer's catalog as being a visual angle of 0.1 degrees (see Fig. 3). Since it is an accurate gaze tracking measuring equipment, we select it in eye-movement recording. Eye-movement coordinate data was recorded synchronously on another laptop computer as time course data, while teacher read the text content of each slide. The tracking data recorded in 640 by 480 pixels was then converted into the same resolution with slides in order to overlay the eye-mark on the e-learning contents.

3.2 Gaze Tracking Integrated e-Learning Content

In order to duplicate the teacher's gaze tracking, an eye mark of 30 by 30 pixels transparent gif is conducted by Adobe Flash. It could wink like a real person.

Since movement of mouse cursor could be recorded and showed on e-learning content by StreamAuthor tool, in the study, we use the eye mark represent mouse movement. The gaze tracking data was converted into mouse movement data so that the eye-mark could be showed at the gaze position on teaching slides to indicate where the teacher is teaching. As most of the e-learning contents in our e-learning practice are recorded in a professional studio without students, the teacher just watches his teaching material displayed on a touch panel to have a lecture. When he moves his eyes inside the touch panel display, the gaze tracking data and time code (see Fig.3) could be recorded and converted into eye-mark moving data. When he moves his eyes outside the display, the over-data are recognized and there is no eye-mark displayed in the e-learning content (see Fig.4).

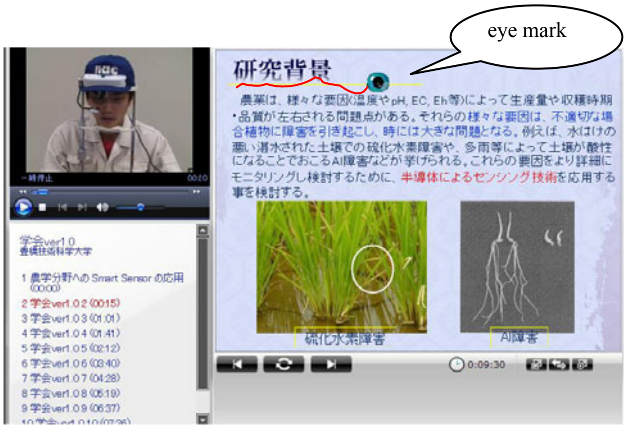


Fig. 4. Gaze-tracking integrated e-learning contents

4 Results

In order to verify the gaze tracking effective on promoting attention in e-learning, a similar questionnaire as the previous research was conducted (see Table 2). With only three participants' answer, we got positive results of gaze tracking as previous research on handwriting which students thought the gaze tracking was important and they have concentrated on the eye-mark and they could know where the teacher is talking about. Since the little sample of the study, further studies with large samples are needed to have an in detail evaluation and comparison.

Table 2. Questionnaire on impression of gaze tracking

Question items	M	SD
I think gaze tracking in e-learning content is important.	3.67	0.58
I became to concentrate on the positions of teacher's eye-mark.	4	0
I am satisfied with teacher's eye-mark.	3.67	0.58
I prefer teacher's eye-mark in e-learning content.	3.33	0.58
I have a deep impression on teacher's eye-mark	4	0
I became concentrate on the learning content	2.67	0.58
I became to know the key points of the content	2.67	0.58

5 Limitations of the Study

Since gaze tracking coordinate data (see Fig. 3) would be changed with little head movement, in order to get an accuracy data of eye-movement, in this study, we asked the teacher to wear a commercial gaze system on his head and his head was fixed on a stand. By fixing his head on the stand, we could get a reliable gaze tracking data for study, but it is discommodious to the teacher in video recording. Furthermore, the commercial gaze system used in the study is a high-accuracy device but too expensive. A low-cost gaze tracker and unfixed gaze tracker should be developed in later studies.

And some features of eye-movements should be extracted by trials and errors. Since the measured point of gaze has drifted a few pixels off from the actual focus point, it is difficult to place the eye-mark exactly on the desired location in the text material with drifted data.

Also, further study is needed to compare gaze tracking with other enhanced input such as handwriting and mouse in terms of learning efficiency and accuracy with large samples.

6 Conclusion

In this study, gaze tracking as a kind of enhanced input was integrated in e-learning content to promote learning attention. Teacher's eye-movement was recorded and represented where he is teaching. As reported, enhanced input could improve attention and memory. This pilot study is the first step towards promoting attention by gaze tracking. We believe this area offers a rich set of opportunities for future research and development.

Acknowledgements

This research has been supported in part by the Grant-in-Aid for Encouragement of Young Scientists (B) No.21700813 from the Ministry of Education, Science, Sports and Culture in Japan.

References

1. Duchowski, A.T.: Eye tracking methodology: Theory and practice, 2nd edn. Springer, New York (2007)
2. Jacob, R.J.K.: What you look at is what you get: eye movement-based interaction techniques. In: CHI 1990: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Seattle, Washington, United States, pp. 11–18. ACM, New York (1990)
3. Majaranta, P., Räihä, K.-J.: Text Entry by Gaze: Utilizing Eye-Tracking. In: MacKenzie, I.S., Tanaka-Ishii, K. (eds.) Text Entry Systems: Mobility, Accessibility, Universality, pp. 175–187. Morgan Kaufmann, San Francisco (2007)
4. Zhang, X., MacKenzie, I.S.: Evaluating eye tracking with ISO 9241 - Part 9. In: Proceedings of HCI International 2007, pp. 779–788. Springer, Heidelberg (2007)

5. San Agustin, J., Skovsgaard, H.H., Hansen, J.P., Hansen, D.W.: Low-Cost Gaze Interaction: Ready to Deliver the Promises. In: CHI 2009 Extended Abstracts on Human Factors in Computing Systems. ACM, New York (2009)
6. Tomlin, R.S., Villa, V.: Attention in cognitive science and second language acquisition. Studies in Second Language Acquisition 16(2), 183–203 (1994)
7. Carr, T.H., Curran, T.: Cognitive factors in learning about structured sequences. Applications to syntax. Studies in Second Language Acquisition 16(2), 205–230 (1994)
8. Robinson, P.: Attention, memory, and the ‘noticing’ hypothesis. Language Learning 45(2), 283–331 (1995)
9. Schmidt, R.: Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In: Schmidt, R. (ed.) Attention and Awareness in Foreign Language Learning, pp. 1–63. University of Hawai’i Press, Honolulu (1995)
10. Corbeil, G.: Can PowerPoint Presentations Effectively Replace Textbooks and Blackboards for Teaching Grammar? Do Students Find Them an Effective Learning Tool? CALICO Journal 24(3), 631–656 (2007)
11. Leow, R.: Attention, awareness and foreign language behavior. Language Learning 51(1), 113–155 (2001)
12. Li, K., Akahori, K.: Development and Evaluation of a Feedback-giving Support System with Audio and Playback Strokes. CALICO Journal 26(1), 91–107 (2008)

System for Creative Distance Learning Environment Development Based on Competence Management

Przemysław Różewski and Bartłomiej Małachowski

West Pomeranian University of Technology in Szczecin,
Faculty of Computer Science and Information Systems,
ul. Żołnierska 49, 71-210 Szczecin, Poland
{prozewski,bmalachowski}@wi.zut.edu.pl

Abstract. The student's creativity comes into action when student produces an outcome or product that is both novel and useful as defined within some learning context. The creative distance learning environment supports student's creativity on the social, cognitive and technology levels. In this paper the method for creativity environment development is presented. Based on the competence set theory the creativity level can be measured. Moreover, competence modelling allows to created individually adapting creative tasks and projects as well as creativity supporting knowledge network. The authors propose the model of e-learning information system for competence-based creative space development.

Keywords: competence management, competence set theory, creative distance learning environment, open and distance learning.

1 Introduction

Nowadays the educational system has tried to reach synergy with the labour market. The requirements for graduate's competence from labour market should be reflected in knowledge pass to students during his/her learning process. Additionally, the educational process is influenced by the Life-Long Learning (LLL) idea. That is the reasons for educational system flexibility and openness improvement. The Open and Distance Learning (ODL) concept combines all this issues in one idea [5]. The ODL concept assumes open and freely access to didactical material and services based on the Internet and e-learning technology. Moreover, in the ODL framework students can freely choose the place to obtain next educational level and competencies. The European Higher Education Area (EHEA) is an example of real implementation of ODL concept.

The important supplement of ODL based educational system is a competence-based learning [9]. In the competence-based learning the learning/teaching process focused on the competences transfer to student based on the didactical material which consists different knowledge types [10]. The competence itself is defined as demonstrated ability to apply knowledge and skills (ISO 24763). The main idea is to record student's competence in comparable, compatible and coherent systems based on the standardized system's interface and metadata system. In the EHEA the Common Qualification

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 180–189, 2010.

Framework is developed to reach this goal. The Qualification Framework is a whole competence ecosystem supported competence-based learning (research project: Ten-Competence: www.tencompetence.org, ICOPER: www.icoper.org). The authors already introduced Competence Object Library [8], which integrating competence description standards with existing mathematical methods of competence analysis into one Java-based software library. Moreover, the model of a social collaboration environment for knowledge management in competence-based learning [9] has been proposed.

In the paper the authors described the method of information system development for creative distance learning environment. The system's goal is to match task (or project) with student's creativity potential. In addition the system is able to create and maintain knowledge network, which support creative process based on the learning objectives. Based on the competence set theory [16] the method for task matching problem and knowledge network development is proposed.

The approach based on the competence set theory allows to make quantity calculation of student's and task's creativity potential. The competence set's optimal expansion cost algorithm make possible to compare acquired competence set with required competence set. To make this computation possible the information about student's competence set as well as competence set required for task's (project's) set have to be accessible to the system. The authors proposed information system architecture which fulfills this requirement.

2 The Issue of Creativity in Distance Learning Environment

The student's creativity in distance learning environments is limited by the technological features of learning information systems. The contact or connection with other students is possible only in the framework of information systems. The student works with the resources based on the provided tools. In the distance learning scenario student is not able to maintain face-to-face contact with teacher and others students.

The concept of creativity itself is a filed of intensive research work, especially on the psychology [11] and problem solving [12] field. Generally speaking creativity is the interplay between ability and process by which an individual or group produces an outcome or product that is both novel and useful as defined within some social context [7].

In the distance learning it is important to recognized proper context of creativity. The question is that creativity is domain-general (applicable to all disciplines and tasks) or domain specific (tailored to specific disciplines and tasks) [4]. In the authors opinion the hybrid models [6] fulfilled the distance learning needs in the best way. According to cognitive research most notably, evidence suggests that transfer of learning is hindered when learning occurs in a tightly focused context (i.e., too much specificity) or content neutral context (i.e., too much generality) [6]. In the proposed solution on the first step the material is introduced in a specific domain or task area but then system assists students as they apply the new knowledge or skills to different tasks and domains through knowledge network.

The fig. 1 presents main aspects of creativity in distance learning. The creative dimension is based on the work [1] and technology on [3]. There are no personal characteristics of creative individuals on the figure. Such characteristics can be found in [4].

The creative distance learning environment should combine all dimensions presented on the fig.1. In the discussed approach the cognitive dimension is supported by the individually generated student's tasks. Every task is designed to address individual student's solving skills. The competence set method allows to choose the task in order to match student's knowledge and competence capabilities. In the cognitive dimension creativity has following meaning: on the one hand task's solution should required best student's effort, on the other hand the student should possess all initial competences and knowledge necessary to solve the task.

The social dimension of creative distance learning environment is maintained by the creation of knowledge network. The knowledge network combines the main characteristics of social network and production network [2]. The social network is reflected in the relations and relations' content between agents, the logic of this relation comes from production network. Knowledge network takes knowledge as the main flow elements. Knowledge network operation includes the knowledge sharing and knowledge transfer [2].

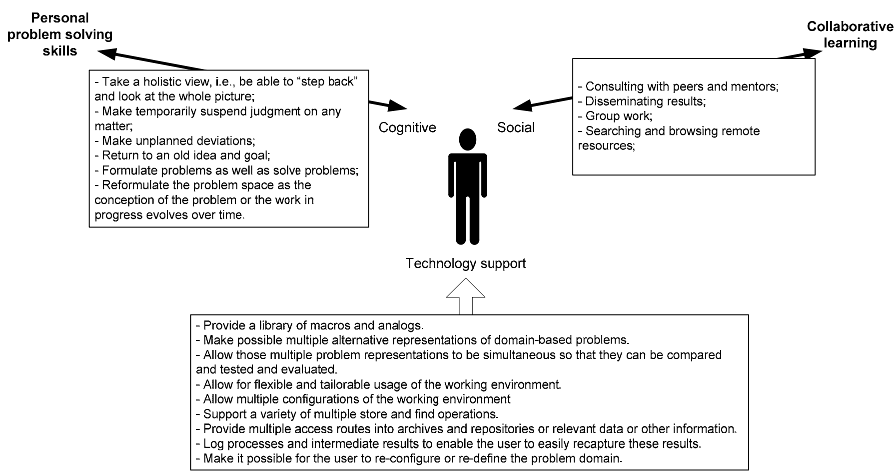


Fig. 1. Different aspects of creative distance learning environment

The technology dimension is supported by the tools, which are designed to maintained creativity process. The functional requirements for information and computer tools are presented on fig. 1. The information and computer tools have to allow students to work on the knowledge level, where communication based on the ontology modeling [9]. Moreover the tools take an advantage of object approach to competence modeling [8].

3 Expressing Creativity with Formal Competence Models

3.1 Formal Competence Models

Survey of the scientific literature in the fields of knowledge modeling, human resource management and learning management provides many different and sometimes unambiguous definitions of the notion "competence". This vast set of definitions was put together by International Standard Organization that gives very coherent and brief definition of competence. According to ISO 9000:2005 the competence is "demonstrated ability to apply knowledge and skills". This simple description of the competence expresses the main idea lying behind the research described in this article. Basing on ISO definition of competence we assumed, that since competence is described as "demonstrated ability" it can be used as a measure of personal performance. There are many studies following this approach in different fields of business oriented literature.

The main problem with competence-based formal models is how to quantitatively measure and process human competence. There are not many quantitative models for competence representation. One of the most advanced idea of this type is the approach called *competence sets* (CS). This approach was for the first time introduced by Yu and Zhang [15], [16]. These authors model competence as the set containing skills, information and knowledge possessed by a person (acquired competence set denoted *Sk*) or required to successfully perform a given job or a task (required competence set *Tr*).

In the early stage of the research on CS, competence was modelled as a classical set containing knowledge, skills and information necessary to solve a problem. However, expressing the presence of a competence in binary terms – one has a competence or not at all (or it is necessary to have a certain competence to solve a problem or not) – turned out to be insufficient regarding the continuous nature of competence. Taking this fact into consideration it was proposed to present human competence as a fuzzy set, defined as follows [13]:

$$A = \{(x, \mu_A(x)) \mid x \in X\}$$

where: $\mu_A(x)$ is the membership function assessing the membership of an element x in relation to set A by mapping X into membership space [0; 1], $\mu_A : X \to [0;1]$.

Basing on the definition of the fuzzy set it is possible to define the notion of *fuzzy competence strength*, that expresses the level of competence presence or requirement. For each competence g, its strength is a function of a person P or a task E in the context of which the competence is assessed: $\alpha : \{P \text{ or } E\} \to [0;1]$. Expansion optimization methods of fuzzy competence sets are computationally more demanding but provide better accuracy and reproduction of nature of competence.

CS methods provides quantitative measure of human competence through optimization and cost analysis of the competence set expansion process [13], [15], [16]. This process is described as obtaining new skills and adding them to the actual acquired competence set *Sk* of a person. The cost and pace of obtaining new skills depends on elements of actual competence set and how close these elements are related with the new skill. Methods of optimal competence set expansion consist of determining the order of obtaining successive competences that provides minimal cost. Competences that need to be obtained are defined by set $Tr(E) \backslash Sk(P)$, where $Sk(P) \subseteq Tr(E)$. The optimisation problem is usually solved by finding the shortest path in an oriented graph, in which vertices represent competences and arcs represent the relations between them [15], [16]. The general cost of expanding competence set is given by the cost function $c(Sk(P),Tr(E))$. The form of the cost function varies in different methods for competence expansion cost analysis that can be found in the literature of CS [13], [15], [16] and can be chosen individually according to application requirements.

The methods for competence set analysis with its quantitative models provide solid background for development of competence management system that can be easily implemented and applied in many domains. At the present time the competence set methods are used mainly in decision theory and knowledge management application. In the literate following competence set methods can be recognized [13], [14], [15], [16]: competence set analysis, competence set expansion, competence set expansion costs and expansion rewards, optimal expansion process. Typical research problems in knowledge management area are [14]: measuring e-business capability, e-business performance indicators, Habitual Domains approach to knowledge management. Typical research problems in decision theory area are [15]: consumer decision problem, generating learning sequences for decision makers, model of decision problems in fuzzy environments.

3.2 Creative Task vs. Consolidation Task – Competence Set Theory Perspective

In this article we propose a division of tasks given to students in their learning process into regular *consolidation tasks* and *creative tasks*. The creative task is aimed at stimulating a student to individually search and get new knowledge through completing untypical assignments.

We assume, that knowledge consolidation process is related to tasks, that can be completed individually by a student. Individual task completion is possible only if a student has collected so far in his/her learning process necessary skills and knowledge. The aim of consolidation tasks is to reinforce knowledge and assimilation of already acquired skills through their practical application in solving typical learning problems. Competences possessed by a student should reinforce as a result of completing similar tasks, that do not require any new knowledge or skill but need only proper application of what she or he has learned before. From the point of view of CS approach it corresponds with the situation where student's acquired competence set *Sk* contains all elements of the required competence set *Tr* of a given task. A student, that was given a consolidation task does not need to acquire any new competence (expand her/his acquired competence set *Sk*), she or he only need to reinforce her/his knowledge and skills (increase strengths of competences in her/his acquired competence set *Sk* to levels defined in required competence set *Tr* for a given task).

Creative tasks – are learning problems or projects that are more demanding than consolidation tasks and play completely different role in didactic process. These tasks are used by teachers or tutors to prepare students to individually solve learning problems and to learn how to manage unfamiliar and atypical situations. The main idea laying behind creative tasks is individual searching for solutions through discovering new knowledge and developing new skills. Using CS vocabulary creative task is a task for which acquired competence set *Sk* of a student challenging a problem does not contain all competences defined in required competence set *Tr* for this problem.

One of the main problems in assigning creative tasks to students is proper selection of difficulty level for each person. In case if difficulty level is too hard, a very demanding task can discourage a student from individual looking for a solution and in effect does not develop student's creativity. Whereas task that are too easy may have weak stimulating effect for student's creativity and development. It is decision of tutor to choose proper difficulty level of tasks selected for students.

The significant benefit of applying CS to learning task selection is the possibility of quantitative evaluation of subjective difficulty level of tasks assigned to students in the learning process. This is achieved by employing models and computational methods provided by CS theory, namely the methods for cost analysis of the competence set expansion process. These methods provides exact quantitative measures of difference between competences of a person and competences required to complete a learning task by this person. Measures of competence difference or in other words – measures of lack of competence can be used as measures of difficulty of both creative and consolidation tasks.

Most of publications on CS analysis use cost functions expressing the difference in competence in financial dimension, which requires assuming certain cost ratios. In the context of the problem discussed in this article information about exact financial cost of competence expansion is not required, because a tutor selecting task to students needs only some universal quantitative scale that enables him to assess difficulty level of tasks selected to a student taking into account her/his competence set *Sk*. Thus, the authors of the article propose simpler form of the cost function that do not requires assuming any cost ratios. This function is denoted $d(Sk(P),Tr(E))$ and gives measure of person's *P* competence deficiency to solve task *E*. The value of this function can be computed using different methods, for example [13], [15], [16].

Apart from quantitative competence modeling and processing, designing a complete model of the information system for competence-based individual learning task selection requires solving several other important issues, namely: competence data acquisition, competence data storing, information system structure and data flows etc. These topics are covered in the Section 4 of the article.

3.3 The Problem of Individual Learning Task Selection

In section 3.2 it was presented, that the decision about learning task selection to a student involved in learning process can be supported by providing information about student's current competence and competence requirements defined for learning tasks. The problem can be formalized in the following way:

$P = \{p_i\}$ - set of persons (students) involved in learning process

$S = \{s_j\}$ - set of learning tasks (or projects)

$Sk(p_i) = Sk_i$ - acquired competence set of a person p_i

$Tr(s_j) = Tr_j$ - required competence set of a learning task s_j

$d(Sk(p_i),Tr(s_j)) = d_{ij}$ - person's p_i competence deficiency to complete task s_j

$d_{\min}^{C}$ - minimum competence deficiency for creative tasks
$d_{\max}^{C}$ - maximum competence deficiency for creative tasks
$d_{\min}^{N}$ - minimum competence deficiency for consolidation tasks
$d_{\max}^{N}$ - maximum competence deficiency for consolidation tasks
$C(p_i) = C_i = \{s_n : s_n \in S; Tr_n \setminus Sk_i \neq \varnothing; d_{\min}^{C} < d(Sk_i, Tr_n) < d_{\max}^{C}\}$ - set of creative tasks of person p_i
$N(p_i) = N_i = \{s_k : s_k \in S; Tr_k \setminus Sk_i = \varnothing; d_{\min}^{N} < d(Sk_i, Tr_k) < d_{\max}^{N}\}$ - set of consolidation tasks of person p_i

It can be seen from the formalization, that obtaining the sets of consolidation and creative tasks requires from tutor to assume certain thresholds for competence deficiency, that selects tasks only on the specified difficulty level.

3.4 The Knowledge Network Development Issue

The goal of knowledge network is to create a relationship between students in order to increase creative potential on the group's level. In the framework of knowledge network the sub-groups of students are formed. The common competence(s) (with different level of fuzzy competence strength) is main characteristics of the sub-groups. Every student facing a problem can be supported by the other student, who already obtained competence(s) required to solve this problem. The sub-groups allow to achieve synergetic learning process maintained in collaboration environment. In every sub-groups some roles are attributed to the students e.g. mentor, coordinator, participant.

In order to illustrate this idea assume, that a student p_x was assigned a learning task s_y. According to definitions in sections 3.2 and 3.3 the student has certain amount of competence deficiency $d(Sk(p_x),Tr(s_y)) > 0$, that need to be supplemented in order to complete the given task. The student can make his work more effective by taking advice from other students, who may have bigger experience.

The collaboration tool build in learning management system could look for students and associate them in groups according to their learning needs and competence requirements. Therefore, the collaboration tool could suggest student s_y to contact student from one or all of the following groups:

1. Group of students, who already completed learning task s_y:

$$G_1 = \{p_m : p_m \in P; m \neq x; G_1 \subset P; d(Sk(p_m), Tr(s_y)) = 0\}$$

2. Group of students, who have lower competence deficiency towards task s_y:

$$G_2 = \{p_r : p_r \in P; r \neq x; G_2 \subset P; d(Sk(p_r), Tr(s_y)) < d(Sk(p_x), Tr(s_y))\}$$

3. Group of students having competences required by student p_x:

$$G_3 = \{p_w : p_w \in P; w \neq x; G_1 \subset P; Tr(s_y) \setminus Sk(p_x) \cap Sk(p_w) \neq \varnothing\}$$

Encouraging students to work together in groups and to exchange their experience may have positive influence on pace and quality of learning. The collaboration tool having access to information about students competence sets could initiate social networks that may play supportive role to classical learning methods, especially in e-learning environment.

4 Information System Model for Competence-Based Learning Task Selection and Collaboration Support

4.1 Personal Competence Profile

Previous sections of the article has presented CS theory as theoretical background for competence modeling in competence-based information system. However, applying theory into working system requires additional work involving software design and implementation. In order to successfully implement analytical methods it is necessary to elaborate suitable data structures covering all aspects of the theory and the approach used for data modeling.

The recent research initiatives on competence provide more and more elaborated competence data models. For example TENCompetence project group developed the TENCompetence Domain Model (TCDM) covering many important issues related to the notion of competence (www.tencompetence.org). TCDM provides data structure built according TENCompetence proprietary definition of competence, which defines this term as “effective performance within a domain/context at different levels of proficiency”. TCDM can been seen as general framework model of competence, that facilitates software development by providing detailed data structure model and data interchange by assuming data export and import in XML metadata formats like IEEE RDC (http://ltsc.ieee.org/wg20/) and HR-XML (www.hr-xml.org).

In [8] you can find the integration of the TCDM with CS approach and the data model that adjusts TCDM to requirements of CS analytical modeling methods. This approach was implemented as Java object library called Competence Object Library (COL) and can be used as component for rapid application development.

In case of the system for competence-based individual learning task selection the COL, thanks to implementation of CS methods can be directly used to implement necessary processing and storing of Personal Competence Profiles (PCPs). The main idea behind PCP is to keep a record of every competence obtained during a student lifecycle.

4.2 Competence Acquisition Tracking

Competence-based learning management would be impossible without detailed and up-to-date information about competence of students involved in learning process. In the Section 4.1 it was presented how to model and store students’ PCPs in a database. The complete information about competence important in learning process can be collected from three sources: (i) enrollment analysis, (ii) continuous learning progress evaluation, (iii) individual work evaluation.

The proposed system is able to create and maintain knowledge network, which support creative process, based on the students’ competence profiles. Moreover, the open nature of the information system make possible to use student’s cognitive characteristics as well. Social aspect of the knowledge network is supported by the predefined role. Every student has some system's established obligation to other students.

4.3 Model Description

The concept of the system is built around two functional databases: personal competence database and projects repository (Fig. 2). The first of them contains students competences stored as PCPs and updated continuously throughout student lifecycle. The second database stores descriptions of all learning tasks and projects that can be given to students in the learning process. These descriptions contains competence requirements generated with COL library.

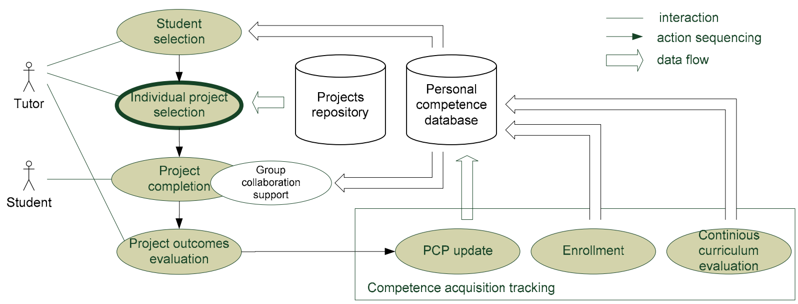


Fig. 2. Model of the system for competence-based individual project selection and collaboration support

The central item of the system is individual project selection component. The component supports a teacher in optimal task assignment taking into account students' PCPs and competence requirements for learning tasks taken from project repository database. Decision about task assignment is made by comparing students acquired competence sets from PCP with tasks required competence sets using methods of competence expansion cost analysis. The item "Group collaboration support" is responsible for knowledge network development and initiating students contacts and interactions basing on their current competences and learning assignments.

5 Conclusion

The student's creativity becomes important factor in e-learning systems development issue. Proposed system takes advantage of formal model of the competence set theory. The system, basing on the competence sets calculations, matches tasks to students to challenge them with creative and demanding learning problems. Moreover, the student's competence profile makes possible to develop creativity environment on the knowledge network level. Future work includes: student's cognitive characteristic support, method for student's competence profile development optimization, and competence set approach to examination.

References

1. Candy, L., Edmonds, E.A.: Supporting the creative user: a criteria based approach to interaction design. Design Studies 18(2), 185–194 (1997)
2. Chen, J., Chen, D., Li, Z.: The Analysis of Knowledge Network Efficiency in Industrial Clusters. In: International Seminar on Future Information Technology and Management Engineering, pp. 257–260 (2008)
3. Hewett, T.T.: Informing the design of computer-based environments to support creativity. International Journal of Human-Computer Studies 63(4-5), 383–409 (2005)
4. Nguyen, L., Shanks, G.: A framework for understanding creativity in requirements engineering. Information and Software Technology 51(3), 655–662 (2009)
5. Patru, M., Khvilon, E.: Open and distance learning: trends, policy and strategy considerations. UNESCO, kod: ED.2003/WS/50 (2002)
6. Plucker, J., Zabelina, D.: Creativity and interdisciplinarity: one creativity or many creativities? ZDM Mathematics Education 41(1-2), 5–11 (2009)
7. Plucker, J., Beghetto, R.: Why creativity is domain general, why it looks domain specific, and why the distinction does not matter. In: Sternberg, R.J., Grigorenko, E.L., Singer, J.L. (eds.) Creativity: From Potential to Realization, pp. 153–167. American Psychological Association, Washington (2004)
8. Różewski, P., Małachowski, B.: Competence Management In Knowledge-Based Organisation: Case Study Based On Higher Education Organisation. In: Karagiannis, D., Jin, Z. (eds.) KSEM 2009. LNCS (LNAI), vol. 5914, pp. 358–369. Springer, Heidelberg (2009)
9. Różewski, P., Ciszczyk, M.: Model of a collaboration environment for knowledge management in competence based learning. In: Nguyen, N.T., Kowalczyk, R., Chen, S.-M. (eds.) ICCCI 2009. LNCS (LNAI), vol. 5796, pp. 333–344. Springer, Heidelberg (2009)
10. Sampson, D., Fytros, D.: Competence Models in Technology-Enhanced Competence-Based Learning. In: Adelsberger, H.H., Kinshuk, Pawlowski, J.M., Sampson, D. (eds.) Handbook on Information Technologies for Education and Training, 2nd edn., pp. 155–177. Springer, Heidelberg (2008)
11. Sternberg, R.J.: Creativity or creativities? International Journal of Human-Computer Studies 63(4-5), 370–382 (2005)
12. Sriraman, B.: The Characteristics of Mathematical Creativity. Mathematics Edukator 14(1), 19–34 (2004)
13. Wang, H.-F., Wang, C.H.: Modeling of optimal expansion of a fuzzy competence set. International Transactions in Operational Research 5(5), 413–424 (1995)
14. Yu, P.L., Lai, T.C.: Knowledge Management, Habitual Domains, and Innovation Dynamics. In: Shi, Y., Xu, W., Chen, Z. (eds.) CASDMKM 2004. LNCS (LNAI), vol. 3327, pp. 11–21. Springer, Heidelberg (2005)
15. Yu, P.L., Zhang, D.: Optimal expansion of competence sets and decision support. Information Systems and Operational Research 30(2), 68–85 (1992)
16. Yu, P.L., Zhang, D.: A foundation for competence set analysis. Mathematical Social Sciences 20, 251–299 (1990)

A Blended Project-Based Learning Program on Embedded Software Design with Collaboration Support Tools*

Takashi Yukawa[1], Tomonori Iwazaki[1], Keisuke Ishida[1],
Hirotaka Takahashi[1], Yoshimi Fukumura[1],
Makoto Yamazaki[2], Naoki Hasegawa[3], and Hajime Miura[4]

[1] Nagaoka University of Technology
[2] Nagaoka National College of Technology
[3] Niigata Industrial Creation Organization
[4] Techno Holon Corporation

Abstract. The present paper proposes a blended learning program on project-based embedded software design. The paper also clarifies the requirements and design of collaboration support tools for the proposed program. The authors have been conducting research in an attempt to implement e-Learning technology for project-based learning (PBL) on the development of embedded software. We have created a new program that blends face-to-face and e-Learning classes and have developed a computer-supported collaborative learning environment. An integrated repository tool and a unified search tool have been proposed and reported in KES2009. The requirements for additional functions for the collaboration between learners were clarified through a trial of the learning activities using these tools. In the present paper, these requirements are described, and the design of tools that satisfy these requirements is proposed. The tools include a review support tool and an online whiteboard system. A blended learning program using the developed tools is also proposed. In the proposed blended learning program, more than 80% of learning units can be investigated with e-Learning, while face-to-face discussion sessions must be conducted at each juncture. A trial of the proposed program has been carried out and a questionnaire survey has been conducted to evaluate the learning effect of the program and the usability of the tools. The results suggest that the proposed program is feasible for PBL on embedded software design and that the tools facilitate collaborative activities between learners and are effective for enhancing design ability.

1 Introduction

Project-based learning (PBL) is intended to strengthen learners' abilities in design, teamwork, and communication through experiences developed in solving

* The present research is supported by a grant through the Strategic Information and Communications R&D Promotion Programme (SCOPE) from the Ministry of Internal Affairs and Communications, Japan.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 190–199, 2010.

practical problems as a team. As an example, PBL has become popular in engineering education because industry requires new university graduates to have engineering design abilities. However, there are some potential problems. Since group study and meetings are the key to PBL, learners need to spend a great amount of time in the classroom. Especially for part-time students, this can limit their opportunities to attend PBL-based courses. In addition, project-based learners must organize their experiences into systematic knowledge and insight in order to acquire the desired abilities. Otherwise, PBL becomes an ineffective and time-consuming process.

In order to address these problems, the authors have launched a project that implements collaborative e-Learning technology in PBL [1] and have proposed collaboration support tools [2]. The demands for additional functions for the collaboration between learners are clarified through the trial of an embedded software development program using these tools. The present paper proposes functions that satisfy these demands and describes the design of tools that achieve the proposed functions. These tools include a review support tool and an online whiteboard system. A blended learning program using these tools is also proposed, and a trial of the proposed program has been carried out. A questionnaire survey has been conducted in order to evaluate the learning effect of the program and the usability of the tools. The evaluation results are also reported in the present paper.

2 Background

In this section, research related to ICT support of PBL for the development of embedded software and the concept and progress of the present project are briefly introduced.

2.1 Related Research

The demand for the development of embedded software has increased because of the widespread use of microprocessors in various types of equipment. In particular, for Japan to remain globally competitive in the development of industrial products, a greater number of embedded systems engineers have to be trained.

In recent years, a number of training programs on embedded software have been established. In Niigata prefecture, the Niigata Industrial Creation Organization (NICO) and associated organizations have conducted training courses on embedded software since 2006. The NICO program incorporates a PBL course [3]. In a traditional PBL course, since a class cannot be formed unless all members of the team are present, time and/or spatial restrictions become an obstacle. In order to remove this obstacle, collaborative e-Learning technology in PBL should be implemented.

Research on collaborative learning with the support of ICT has become increasingly active, and so research on Computer Supported Collaborative

Learning (CSCL) has evolved [4,5,6,7]. In a CSCL environment, communication between a teacher and a learner and/or between learners, sharing documents and program codes created through the learning process, and the exchange of atmosphere (awareness) among learners are supported by a computer.

2.2 Project Overview

The authors have established a project for implementing CSCL technology for the PBL course on embedded software development. The final goal of the project is to achieve the combination of the following two objectives:

- to construct a blended learning program for the embedded software PBL, and
- to develop supporting software tools for the learning program.

Figure 1 shows the process of developing embedded software. Most existing training courses incorporating e-Learning technology focus on having learners acquire programming skills through a trial-and-error process, as shown on the right side of the figure.

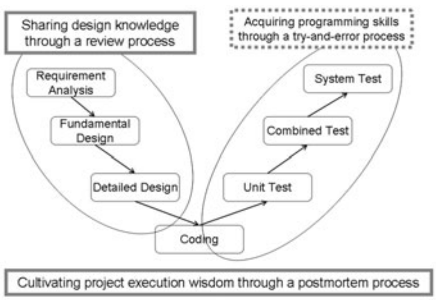


Fig. 1. Embedded Software Development Process

Without ignoring the right side of Fig. 1, the authors focus on sharing design knowledge through a review process, as shown on the left side of the figure, as well as the postmortem process. The postmortem process enables learners to organize their experiences into systematic knowledge about designing embedded software and to strengthen their insight for carrying out a project. For these purposes, collaborative learning technology is required to handle and process the information and knowledge contributed by the learners.

2.3 Progress of the Project

The objectives of the project were accomplished in the following steps:

1. The behaviors of learners were observed in an actual PBL course. Based on these observations, the possibilities of e-Learning support were considered.

2. Based on the findings obtained from the previous step, a PBL-based learning program for embedded software development supported by e-Learning was designed. The supporting computational environment (tools) was also developed.
3. The program and the support system were implemented in an actual course, and their effectiveness was evaluated.
4. Steps 2 and 3 were iterated to improve the tools and the program based on the results of the evaluation in Step 3.

The authors reported the first round of Steps 1 through 3 at the previous conference [2]. The report is reviewed briefly below.

Situations requiring support tools when learners cannot gather in a classroom for an embedded system PBL are:

1. discussion with other learners and teachers,
2. review of the designs of other learners, and
3. sharing design documents and program files.

To support learners in the above situations, the Integrated Repository Tool (IRT) was proposed and developed. The IRT is the integration of file storage and a bulletin board. It is basically a BBS that is able to store multiple files as attachments to a posted message. The IRT can be used as a simple asynchronous communication tool, and can also be used for sharing documents and programs (products) created through the learning process and for annotating the products. This would enable remote parties to perform cross and group reviews. The IRT is integrated in OpenSourceLMS [8].

A training program using the above tool was also proposed and a training trial was performed for evaluation. The program targeted the development of an electronic thermometer using the one-chip microcomputer board. The five-day course focused on the left half of the development process shown in Fig. 1. After completing the trial program, a questionnaire survey was conducted in order to evaluate the PBL program and the IRT. The results are shown in Fig. 2. As shown in the figure, each learner provided a positive response regarding the effectiveness of the cross and group reviews. Although the IRT was reported to be easy to use by more than half of the learners, some learners felt it was inconvenient. In particular, for the cross review processes, most of the learners complained about the insufficient functionality of the IRT.

3 Additional Requirements for the Collaboration Support Tools

Cross review processes are essential for learning embedded software development. As clarified through the trial described in the previous section, additional functionalities are required for smoother communication in the cross review

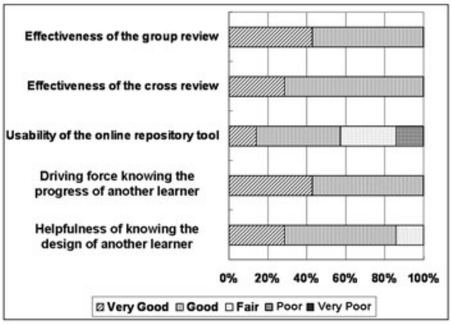


Fig. 2. Evaluation Results of the Trial Training Course

processes. For extracting the functional requirements, a procedure for a review process using the IRT is examined. The review is conducted with the following steps.

1. A reviewee uploads his/her design document files on the IRT.
2. A reviewer downloads the files to his/her local PC and reviews the files. The reviewer and the reviewee then discuss issues in the design documents using the BBS function in the IRT.
3. The reviewee revises the design documents based on the discussion, and then uploads the revised document files.
4. The reviewer downloads the revised files and checks to see whether his/her comments are appropriately reflected in the design documents. If the revised documents are still considered to be unsatisfactory, then the reviewer identifies the unsatisfactory areas and discusses them with the reviewee.
5. Steps 4 and 5 are iterated until the quality of the design reaches a certain level.

The following are considerations related to the procedure using the IRT.

1. The BBS function in the IRT allows text-only discussion. However, drawings should useable in the discussion, because the design documents comprise numerous diagrams.
2. Checking the revised documents is a heavy burden for a reviewer because he/she must recognize changes in the documents through visual comparison.

In response to the first consideration, an online whiteboard tool with a chat function and a playback function (OWB) is proposed. Although several online whiteboard tools have been developed, these tools can only be used for synchronous (real-time) discussion. In our learning program, cross review processes should able to be performed asynchronously. Adding a chat function in addition to a playback function, which can replay the chat contents sequentially as well as display previous drawings on the whiteboard, enables asynchronous discussion to be achieved using drawings. The OWB also has a function whereby

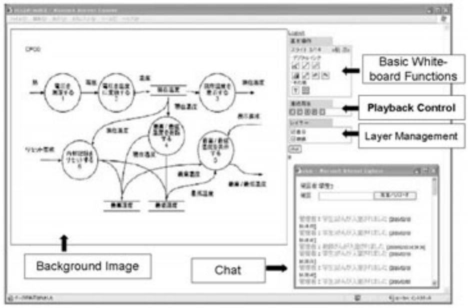


Fig. 3. Screenshot of the Online Whiteboard Tool

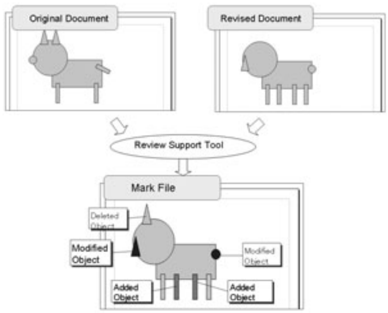


Fig. 4. Function of the Review Support Tool

users can upload an image file and display the file as a background image of the whiteboard.

A screenshot of the OWB is shown in Figure 3 A scenario for using the OWB in an asynchronous discussion is described below.

A learner, learner A, invokes the OWB, uploads a diagram in his/her design document and draws comments. Learner A also writes text comments in the chat area. After drawing and posting comments, learner A can close the OWB. At a different time, another learner, learner B, invokes the OWB. Learner B can use the playback function to read the comments and add drawings and text messages in reply to the comments of learner A. Learner A invokes the OWB again and reads the comments of learner B using the playback function. Learner A then draws and writes a reply to the comments of learner B. The discussion is traced using the playback function and several rounds of replies with drawings and text messages are sent.

A Review Support Tool (RST) is proposed to address the second consideration. The RST displays differences between an original design document and a revised design document. There are several tools that can clarify the differences between text files. However, differences between diagrams must also be displayed in this learning program. Figure 4 demonstrates the function of the RST.

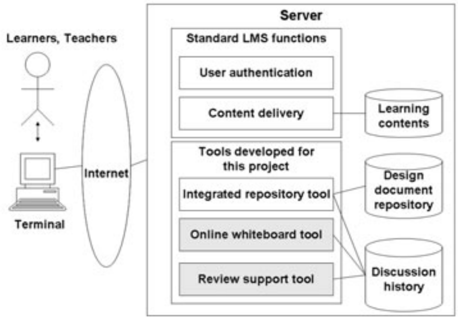


Fig. 5. Overview of the e-Learning Environment for Embedded Systems PBL

For detecting added, deleted, and modified objects, the tool uses the edit graphs algorithm [9].

These tools are intended to be used for discussions on design documents for embedded software development including data-flow diagrams, hierarchical structure diagrams, and module diagrams. Integration of these tools with the existing LMS (OpenSourceLMS), the e-Learning environment in this project is shown in Figure 5. In next section, a blended learning program that assumes the use of this environment is proposed.

4 A Blended Learning Program on Project-Based Embedded Software Development

Postulating the use of the proposed e-Learning environment, a blended learning program on project-based embedded software development is constructed as shown in Table 1. The table lists the schedule, the learning modes, and the tools usages for each learning unit or activity. The learning modes include synchronous (sync.) or asynchronous (async.), face-to-face (f2f) or distant (dist.), and individual (indiv.), group of learners (group) or all learners (all).

The blended learning program requires 21 days, whereas the equivalent face-to-face program requires only five days. This is because most learning units and activities are conducted asynchronously, which requires a time allowance. Learners are assumed to be part-time students and have workplace responsibilities. Therefore, learners have limited available time each day for the learning program.

The authors do not propose the learning program as a full e-Learning program, because face-to-face communication is assumed to be important for sharing common perception in learning subjects and achieving a uniform level of understanding. Therefore, the learning program includes a face-to-face session at the end of each phase.

Table 1. Blended Learning Program on Project-based Embedded Software Development and Supporting Tool Applicability to the Learning Units

Learning Unit/Activity	Day	Learning Modes	IRT	OWB	RST
Requirement Specification					
a lecture on a development process	1	async./dist./indiv.			
a lecture on device drivers	1	async./dist./indiv.			
a lecture on requirement specification	1	async./dist./indiv.			
a task description	2	async./dist./indiv.			
writing an event list	2	async./dist./indiv.	x		
cross review	3	sync./dist./group	x	x	x
general review	3	async./dist./all	x		
writing a content diagram	4	async./dist./indiv.	x		
cross review	5	async./dist./group	x	x	x
general review	4	sync./f2f/all			
Structured Design					
a lecture on structured design	6–7	async./dist./indiv.			
a lecture on data flow and state transition diagrams	6–7	async./dist./indiv.			
writing a data flow diagram	6–7	async./dist./indiv.	x		
cross review	8	sync./dist./group	x	x	x
general review	8	async./dist./all	x		
a lecture on a module structure diagram	9–10	async./dist./indiv.			
writing a module structure diagram	9–10	async./dist./indiv.	x		
cross review	11	sync./dist./group	x	x	x
general review	11	async./dist./all	x		
a lecture on module specifications	12–13	async./dist./indiv.			
a lecture on a timer interrupt	12–13	async./dist./indiv.			
writing module specifications	12–13	async./dist./indiv.	x		
cross review	14	async./dist./group	x	x	x
general review	14	sync./f2f/all			
Test Design					
a lecture on test design	15–16	async./dist./indiv.			
writing system test specifications	15–16	async./dist./indiv.	x		
writing linkage test specifications	15–16	async./dist./indiv.	x		
writing unit test specifications	15–16	async./dist./indiv.	x		
cross review	17	sync./dist./group	x	x	x
general review	17	sync./f2f/all			
Coding, Debugging, and Testing					
coding, debugging, and testing	18–20	async./dist./indiv.	x		
Postmortem					
individual retrospect	21	async./dist./indiv.	x		
overall retrospect	21	sync./dist./all	x	x	
summarization	21	sync./f2f/all	x	x	x

5 A Trial and Evaluation of the Learning Program

A trial of the proposed learning program involving five learners has been conducted. The learners engaged in a face-to-face learning program on embedded software development before participating in the trial. The assignment of the face-to-face program was different from the assignment of this trial. After completing the trial, a questionnaire survey was conducted to evaluate the blended learning program and the supporting tools for the reviewing processes.

The results of the survey are shown in Figure 6. As shown in the figure, the learners provided positive responses to most questions. The question on the effectiveness of the supporting tools received positive responses (very good and good) from 80% of the learners and received no negative responses. On the other hand, the question on reduction of burden for the review with the supporting tools received negative responses from 20% of the learners, whereas 40% of the learners provided positive responses. This suggests that the usability of the supporting tools should be improved. All trial learners reported that the review quality of this program is equivalent to that of the face-to-face program. Although the blended learning program takes longer to complete, the learning quality is not far behind that of the face-to-face program. Positive responses regarding the effectiveness of the face-to-face sessions in the program were reported by all learners, which suggests that blending e-Learning and face-to-face achieves better learning effects.

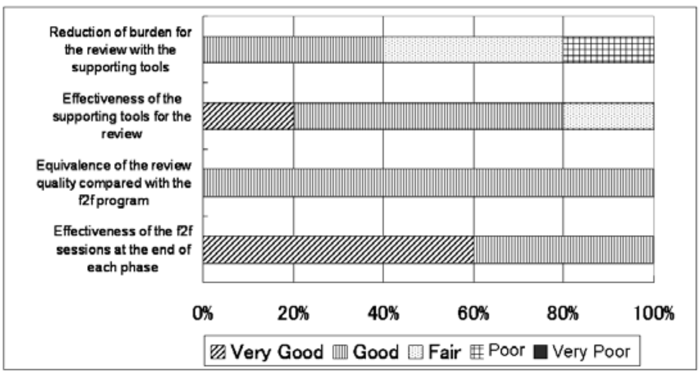


Fig. 6. Evaluation Results of the Proposed Learning Program

6 Conclusions

The present paper finalizes a research project to implement e-Learning technology in PBL for the development of embedded software. In addition to the Integrated Repository Tool (IRT), which was reported last year, an Online Whiteboard (OWB) Tool and Review Support Tools (RST) were proposed in order to achieve smooth communication in the review processes.

The OWB is a combination of an online whiteboard and a chat function that can playback previous drawings and communications. The OWB can be used for both synchronous communication and asynchronous discussion. The RST is a tool for displaying changes between an original document and a revised document. The RST can reveal differences in both the text and diagrams in documents.

A blended learning program using the proposed tools is also proposed and implemented. The results of the questionnaire survey after the actual trial of the learning program suggest that the learning effect of the blended learning program is equivalent to that of the face-to-face program. Although the blended learning program takes longer to complete, it can provide learners who are not able attend a classroom due to geographical and/or temporal limitations with the opportunity to learn embedded software PBL at a quality level similar to that provided by a face-to-face course.

The evaluation results suggested that seamless operation of the tools would reduce the time to finish the learning program. Improvement of coupling of the tools and the LMS is an area for future study.

References

1. Yukawa, T., Takahashi, H., Fukumura, Y., Yamazaki, M., Miyazaki, T., Yano, S., Takeuchi, A., Miura, H., Hasegawa, N.: Implementing e-learning technology for project-based learning for the development of embedded software. In: Proceedings of 20th Annual Conference of the Society for Information Technology and Teacher Education, pp. 2208–2212 (2009)
2. Yukawa, T., Takahashi, H., Fukumura, Y., Yamazaki, M., Miyazaki, T., Yano, S., Takeuchi, A., Miura, H., Hasegawa, N.: Online collaboration support tools for project-based learning of embedded software design. In: Velásquez, J.D., Ríos, S.A., Howlett, R.J., Jain, L.C. (eds.) Knowledge-Based and Intelligent Information and Engineering Systems. LNCS, vol. 5712, pp. 531–538. Springer, Heidelberg (2009)
3. Niigata Industrial Creation Organization: Nico advanced training programs (2009) (in Japanese), `http://www.nico.or.jp/modules/list2/kouza_kensyu/koudoit.html`
4. Alavi, M.: Computer-mediated collaborative learning: An empirical evaluation. MIS Quarterly: Management Information Systems 18(2), 159–174 (1994)
5. Brandon, D.P., Hollingshead, A.B.: Collaborative learning and computer-supported groups. Communication Education 48(2), 109–126 (1999)
6. Gillet, D., Nguyen-Ngoc, A.-V., Rekik, Y.: Collaborative web-based experimentation in flexible engineering education. IEEE Transactions on Education 48(4), 696–704 (2005)
7. Kojiri, T., Kayama, M., Tamura, Y., Har, K., Itoh, Y.: CSCL and support technology. JSiSE Journal 23(4), 209–221 (2006) (in Japanese)
8. eLearning Consortium: Scorm2004 engine (2004) (in Japanese), `http://www.elc.or.jp/cgi-bin/scorm_engin/lms/index-scorm.html`
9. Myers, E.W.: An O(ND) difference algorithm and its variations. Algorithmixa 1, 251–266 (1986)

Multilingual Discussion in Metaverse among Students from the USA, Korea and Japan

Hideyuki Kanematsu[1], Yoshimi Fukumura[2], Dana M. Barry[3], So Young Sohn[4], and Ryosuke Taguchi[5]

[1] Dept. MS & E., Suzuka National College of Technology, Shiroko-cho, Suzuka, Mie 510-0294, Japan
kanemats@mse.suzuka-ct.ac.jp

[2] Dept. Management & Information Systems Science, Nagaoka University of Technology, 1603-1, Kamitomioka-cho, Nagaoka, Niigata, Japan
fukumura@oberon.nagaokaut.ac.jp

[3] CAMP, Clarkson University, P.O. Box 5665, Potsdam, NY, USA
dmbarry@clarkson.edu

[4] Dept. Information & Industrial Engineering, Yonsei University, 134 Sinchon-dong, Seodaemun-gu, Seoul, 120-749, Korea
sohns@yonsei.ac.kr

[5] Graduate Student, Nagaoka University of Technology, 1603-1, Kamitomioka-cho, Nagaoka, Niigata, Japan
jp46bb@bma.biglobe.ne.jp

Abstract. E-learning continues to be used as an educational tool all over the world. However, the typical e-learning systems are sometimes not enough for students to feel a close connection with others. Now engineering education needs to be carried out from a global viewpoint, since there are lots of opportunities for international collaboration in higher education organizations. In this study, we investigated multilingual discussion in Metaverse and its possibility to enhance the skills for international discussion and international characteristics. Students from Edwards-Knox high school in the USA, graduate students from Yonsei University, Korea, and Nagaoka University of Technology (NUT), Japan, participated in this discussion held on the virtual island (owned by NUT) of Second Life. A language grid system was incorporated into the Metaverse system. This made it possible for the discussion to be carried out with and without the translation system. These sessions were carried out and compared with each other. Also the effectiveness of the international discussion project in Metaverse (with the language grid system) was discussed.

Keywords: Multilingual discussion, e-learning, Metaverse, Language grid, Second life.

1 Introduction

Today we live in a trend of globalization. This trend has the power and potential to rapidly change many aspects of our daily lives. If engineering could be defined as the

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 200–209, 2010.

discipline to solve daily problems in our society, then engineering education should be designed to deal with the trend of globalization. Many innovative approaches can be mentioned. E-learning is one of these approaches. This educational tool makes it possible for us to overcome the geographical and temporal obstacles[1-6]. One can exchange messages and information with others by the e-learning process asynchronously. It enables us to carry out discussions, even though we might be far apart from each other geographically and have significant time differences. Also when language barriers exist, one can avoid the direct conversation and exchange messages, by looking up words in language dictionaries. Individuals may be able to relax and be confident, by avoiding the face-to-face contacts. When one needs international communication beyond the differences of languages and cultures, e-learning can make a significant contribution.

However, the typical e-learning systems are sometimes not enough for students to feel the close connection with others. Here they do not feel that they are really represented because the information is only delivered through characters. Currently there seems to be lots of opportunities for international collaboration in higher education organizations. For example, students often go abroad not only for overseas education, but also for internships, co-op education, and for a chance to work in foreign countries. Such opportunities give those students some concrete skills for international collaboration. At the same time it should be mentioned that higher education organizations very often have foreign students. In such cases, the foreign students need a support system with services such as tutoring, counseling, opportunities to makeup classes, etc. For this type of situation, e-learning should be useful due to its asynchronous and anonymous characteristics. However, it sometimes fails to reach the goal, due to its lack of reality and the language barrier. Therefore, we decided to utilize Metaverse[7-10] with the translation function to make the multilingual discussion smooth, so that students could communicate with each other in their own languages. Metaverse is one of the social media characterized as a synchronous, 3D virtual world resembling reality[11-16]. In this study, we introduced the language grid system into Metaverse and investigated the multilingual discussion there as the first step for the application to virtual multilingual Problem Based Learning or some synchronous e-learning classes in the near future.

2 Experimental

2.1 Discussion Environment in Metaverse

Metaverse can be defined as a virtual three dimensional space where the character called avatar is active instead of a person. As a commercial service of Metaverse, Second Life (of Linden Research Inc. in San Francisco, California, USA) is the most well known. In this study, the Second Life provided by Linden Research Inc. was used. Nagaoka University of Technology (NUT), to which some of the authors belong, owns a virtual island in Second Life and has built different kinds of classrooms where some e-learning classes are already available. For this project, three new classrooms were built, as already described in some previous papers[17-22].

Each of them contains the same structures and facilities. They all have two big tables with chairs around them to carry out discussions. Each classroom also includes a small blue item on the wall that records all of the chatting. The record of conversations can be sent to a server and written out in the HTML language.

2.2 Multilingual Discussion

Participants in the project are from the USA, Korea and Japan. There are three students from each Country, giving a total of nine. The students were divided into teams with three members each. Each of the three teams included a student from Japan, a student from the USA, and a student from Korea. Each team met twice separately (one team per classroom) in the virtual island at an interval of several weeks, so that the results for both sessions would not affect mutually. Here they discussed general topics for an hour. Two discussion sessions were held for the international discussion. Each team was assigned in advance a specific and different topic to discuss for each session. This project focused on the possibility of the smooth multilingual discussion as the preliminary investigation for e-learning classes such as PBL. Therefore, the topics were chosen from daily lives matters including songs, sports, and movies with which any students could grapple easily.

Fig. 1. Language chairs for a language grid system

For the first session, one of the teachers (by way of an avatar) gave an opening message before the students entered their classrooms according to their assigned topics. A topic title was displayed in front of each classroom, so the students could easily find their appropriate classroom. Once entering the building, the students sat in chairs around a table. Then they began the discussions about the specific topic that was assigned to them (each team had a specific topic). The discussion began as follows.

Songs classroom: What is the most popular song in your country?
Movies classroom: What is the most popular movie in your country?
Sports classroom: What is the most popular sport in your country?
Students discussed these topics in English. The session continued for about one hour.

For the second session, the process was almost the same as that for the first session. Each team met again in the virtual island of NUT two weeks later. After the opening message, they entered their appropriate classrooms to discuss their second and different topics (sports, movies, or songs) assigned to them. However, this time discussions were carried out in their own languages. For the second session, the language grid system was introduced. Visually, it was realized as a language chair. Students sat down (around a table) on specially prepared language chairs. The US students sat on green chairs, Korean students sat on red chairs and the Japanese students sat on yellow chairs. The language chairs are shown in Fig.1. The language system was introduced and arranged, so that students chatted by using their own languages (English, Japanese, and Korean)[23,24]. Here each language statement that was chatted would be translated into the other two languages simultaneously.

After each session, students answered questionnaires. From the two questionnaires and recorded conversations, the record was analyzed from various viewpoints.

3 Results and Discussion

3.1 The First Session in English

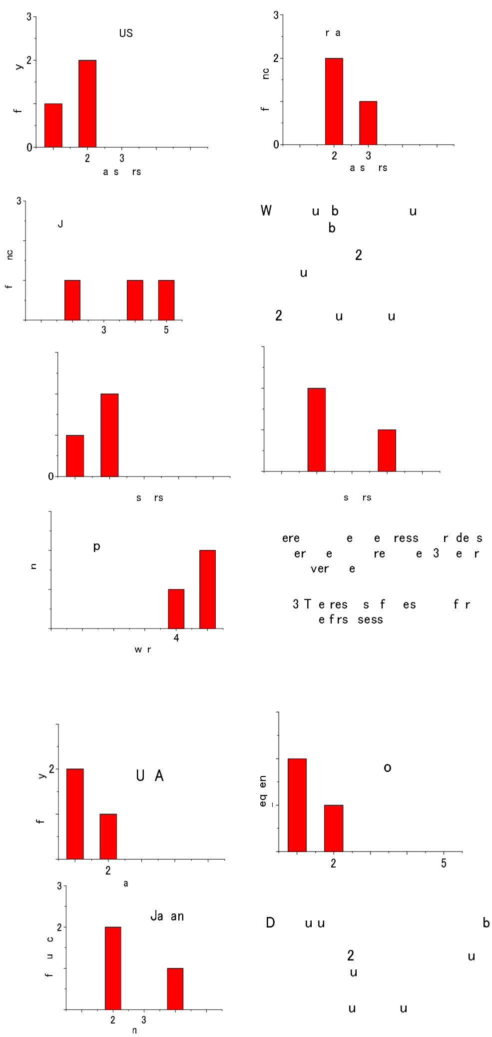
The first session was carried out on a day in January 2010. The gathering time was set in advance, considering the time differences among the three countries. After the opening announcement, students entered their classrooms and the discussions began. An example of the conversation is provided below:

Table 1. An example of the chatting for the first session

A: a US student, B: a Korean student, C: a Japanese student
Classroom: Songs
B: hi, all
C: hi
A: Hi
B: good to see you guys. I am peterelli from Korea.
A: My name is XXX and I am from the United States of America.
C: I am XXX from Japan.
B: oh I see
B: what do we discuss?
A: Currently in the United States the most popular song is Bad Romance by Lady GAGA.
B: oh, I know that song.
A: What is the most popular song in your countries?
B: I enjoy that song these days.
C: O

As mentioned above, the chatting began with the self introduction (by each team member). Then one of the members introduced the topic and the discussion began. After one hour passed, the discussion was stopped and the students gathered again in the yard (in front of the classrooms). They broke up according to the final announcement by the teachers. After the students logged off of their computers, they submitted the questionnaires to the teachers by email. The questionnaire for the first session is as follows.

Table 2. The questionnaire for the first session



Please write the name of the Country that you are representing on the line provided.

____

<u>#1: Were you able to communicate with the other members?</u>

1. Very well 2. Pretty well 3. Neutral 4. Not very well 5. Not at all

<u>#2: Were you able to express your ideas?</u>

1. Very well, 2. Pretty well 3. Neutral 4. Not very well 5. Not at all

<u>#3: Did you understand the other members' ideas?</u>

1. Very well 2. Pretty well 3. Neutral 4. Not very well 5. Not at all

<u>#4: Do you have experience in Second Life?</u>

1. Very much 2. Pretty much 3. Neutral 4. Not very much 5. Not at all

<u>#5: Do you have experience chatting by computer?</u>

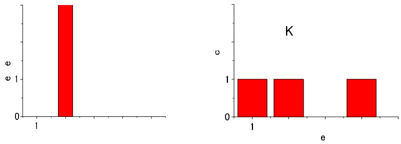
1. Very much 2. Pretty much 3. Neutral 4. Not very much 5. Not at all

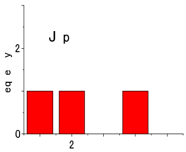
<u>#6: Did you feel more comfortable in Second Life than in other communication systems?</u>

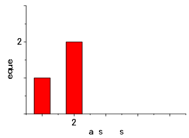
1. Very much 2. Pretty much 3. Neutral 4. Not very much 5. Not at all

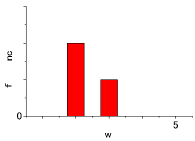
<u>#7: If you found any problems in NUT island (the Japanese island), then please feel free to write them in the space provided.</u>

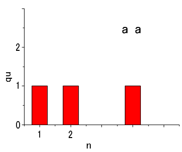
<u>#8: Please feel free to write down your impression of this session in SL.</u>











The results of question #1 are summarized in Fig.2. While the US students felt the communication went pretty well, there were some Korean and Japanese students who did not feel their communication was so well.

Fig.3 shows the results of question # 2. While the US students expressed their ideas very easily, Asian students felt it difficult to express their ideas to some extent. Particularly, the tendency for Japanese students was very remarkable.

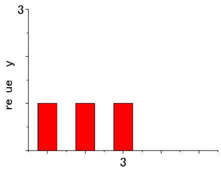
Fig. 4 shows the results of question #3. Most of participants including the Japanese students who found it hard to express their ideas could understand other members' ideas pretty well. From all of these results (Figs. 2, 3, 4), Asian students, particularly Japanese students felt that they could not express their ideas fully, even though they could understand others' ideas. Actually, most of Japanese and some of Korean students confessed in question #8 that they felt frustrated for the chatting in English. As for question # 7, most of students did not feel any problems for the system itself. However, only one participant (Japanese) wrote that it should have been more clearly specified who made remarks, e.g. by using sounds, etc.

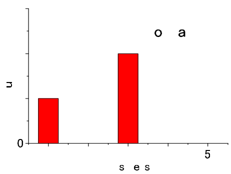
3.2 The Second Session in Their Own Languages

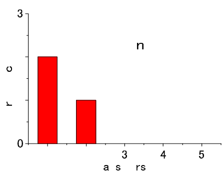
Two weeks after the first session, the second session was carried out in the same virtual place. After the short opening announcement, students entered their classrooms that were assigned in advance. The group members were quite the same with those in the first session. However, their topics were different from those of the first session, but once again selected from the same group (songs, movies and sports). The chatting at the beginning for the same group in Table 1 was going on in the following Table 3, as for the second session.

Table 3. An example of chatting for the second session

U u r u u
r m r
こ . 안 십니
w r y u? こ どうで ? 안 십니 ?
です 니다
h r すば し 다.
안 . こ .
Wh r r u r i y ur u ry? ど ー で ー
で ?
라 적 ?
だ す
라고 니다.
장 적입니다.
も です
i r m .
i u r i
m ri h u h m ri
i mu h m r u r
もっ ずっ
ーで ども
で ーで
미국 적 미국 적 다.
국 장 기 니다. でも も す
고 족 공 니
本でも です i ry u r i .
본 기 니다.







Did d r t d t t r r
id
1 r Pr tt tr
4 t r t t
i r t q ti r t
d i

Fig. 8. Translation for chatting

The same amount of the conversation with that for the first session was done in several minutes for the second session. The Japanese members became very active in this case, being compared with the former case. It suggests that the translation system made the conversation by chatting very lively.

Fig. 5 shows the results of question #1 for the second session. Even though there were still some members who were not satisfied with the communication among Asian students, communication became lively to some extent, in comparison with Fig. 2.

Fig.6 shows the results of question #2 for the second session. While US students showed their positive attitude for expressing ideas very clearly, Asian students showed their feeling of accomplishment for that a little bit higher than in the first session. From the viewpoint, the conversation seems to be improved to some extent.

Fig.7 shows the results of question #3 for the second session. The results for US and Korean students suggest a slightly negative evaluation for the holistic impression about the conversation, in comparison with Fig. 4. It could be attributed to that the English-Korean dictionary was not good enough to satisfy the Korean students. Actually, some of them had that impression after the second session, according to the survey question #8. However, the evaluation for Japanese students was improved remarkably, being compared with Fig.4. It may depend on the good quality of the Japanese – English dictionary. However, the Japanese students' positive conversation improved the entire conversation to some extent. These analyses were carried out in detail by conversation analysis.

As for the system itself, there were some other assignments for negative aspects by teachers and students who participated to the discussion. For example, one of the students pointed out that the system should have notified other members for every remark by sound or some other method. And one of us teachers felt that the current translation displays would be too annoying, since the multi-lingual conversations among many countries might need the corresponding numbers of translation pop-up windows shown in Fig.8. The detailed current system was proposed by Yoshino and Ikenobu firstly. They described the advantage and disadvantage[23]. However, the system should be changed, depending on the purposes and conditions. All of these wish lists or criticisms could be written in the question 7, 8 or in other follow-on questionnaires. The results were analyzed and will be used as feedback to improve the system later.

4 Conclusions

Multilingual discussion in Metaverse was carried out by the incorporation of a language grid system into the Second Life system. This allowed virtual discussions to be held among US, Korea and Japanese students. The introduction of the translation system made students more communicable as a whole. The tendency was the most remarkable in Japanese students. This new tool is expected to be developed for some educational trials including the support system for foreign students. However, the current system should be improved much further for the practical educational use in the future. The improvement of the quality of dictionary or their appropriate choice could be mentioned as the concrete examples.

References

1. Levine, A., Sun, J.C.: Barriers to Distance Education. Distributed Education: Challenges, Choices, and a New Environment, vol. 6. American Council on Education Center for Policy Analysis (2002)
2. Eaton, J.S.: Maintaining the Delicate Balance: Distance Learning, Higher Education Accreditation, and the Politics of Self-Regulation, p. 39. American Council on Education Center for Policy Analysis, Washington (2002)
3. Mehrotra, C.M., Hollister, C.D., McGahey, L.: Distance Learning - Principles for Effective Design, Delivery, and Evaluation. Distributed Education: Challenges, Choices, and a New Environment, vol. 2, p. 242. Saga Publications (2001)
4. Berg, G.A.: Why Distance Learning. ACE/Praeger Series on Higher Education. American Council on Education and Praeger Publishers (2002)
5. Yates, C., Bradley, J.: Basic Education at a Distance. In: Perraton, H. (ed.) World Review of Distance Education and Open Learning, vol. 2. The Commonwealth of Learning, London (2000)
6. Meyer, K.A.: Quality in Distance Education - Focus on On-Line Learning. In: Kezar, A.J. (ed.) ASHE-ERIC Higher Education Report, vol. 29. Jossey-Bass A Wiley Imprint, New Jersey (2002)
7. Kanematsu, H., Nakayama, S., Kobayashi, T.: Patent Education by Using PBL. In: Proceeding of Annual Meeting of Japanese Society for Engineering Education, pp. 239–242. The Japanese Society for Engineering Education, JSEE (2003)
8. Kanematsu, H., Tanaka, T., Hirai, N., Lee, J.: Introduction of Patent Education by PBL into Laboratory's Seminar and It's Educational Effect. Bulletin of Tokai Kagaku Kougyoukai (The Society of Chemical Industry in Midland of Japan) 239, 1–8 (2003)
9. Kanematsu, H., Barry, D.M.: World First Marslink Mission As Educational Material for Design Skills - As a PBL Class. In: Proceeding of JSEE Annual Meeting, pp. 250–251 (2007)
10. Kuroda, D., Kanematsu, H.: PBL Type Creative Engineering Class Using Bicycle in the Department of Materials Science and Engineering. In: Proceeding of Teachers' Workshop 2006 by Institute of National Colleges of Technology, Japan, pp. 317–318 (2007)
11. De Lucia, A., Francese, R., Passero, I., Tortora, G.: Development and evaluation of a virtual campus on Second Life: The case of SecondDMI. Computers & Education 52(1), 220–233 (2009)
12. Jarmon, L., Traphagan, T., Mayrath, M., Trivedi, A.: Virtual world teaching, experiential learning, and assessment: An interdisciplinary communication course in Second Life. Computer & Education 53, 169–182 (2009)
13. Kaplan, A.M., Haenlein, M.: The fairyland of Second Life: Virtual Social Worlds and How to Use them. Business Horizons 52, 563–572 (2009)
14. Messinger, P.R., Stroulia, E., Lyons, K., Bone, M., Niu, R.H.: Virtual worlds - past, present, and future: New Directions in Social Computing. Decision Support Systems 47, 204–228 (2009)
15. Shaf, F.M., Mueller, D., Bruns, F.W., Pereira, C.E., Erbe, H.-H.: Collaborative learning and engineering workspaces. Annual Reviews in Control 33, 246–252 (2009)
16. Young, J.R.: After Frustrations in Second Life, Colleges Look to New Virtual Worlds. In: The Chronicle of Higher Education. The Chronicle of Higher Education, Washington (2010)

17. Dana Barry, M., Kanematsu, H., Fukumura, Y., Ogawa, N., Okuda, A., Taguchi, R., Nagai, H.: Problem Based Learning Experiences in Metaverse and the Differences between Students in the US and Japan. In: International Session Proceedings, 2009 JSEE Annual Conference - International Cooperation in Engineering Education, pp. 72–75 (2009)
18. Barry, D.M., Kanematsu, H., Fukumura, Y., Ogawa, N., Okuda, A., Taguchi, R., Nagai, H.: International Comparison for Problem Based Learning in Metaverse. In: The ICEE and ICEER 2009 Korea, International Conference on Engineering Education and Research, vol. 1, pp. 60–66 (2009)
19. Kanematsu, H., Fukumura, Y., Ogawa, N., Okuda, A., Taguchi, R., Nagai, H.: Practice and Evaluation of Problem Based Learning in Metaverse. In: ED-MEDIA 2009, World Conference on Educational Multimedia, Hypermedia & Telecommunications, pp. 2862–2870 (2009)
20. Kanematsu, H., Fukumura, Y., Ogawa, N., Okuda, A., Taguchi, R., Nagai, H.: The Study on PBL Models for e-Learning – As a Series of eHelp Project. In: Proceedings of JSEE Annual Meeting, pp. 584–585 (2009)
21. Kanematsu, H., Fukumura, Y., Ogawa, N., Okuda, A., Taguchi, R., Nagai, H., Barry, D.M.: Problem Based Learning in Metaverse As a Digitized Synchronous Type Learning. In: Proceedings of the ICEE and ICEER, International Conference on Engineering Education and Research, Korea, vol. 1, pp. 330–335 (2009)
22. Taguchi, R., Fukumura, Y., Kanematsu, H., Nagai, H.: Construction of Problem Based Learing Environment in Metaverse. In: Proceedings of the 34th Annual Conference of Japanese Society for Information and Systems in Education, pp. 476–477 (2009)
23. Yoshino, T., Ikenobu, K.: Availability of Multilingual Chat Communication in 3D Online Virtual Space. In: Proceedings of The First International Conference on Culture and Computing - Culture and Computing, pp. 77–89 (2010)
24. Language grid portal site, `http://langrid.nict.go.jp/en/index.html`

Improvement of an Annotation Sharing System on Web-Based Materials to Activate Discussions

Hisayoshi Kunimune[1], Yuuji Gonda[2], Yuuki Tominaga[2], and Masaaki Niimura[2]

[1] Faculty of Engineering, Shinshu University
[2] Graduate School of Science and Technology, Shinshu University
Wakasato 4–17–1, Nagano City, Nagano 380–8553 Japan
kunimune@cs.shinshu-u.ac.jp

Abstract. This paper describes improvements of an annotations sharing system on web-based materials named *"Writable Web"*. According to experimental results, the users of the system insufficiently made asynchronous discussion, and annotated passages in a material were not properly organized because large part of the material are colored when there are many annotations in the material. We implement social tagging function and emphasizing passages function to solve these problems. We also conduct an experiment to evaluate the usefulness of these functions.

Keywords: annotation sharing, web-based material, asynchronous discussion, social tagging, collective intelligence.

1 Introduction

We have developed an annotation sharing system for web-based learning materials named *"Writable Web"* [4,5,6]. The system offers its users to write annotations directly on web pages and to share them among users. Moreover, the users can discuss by writing comments to shared annotations.

We think that to write annotations organizes the student's thinking and clarifies the important words or sentences in learning materials. The system also show others' shared annotations to a student, and the student can also know important parts from the point of view of others. These activities are important to acquire knowledge.

Furthermore, asynchronous discussions foster learners' knowledge construction [9]. Students deeply understand and construct their knowledge through discussions.

Thus, the targets of the system are 1) to clarify the students' thinking and their important part, 2) to show others' important part, and 3) to deepen their understanding through asynchronous discussions.

We conducted an experiment to confirm the usefulness and the learning effect of the system in 2007 [5]. According to the experimental result, we found two problems that are 1) the activation level of discussions among users was

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 210–219, 2010.

insufficient, and 2) user interfaces to organize others' annotations does not show proper information to the users.

We consider that to increase opportunities to find other users' opinions or ideas is the first step to activate discussions. The system manages the thread of an annotation and comments for the annotation; however, the system does not manage the relationship between an annotation and others. It is difficult to the users to find others' annotations with the system.

We also assume that the passages of a material where are annotated by many users are important parts of the material. The system shows a user the others' annotations by overlaying all annotations in the same color on a material; therefore, there are many annotations on the material, and many parts of the material are colored. It is difficult to find important parts of materials with others' annotations.

This study implements some functions to solve these problems and conducts an experiment to evaluate the functionality.

2 Related Works

Several systems that enable users to insert annotations in web-based materials and share these annotations have been proposed [1,2,3,7,8].

The systems proposed in [1], [2], [3], and [8] need specified web server or client to use these systems. All of these systems do not organize shared annotations and do not make relationship among annotations. *"Writable Web"* offers a feature to organize others' annotations and makes relationship among annotations to show users others' annotations.

Moreover, these annotations are shifted when the materials are updated. Because, these systems locate annotations on fixed positions when the annotations are written on original materials. We have proposed a method for realigning annotations in updated learning materials by using natural language processing techniques and conduct an experiment to evaluate the accuracy of proposed method [6]. Using this method, this system can determine the correct position of annotations in updated learning materials and move annotations to correct positions.

3 Overview of Writable Web

"Writable Web" has the following features:

- Works on common web browsers and does not require the installation of any special software to run.
- Offers writing marks, memos, and questions on texts (Fig. 1) and freehand drawings and memos on figures.
- Annotations are in proper positions after materials are updated.
- Supports online asynchronous discussion with shared annotations.

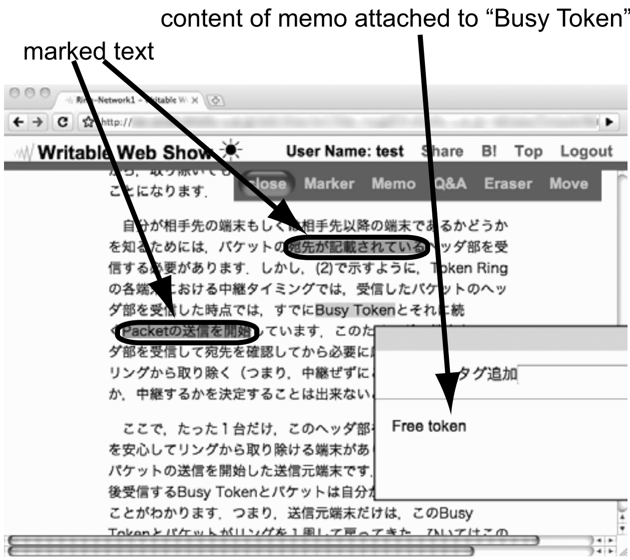


Fig. 1. Marks and memo on text

“Writable Web” works as a server-side web application between the web browser of a user and web servers that provide web-based learning materials. The system works on commonly used web browsers, and the annotations written by each user are stored in the same database on the server, allowing users to share their annotations.

When a user writes an annotation on a page, and *“Writable Web”* stores information, which specifies the position, the type, the owner, and the content of the annotation, in a database. When users browse that page, the system inserts HTML tags to draw annotations and a client-side program into the original material only on memory, and does not make any changes to the original material itself. The client-side program is written in Javascript and works on the user's web browser. Moreover, the system stores the HTML file of the original material as a cache, and uses the cache, instead of the original material, when a user browses the same material, and will use the cache to detect for any updates.

4 Implementation

We have conducted a trial experiment of *“Writable Web”* with 88 undergraduate students in 2007 [5]. According to the experimental result, we found that the students mainly intend to use the system to organize their thoughts, not to communicate or discuss with others. Online asynchronous discussions foster students' knowledge construction. Thus, we attempt to implement functions to activate discussions to the system.

The users of the system can discuss with others by attaching comments to others' memos and comments. The system provided to the students in the experiment shows the list of all memos and comments on a page. The users can find a memo or a comment from the list, and a memo and attached comments to the memo are organized in a thread structure; thus, the users can easily follow a thread of a discussion. However, a memo and attached comments have no relationship with memos and comments in the other threads.

We implement some function to make relationship among memos. This function may increase opportunities to browse others' opinions in memos and comments, and this is the first step to activate discussions.

4.1 Social Tagging

Social tagging is a grouping method with tags, which are attached by users, and vocabularies used to tags are freely chosen by users. We implement functions to attach social tags to memos and to search memos with tags. Memos, which are attached the same tag, are in a group, and multiple tags can be attached to a memo. Users can search other memos with the same tag, and they have opportunities to find others' opinions and discussions.

Tags are spelled in several different ways because users freely choose the vocabulary in tags. Users cannot find memos with other tag although the tag expresses the same thing. Thus, we also implement a function to recommend appropriate tags to users. This function extracts and recommends to users important words from the paragraph, which is attached an annotation, and the content of web page in which a user is browsing. This function uses "Key phrase extraction API" provided by Yahoo! Japan [10].

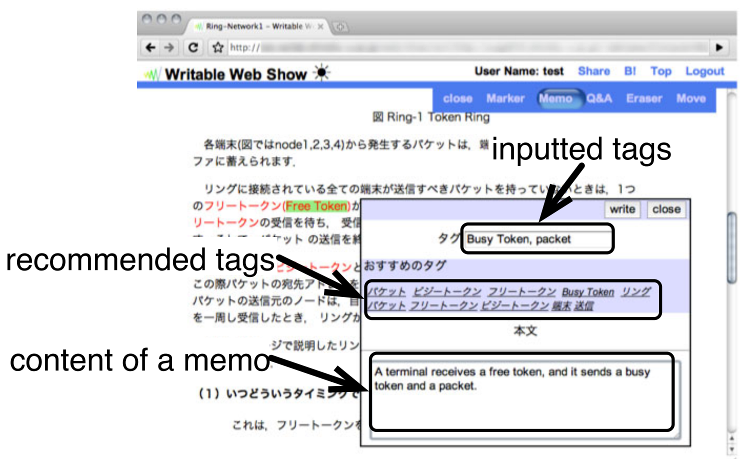


Fig. 2. Dialogue box to write a memo

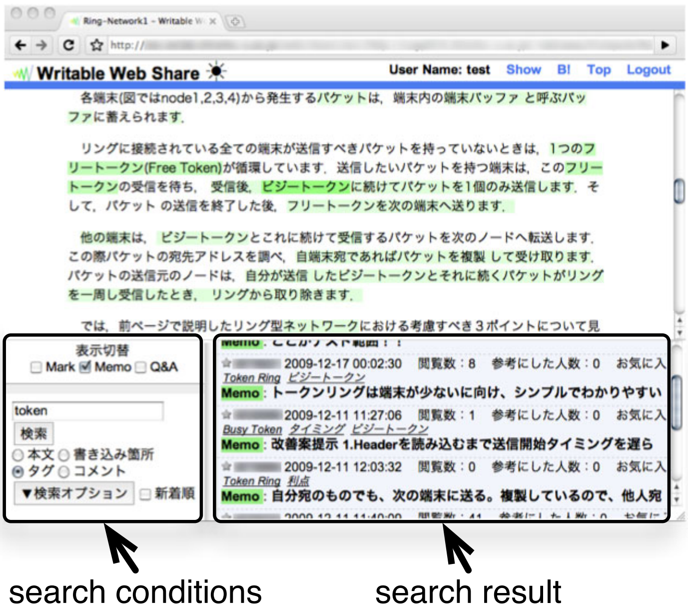


Fig. 3. Searching function

When a user writes a memo, *"Writable Web"* pops up a dialogue box as shown in Fig. 2. This dialogue box shows the user recommended tags and offers text forms to enter tags and content of the memo. The user clicks a recommended tag, and the tag is entered to the form to enter tags, and the user also can freely enter tags. The system also provides users a function to search annotations by attached tags, content, annotated string, or attached comment (Fig. 3).

4.2 Emphasizing Annotated Passages

In order to organize others' annotations, we implement a function to change the color density of annotations according to the number of annotations. There are five levels of density, and three colors depending on the type of annotations composed by marks, memos, and questions. There are two or more annotations in different types on the same character; then, the system determines the color according to the priority of these types of annotations. The system gives first priority to questions and second priority to memos; because the content of a question and a memo can be a wedge to discuss.

Fig. 4 shows an example of emphasized annotations. Users can easily recognize parts where many users annotated.

Fig. 4. Emphasized annotations

5 Evaluation

We conduct an experiment with 97 undergraduate students and a lecturer to evaluate the features mentioned above from December 11th, 2009 to January 6th, 2010. The students participated in a class "Computer network" that was held as blended learning, and they used *"Writable Web"* to annotate on five pages of web-based material in four times face-to-face lectures and to have online asynchronous discussions on the materials between the lectures. The materials used in these lectures and the configuration of these lectures are the same as them in the experiment conducted in 2007 [5]. They learn three types of ring network and write an assignment paper including two types of questions. The first type is to summarize the materials, and another one is to confirm a student's understanding of the mechanism of these networks.

After the experiment, we carried out a questionnaire survey for the students. The questions ask about the helpfulness and intended purpose of the features provided by the system, and 43 students (44.3%) answered these questions.

5.1 Overall of the System

In this experiment, 57 students wrote 368 markers (avg. 6.5 markers for each student), 97 students wrote 388 memos (avg. 4.0 memos), and 73 students wrote 75 comments (avg. 1.0 comments) to 34 memos (avg. 2.2 comments for each memo). These results indicate that discussions did not activate. We guess, one of the reasons of this result is that the system was down in some early lectures,

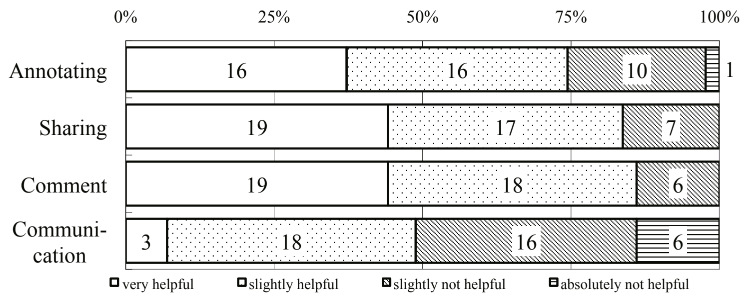
Table 1. Helpfulness of the features of the system

	2007 (N=69)	2009 (N=43)	p-value of U test
Annotating	2.12 (0.96)	1.91 (0.84)	0.167
Sharing	1.74 (0.78)	1.72 (0.73)	0.504
Comment	1.96 (0.84)	1.70 (0.71)	0.0637
Communication	2.78 (0.89)	2.58 (0.82)	0.0900

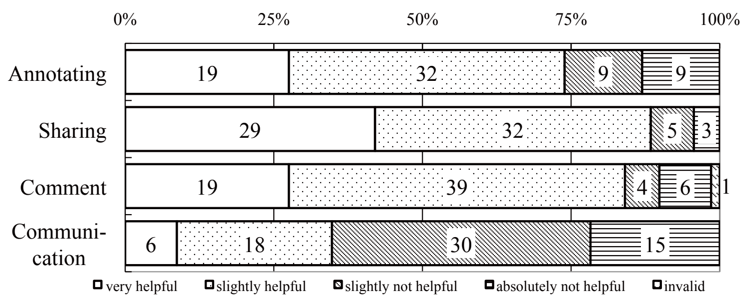
and the students might have a feeling of dislike to the system. Moreover, we will analyze the content of memos and comments to confirm the quality of the discussion. We also analyze the score of papers, which are submitted after this experiment, to confirm their achievement.

Fig.5(a) shows the results of the survey about the helpfulness of the annotating features for learning. The first question asks about writing annotations directly on web-based materials. The second question asks about sharing annotations among users, and the third question asks about writing comments to others' memos. The last question asks about the helpfulness of sharing annotations and writing comments to communicate with other users. We also asked the students of the class on 2007 the same question. Fig.5(b) shows the result of the question on 2007.

Table 1 shows the mean scores and the standard deviations (in parentheses) of each answers shown in Fig 5. In this table, we grade the answers from "very



(a) Result in 2009



(b) Result in 2007

Fig. 5. Helpfulness of the features related to annotation

helpful" to "absolutely not helpful" on a scale of one to four; thus, low score means highly evaluated by students.

To determine where there are differences in the representative values in the present and the former experiments, the Mann-Whitney U test is performed with the null hypothesis (H_0), which indicates that the "representative values are the same", and the alternative hypothesis (H_1), which implies that the "representative value in 2009 is less than that in 2007." There are no significant differences in all tests according to the result.

These results indicate that the features of writing and sharing annotations and writing comments are highly valued by the students in both experiments. However, the improved features related to activate communications are lowly evaluated.

5.2 Social Tagging

In this experiment, 52 users attached 194 tags to 130 memos (avg. 3.7 tags for each user and 1.5 tags for each comment), and the average number of browsed time was 5.3 to tagged memos and 4.9 to memos without tags. We found that the students insufficiently used the social tagging function according to these results.

Consequently, we carried out an additional questionnaire survey for students. The question ask students, who did not attach tags, about the reason why they did not attach tags to memos, and 15 students answered this question. The main answers are as follows:

(1) Tags are not important information.
(2) Tags are unnecessary for few comments.
(3) To attach tags is bothersome.
(4) Too many tags interfere to organize comments.
(5) Searching memos with other information is more efficient than searching with tags.
(6) If the user interface to search related comments improves, I would use.

According to these answers, we found that the current system and its features related to social tagging are not acceptable to the some students. We think that there are three reasons why they didn't use tags based on their answers. Firstly, there are some students who think tags are just useless things according to the comments (1), (3). Secondly, tags do not have significance with few tagged contents according to the comment (2). Thirdly, the user interface of the system is not useful for the students according to the comments (4), (5), and (6).

We should consider whether the social tags are appropriate to solve our problem according to the first reason. However, we also should make a plan to increase comments by students according to the second reason. For example, we have a plan to provide the system for more classes, which have relationships with each other classes with more students. Then, tagging will become useful feature for students to organize information. We should also improve the user interface of the system according to the third reason.

5.3 Emphasizing Annotated Passages

In order to examine the difference of importance of passages between a lecturer and learners, the lecturer annotated to important passages before the experiment. We compared the lecturer's important passages with passages that are annotated by the learners after the experiment. The learners annotated to 89 percent of annotated characters by the lecturer. They also annotated to 57 percent of not annotated characters; however, over 90 percent of them are emphasized in level one or two (lower levels).

Moreover, we asked the students about the meanings of emphasized passages for the students in the questionnaire survey. The students made multiple selections from "informative sentences", "important word", "passage that is difficult to understand in lectures", "not informative sentences", and "others." Fig.6 shows the result of the question.

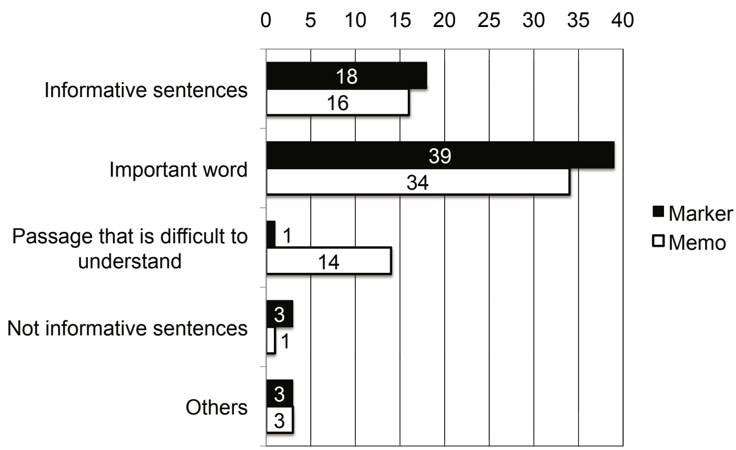


Fig. 6. Meaning of emphasized annotations

According to these results, emphasizing annotated passages are useful to organize annotations for the students; however, we should improve the determining method of the color density to emphasize more important and meaningful parts of materials.

6 Conclusion

We improved *"Writable Web"* to make opportunities to find others' memos as the first step to activate asynchronous discussions and to organize annotations by users helping to find important information.

From the experimental results, we found that the students used the social tagging functions insufficiently. We should reconsider the design of the environment, for which the system is properly utilized, and should improve the user interface of the system. We also should analyze the contents of discussions to confirm the quality of the discussions. We found that emphasizing passages functions are useful for students to organize important or meaningful parts in materials. However, some parts, which are not so important, are also emphasized; thus, we should improve the method to determine the emphasizing levels.

Acknowledgement

This research was partially supported by the Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Young Scientists (B), 21700808, 2009–2011.

References

1. Cadis, J., Gupta, A., Gruding, J.: Using Web Annotations for Asynchronous Collaboration Around Documents. In: Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work, pp. 309–318 (2000)
2. Denoue, L., Vignollet, L.: An Annotation Tool for Web Browsers and Its Applications to Information Retrieval. In: Proceedings of RIAO 2000, pp. 180–195 (2000)
3. Koivunen, M.: Annotea and Semantic Web Supported Collaboration. In: Proceeding of UserSWeb, 1st Workshop on End User Aspects of the Semantic Web, vol. 137(1), pp. 5–17 (2005)
4. Kunimune, H., Yokoyama, K., Niimura, M., Fuwa, Y.: An Annotation Sharing System for e-Learning Materials. In: Proceedings of ED-MEDIA 2007, pp. 3366–3371. Association for the Advancement of Computing in Education, Norfolk (2007)
5. Kunimune, H., Yokoyama, K., Takizawa, T., Fuwa, Y.: Development and Evaluation of a Web-Based Asynchronous Discussion System on e-Learning Materials. International Journal of Advanced Intelligence Paradigms 1(2), 163–177 (2008)
6. Kunimune, H., Yokoyama, K., Takizawa, T., Fuwa, Y.: A Following Method of Annotations on Updated Contents and Its Evaluation. In: Velásquez, J.D., Ríos, S.A., Howlett, R.J., Jain, L.C. (eds.) Knowledge-Based and Intelligent Information and Engineering Systems. LNCS, vol. 5712, pp. 555–563. Springer, Heidelberg (2009)
7. Nokelainen, P., Miettinen, M., Kurhila, J., Floreen, P., Tirri, H.: A Shared Document-Based Annotation Tool to Support Learner-Centred Collaborative Learning. British Journal of Educational Technology 36(5), 757–770 (2005)
8. Rau, P.L.P., Chen, S.H., Chin, Y.T.: Developping Web Annotation Tools for Learners and Instructors. Interacting with Computers 16(2), 163–181 (2004)
9. Schellens, T., Valcke, M.: Fostering Knowledge Construction in University Students Through Asynchronous Discussion Groups. Computer & Education 46(1), 349–370 (2006)
10. Yahoo! Japan: Key phrase extraction API, `http://developer.yahoo.co.jp/webapi/jlp/keyphrase/v1/extract.html` (accessed on March 20, 2010)

Where to Crawl Next for Focused Crawlers

Yuki Uemura, Tsuyoshi Itokawa,
Teruaki Kitasuka, and Masayoshi Aritsugi

Computer Science and Electrical Engineering
Graduate School of Science and Technology
Kumamoto University, Kumamoto 860-8555, Japan
{uemura@dbms.,itokawa@,kitasuka@,aritsugi@}cs.kumamoto-u.ac.jp

Abstract. Since WWW provides a large amount of data, it is useful for innovative and creative activities of human beings to retrieve interesting and useful information effectively and efficiently from WWW. In this paper, we attempt to propose a focused crawler for individual activities. We develop an algorithm for deciding where to crawl next for focused crawlers, by integrating the concept of PageRank into the decision. We empirically evaluate our proposal in terms of precision and target recall. Some results show that our system can give good target recall performance regardless of topics on which the crawler system focuses.

1 Introduction

The amount of available webpages in WWW has been increasing these days. Due to the advancement of technologies concerning Blog and Wiki, many webpages are generated easily and dynamically, resulting in making the amount larger and larger. It is thus required for innovative and creative activities of human beings to retrieve necessary information from WWW effectively and efficiently.

Instead of developing a general purpose WWW retrieval system such as Google and Yahoo!, Chakrabarti et al. [3] proposed a focused crawler. A focused crawler attempts to collect webpages on a set of topics in which users are interested. For it, a focused crawler needs to calculate the border between regions where to and not to crawl according to the topics. As a result, the amount of hardware and network resources for implementing a focused crawler can be smaller than that for a general purpose crawler.

In this paper, we consider running a focused crawler on our own personal computer, thereby preserving our privacy and keeping collected webpages up-to-date more easily, for individual activities. We develop an algorithm for deciding where to crawl next for focused crawlers. One way to decide where to crawl next is to exploit text data of already crawled webpages in calculating similarities between unvisited webpages linked by the crawled webpages and specific topics. However, it is naturally hard to calculate the similarity precisely if a crawled webpage has few text data. In other words, we need to find any tip other than anchor texts to be used in calculating the similarity. Since webpages generated with Blog and Wiki tend to have few text data, this can be a serious problem

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 220–229, 2010.

for focused crawlers. We attempt to integrate the concept of PageRank [2] into deciding where to crawl next. Concretely, our algorithm is based on personalized PageRank [9] and a lower bound of PageRank [5]. In this paper, we also evaluate our proposal in terms of precision and target recall [12,15,13]. Some experimental results show that our focused crawler can give good target recall performance regardless of topics on which the crawler system focuses.

The remainder of this paper is organized as follows. Section 2 mentions related work. Section 3 proposes a crawling algorithm based on the concept of PageRank. Section 4 reports some experimental results to evaluate our proposal, and Section 5 concludes this paper.

2 Related Work

A focused crawler was firstly proposed in [3], and many studies of focused crawlers have followed [6,7]. Chakrabarti et al. [3] attempted to identify hubs for specific topics for focused crawling. In [3], topics are specified by using exemplary documents. However, we assume in this paper that topics are specified by a user as keywords for simplicity, and focus on the strategy of deciding where to crawl next. Diligenti et al. [6] developed context focused crawler, in which context graphs were generated as compact context representations for modeling context hierarchies. Ester et al. [7] introduced a unique focused crawler that attempted to select websites instead of webpages. While the existing studies tend to exploit only contents, or text data, of webpages, we exploit not only them but also ranks of webpages calculated with the concept of PageRank. This idea was inspired by [5].

There have been many studies on web crawling ordering [4,10,1,5]. Cho and Schonfeld [5] discussed crawler coverage guarantee and crawler efficiency. They defined RankMass as the sum of PageRank values of crawled webpages, and developed a set of algorithms using RankMass for providing a theoretical coverage guarantee. In [5], they also defined a lower bound of PageRank, and we exploit it in this paper. Main different points of our work from them are to focus on focused crawlers instead of general purpose crawlers and to use precision and target recall [12,15,13] in evaluation.

3 A Prioritization Algorithm for Focused Crawlers

We attempt to integrate the concept of PageRank [2] into deciding where to crawl next for focused crawlers. This can be developed with personalized PageRank [9] and a lower bound of PageRank [5] as follows.

3.1 PageRank for Focused Crawlers

PageRank is based on the random surfer model. The importance of each webpage is calculated as a probability that the webpage is accessed in the random surfer

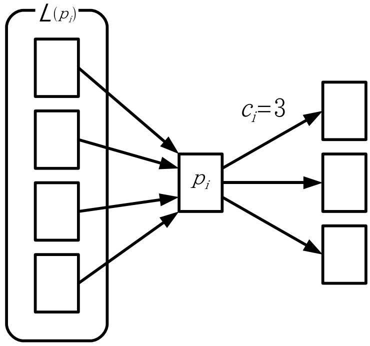


Fig. 1. Example of links of page p_i

model. In other words, the higher the PageRank value of a webpage is, the more important the webpage is supposed to be in WWW.

In [9], the original PageRank is refined as personalized PageRank. Let $L(p_i)$ be the set of webpages that have at least a link to webpage p_i, and c_j be the number of links that webpage p_j has. An example of links concerning webpage p_i is shown in Fig. 1 where c_i is three. Then, personalized PageRank of webpage p_i, which is expressed as r_i, is defined as follows.

$$r_i = d \left[\sum_{p_j \in L(p_i)} \frac{r_j}{c_j} \right] + (1-d)t_i \tag{1}$$

In (1), d is a constant value called a damping factor and is often set to 0.85 [8], and t_i is trust score of webpage p_i. If a webpage is supposed to be trusted, then its trust score is a non-zero value.

In this study, we assume a variation of the random surfer model, in which a Web surfer who attempts to access only webpages that are related to a specified topic, and propose another PageRank based on the model for focused crawlers. The PageRank for focused crawlers is going to be used for deciding where to crawl next in crawling. With respect to the model, the probability that a Web surfer, who is attempting to access webpages about a specified topic T, accesses webpage p_i is calculated as follows:

$$r_i = d \left[\sum_{p_j \in L(p_i)} \frac{sim(a_{ij}, T) \times r_j}{\sum_{a_{kj} \in A_j} sim(a_{kj}, T)} \right] + (1-d)t_i, \tag{2}$$

where A_j is the set of anchor texts that webpage p_j has, a_{ij} is an anchor text for the link to webpage p_i in webpage p_j, and $sim(a_{kj}, T)$ is the cosine similarity between anchor text a_{kj} and topic T.

3.2 Our Prioritization Algorithm

We prioritize unvisited webpages for deciding which of them we should crawl next in our focused crawler. Our prioritization algorithm is based on a lower bound of PageRank proposed in [5].

A candidate way of deciding a webpage to crawl next is to calculate PageRanks of all webpages and to select a webpage with the highest PageRank. Note, however, that it is naturally impossible to calculate precise PageRanks of webpages that have never been accessed yet. Instead of calculating precise each PageRank, Cho and Schonfeld [5] proposed calculating a lower bound of it based on visited webpages.

In this study, we attempt to integrate the idea of a lower bound of PageRank into focused crawlers. Assume that there is a path from webpage p_j to p_i in WWW. Let t_j be the trust score of webpage p_j. According to (2), we can say that webpage p_j has $(1-d)t_j$ as a lower bound of its PageRank, regardless of link structures around the webpages. Let w_{ji} and W_{ji} be a path and the set of all paths between the two webpages, respectively, and $|w_{ji}|$ be the number of clicks to get webpage p_i along path w_{ji}. Then, the probability of being on webpage p_i from webpage p_j along path w_{ji} without being interrupted can be expressed as $d^{|w_{ji}|}$. Let p_k be a webpage on path w_{ji}, and S_k and s_k be the sum of similarities between a certain topic and all anchor texts webpage p_k holds and the similarity between the topic and the anchor text on webpage p_k that the surfer clicks, respectively. Then, the probability that the surfer gets webpage p_i from webpage p_j along path w_{ji}, expressed as $PP(w_{ji})$, is calculated as follows.

$$PP(w_{ji}) = \left(\prod_{p_k \in w_{ji}} \frac{s_k}{S_k} \right) (1-d) t_j d^{|w_{ji}|} \tag{3}$$

Let D_c be the set of webpages crawled already. Then, we can calculate the probability that the surfer accesses webpage p_i, or a lower bound of its PageRank, as follows.

$$r_i \geq \sum_{p_j \in Dc} \sum_{w_{ji} \in W_{ji}} PP(w_{ji}) \tag{4}$$

In (4), a lower bound of PageRank of a webpage is propagated to those linked by the webpage. As a result, we can calculate a lower bound of PageRank of each webpage linked from a webpage by means of (4). This equation is calculated during crawling, and we can decide where to crawl next by selecting the webpage with the highest value of this lower bound.

Figure 2 shows a flow of processing our prioritization algorithm on an example where there are webpages p_1, p_2, and p_3. At the first step, p_1 has been crawled and has 0.1 as its lower bound of PageRank. On the other hand, p_2 and p_3 have not been crawled yet, and thus they have 0 as their lower bounds of PageRank. Then, p_1's lower bound of PageRank is propagated to the uncrawled webpages p_2 and p_3 according to their similarities between their anchor texts and the topic. In the example, 0.028 and 0.057 are propagated to p_2 and p_3, respectively, as

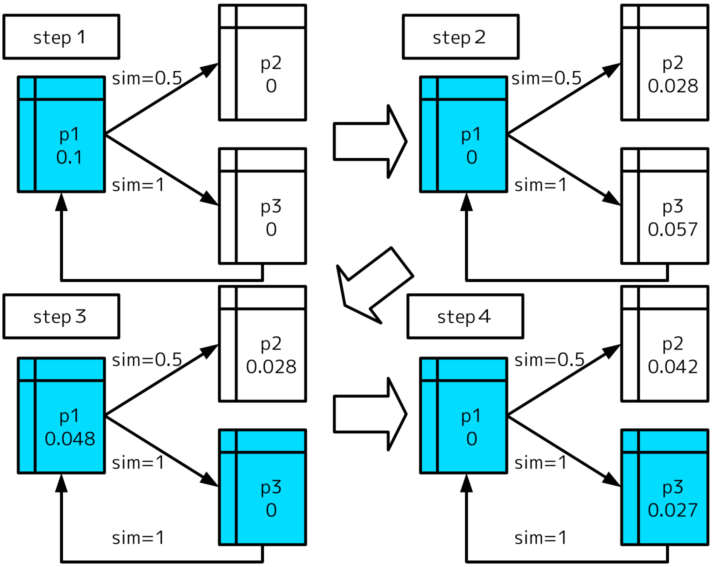


Fig. 2. Example of results of our prioritization algorithm

shown in the second step. After the propagation, p_1's lower bound is set to 0. Since p_3 has the highest value at that moment, the webpage is selected to be crawled. At step 3, the webpage with the highest value is p_1, and the crawler processes our prioritization algorithm without crawling the webpage, since that has been crawled. Finally, p_2 has the highest value at step 4 and is selected to be crawled by our system.

4 Experiments

We empirically evaluated our proposal. In the evaluation we used precision and target recall [12,15,13] as metrics.

4.1 Environment

The data we used in the experiments come from the Open Directory Project (ODP) [11], which is a human-edited webpage directory. Webpages are categorized into topics and the topics construct topic hierarchies in ODP. In the experiments, we assumed that the topics and topic hierarchies of ODP were correct.

The experiments were conducted in Japanese. To evaluate the effectiveness of our proposal, we randomly selected three main topics, namely, camera, cancer, and recipe. As their child topics, camera had classic, Leica, and digital camera, cancer had lung cancer and breast cancer, and recipe had menu and easy cooking, respectively. We used 20 webpages in each topic as its seeds, and the rest of the webpages were used as its targets. The 20 webpages were randomly selected from each directory, the proportion of them were set to the proportion of webpages in each topic hierarchy. For example, assume that topic B is a child of topic A and the numbers of webpages categorized in A and B are 20 and 30, respectively. In this case the numbers of seeds from topics A and B are 8 and 12, respectively. We extracted feature words from the 20 webpages by using tf-idf method, and each main topic was modeled with the feature vector consisting of the words. We calculated the cosine similarity between the feature vector and anchor texts in crawling.

In the experiments, we implemented three crawlers, namely, our proposal, a focused crawler using anchor texts only, and a crawler based on breadth-first crawling. We included the breadth-first crawling in the experiments for the comparison because of the results of [10]. In the anchor texts only strategy, as in [12,14], the score of each linked webpage was estimated as follows,

$$score = \beta \times page\_score + (1 - \beta) \times context\_score, \tag{5}$$

where *page_score* stands for the cosine similarity between the feature vector constructed with words extracted from all seeds using tf-idf method and the feature vector constructed with words extracted from the crawled webpage using tf-idf method, and *context_score* stands for the cosine similarity between the feature vector constructed with words extracted from all seeds using tf-idf method and the feature vector constructed with anchor texts from the crawled webpage using tf-idf method. We set $\beta = 0.25$, which come from [12,14], in the experiments.

To evaluate the three crawlers, we decided to use precision and target recall [12,15,13] as metrics. After crawling N webpages, the precision is defined as follows,

$$precision = \frac{1}{N}\sum_{i=1}^{N} sim(T, p_i), \tag{6}$$

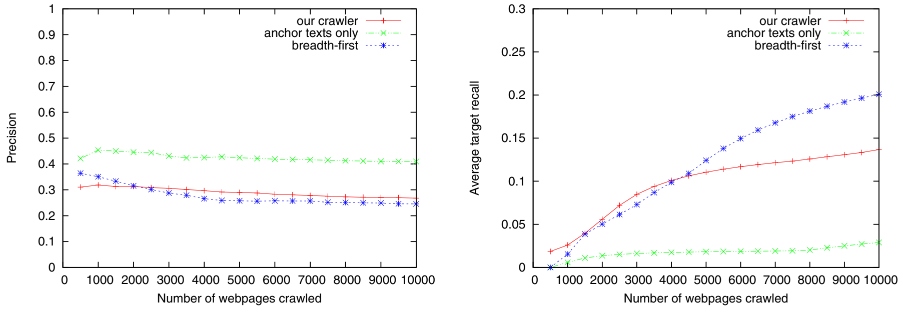
where $sim(T, p_i)$ is the cosine similarity between topic T and webpage p_i. Let T_t be the set of targets of topic t, and C_t^N be the set of crawled N webpages according to the topic. Then, the target recall is defined as follows.

$$target\ recall = \frac{|T_t \cap C_t^N|}{|T_t|} \tag{7}$$

In the following, we report not only values of target recall but also those of average target recall, which is calculated by dividing the sum of the values of target recall by the number of crawled webpages. We think this metric is significant for focused crawlers running on personal computers with poor computing resources because the results can tell us how fast a crawler can collect targets.

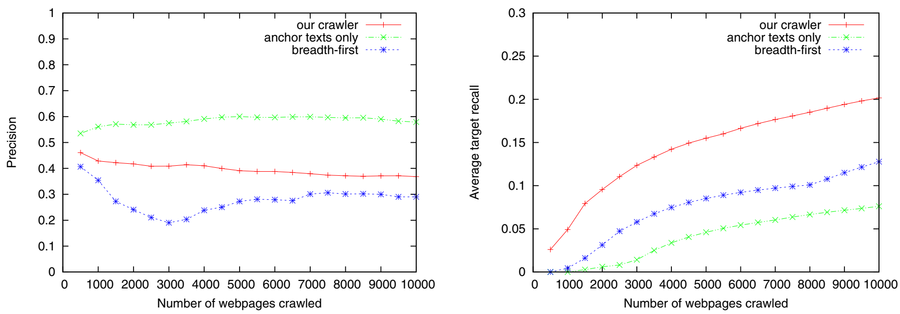
4.2 Results

We run the three crawlers with identical seeds for collecting 10,000 webpages. Figures 3, 4, and 5 show the performances of precision and average target recall



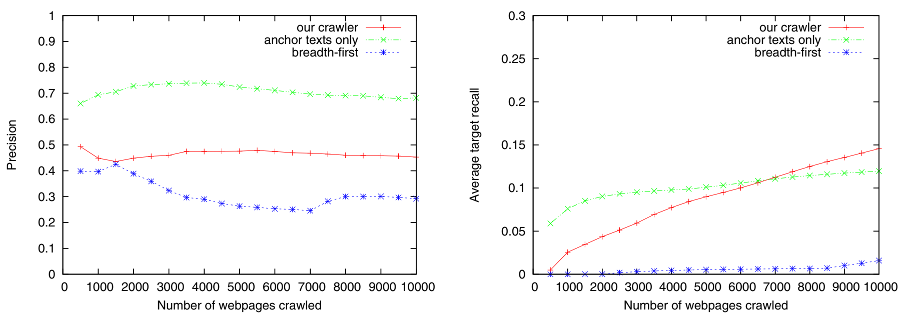
(a) Precision (b) Average target recall

Fig. 3. Results on camera



(a) Precision (b) Average target recall

Fig. 4. Results on cancer



(a) Precision (b) Average target recall

Fig. 5. Results on recipe

Table 1. Target recalls after crawling 10,000 webpages

	camera	cancer	recipe
our crawler	0.213	0.270	0.254
anchor texts only	0.0638	0.127	0.140
breadth-first	0.298	0.254	0.0789

of the three crawlers on the three main topics, respectively. Table 1 reports the values of target recall of the crawlers after crawling the 10,000 webpages.

In the experiments, the anchor texts only crawler could give the best precision performance in the three main topics, as shown in Figs. 3(a), 4(a), and 5(a). The results can indicate that to use anchor texts can work well for focused crawlers. On the other hand, our proposal could give the second precision performance among the three. Incidentally, the precisions of targets of camera, cancer, and recipe used in the experiments were 0.237, 0.303, and 0.505, respectively. The results indicate that the precision performance of our proposal was not so bad in the real world. Note that the precision was calculated with (6), and if we intend to improve the precision then the system will collect only similar webpages; this will not what we expect. We intend to rethink how to assess the precision more suitably for focused crawlers in future.

As shown in Figs. 3(b), 4(b), and 5(b), our crawler could give the best performance of average target recall on topics cancer and recipe after crawling many webpages. On the other hand, the breadth-first crawler could give the best performance on topic camera. This is mainly because webpages on camera might have more links to each other than those on the other topics. Note that the breadth-first crawler could give better average target recall performance on topic cancer and could give worse performance on topic recipe than anchor texts only crawler. We found the following characteristics of webpages in the two topics.

– Many webpages on cancer linked to each other. This is mainly because many authors of webpages on this topic might want to collect valuable information in WWW and share with many WWW users.
– There are many commercial webpages on topic recipe, and we did not find many links between them. We thus thought webpages on certain topics related to business and/or selling of goods and services might not have links with each other.

In addition, we think that the results indicate the performance of average target recall of the breadth-first crawler and anchor texts only crawler can depend on topics, because the structure of links between webpages tends to depend on topics. In contrast, our crawler could give the most stable performances of average target recall and target recall regardless of topics on which the crawler system focused, as shown in Figs. 3(b), 4(b), and 5(b) and Table 1.

5 Conclusion

In this paper, we have developed a prioritization algorithm for focused crawlers. Our algorithm exploits not only anchor texts but also the concept of PageRank in deciding where to crawl next. According to the results of our experimental evaluation, our algorithm can give good target recall performance regardless of topics on which the crawler system focuses.

We included the results of the three topics in this paper, and intend to evaluate our proposal with more topics in details in future. The evaluation results indicate that characteristics of links between webpages can depend on topics, and it would be helpful for focused crawlers to consider them in crawling algorithms.

References

1. Baeza-Yates, R., Castillo, C., Marin, M., Rodriguez, A.: Crawling a country: better strategies than breadth-first for Web page ordering. In: WWW 2005: Special Interest Tracks and Posters of the 14th International Conference on World Wide Web, pp. 864–872. ACM, New York (2005), http://doi.acm.org/10.1145/1062745.1062768
2. Brin, S., Page, L.: The anatomy of a large-scale hypertextual Web search engine. Computer Networks and ISDN Systems 30(1-7), 107–117 (1998), http://www.sciencedirect.com/science/article/B6TYT-3WRC342-2N/2/63e7d8fb6a64027a0c15e6ae3e402889
3. Chakrabarti, S., van den Berg, M., Dom, B.: Focused crawling: a new approach to topic-specific Web resource discovery. Computer Networks 31(11-16), 1623–1640 (1999), http://www.sciencedirect.com/science/article/B6VRG-405TDWC-1F/2/f049016_cf8fefd114f056306b5ae4a86
4. Cho, J., Garcia-Molina, H., Page, L.: Efficient crawling through URL ordering. Computer Networks and ISDN Systems 30(1-7), 161–172 (1998), http://www.sciencedirect.com/science/article/B6TYT-3WRC342-2G/2/122be31915c6e16c444898fb12cfdf87; Proceedings of the Seventh International World Wide Web Conference
5. Cho, J., Schonfeld, U.: RankMass crawler: A crawler with high PageRank coverage guarantee. In: VLDB 2007: Proceedings of the 33rd International Conference on Very Large Data Bases, pp. 375–386 (2007), http://www.vldb.org/conf/2007/papers/research/p375-cho.pdf
6. Diligenti, M., Coetzee, F., Lawrence, S., Giles, C.L., Gori, M.: Focused crawling using context graphs. In: VLDB 2000: Proceedings of the 26th International Conference on Very Large Data Bases, pp. 527–534. Morgan Kaufmann, San Francisco (September 2000), http://www.vldb.org/conf/2000/P527.pdf
7. Ester, M., Kriegel, H.P., Schubert, M.: Accurate and efficient crawling for relevant websites. In: VLDB 2004: Proceedings of the Thirtieth International Conference on Very Large Data Bases, pp. 396–407 (2004), http://www.vldb.org/conf/2004/RS10P3.PDF
8. Haveliwala, T.H.: Topic-sensitive PageRank: A context-sensitive ranking algorithm for Web search. IEEE Transactions on Knowledge and Data Engineering 15(4), 784–796 (2003), http://doi.ieeecomputersociety.org/10.1109/TKDE.2003.1208999

9. Jeh, G., Widom, J.: Scaling personalized web search. In: WWW 2003: Proceedings of the 12th International Conference on World Wide Web, pp. 271–279. ACM, New York (2003), `http://doi.acm.org/10.1145/775152.775191`
10. Najork, M., Wiener, J.L.: Breadth-first crawling yields high-quality pages. In: WWW 2001: Proceedings of the 10th International Conference on World Wide Web, pp. 114–118. ACM, New York (2001), `http://doi.acm.org/10.1145/371920.371965`
11. Open Directory Project: `http://www.dmoz.org/`
12. Pant, G., Menczer, F.: Topical crawling for business intelligence. In: Koch, T., Sølvberg, I.T. (eds.) ECDL 2003. LNCS, vol. 2769, pp. 233–244. Springer, Heidelberg (2003), `http://www.springerlink.com/content/p0n6lh04f4j7y26u`
13. Pant, G., Srinivasan, P.: Learning to crawl: Comparing classification schemes. ACM Trans. Inf. Syst. 23(4), 430–462 (2005), `http://doi.acm.org/10.1145/1095872.1095875`
14. Pant, G., Srinivasan, P.: Link contexts in classifier-guided topical crawlers. IEEE Transactions on Knowledge and Data Engineering 18(1), 107–122 (2006), `http://doi.ieeecomputersociety.org/10.1109/TKDE.2006.12`
15. Srinivasan, P., Menczer, F., Pant, G.: A general evaluation framework for topical crawlers. Inf. Retr. 8(3), 417–447 (2005), `http://dx.doi.org/10.1007/s10791-005-6993-5`

Extraction of Co-existent Sentences for Explaining Figures toward Effective Support for Scientific Papers Reading

Ryo Takeshima and Toyohide Watanabe

Department of Systems and Social Informatics,
Graduate School of Information Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464-8603, Japan
{takeshima,watanabe}@watanabe.ss.is.nagoya-u.ac.jp

Abstract. It is important for researchers/investigators to read and understand scientific papers effectually and effectively. However, it takes much time and many efforts to read and understand many papers related directly to their researches, even if they could refer to necessary papers timely. In this paper, we address a function for supporting the scientific paper understanding process successfully. We focus on figures which can usually explain the important topics along a series of successive paragraphs, and develop an intellectual tool which collects the mutually related sentences, attended interdependently to the focused figure, and supports a paper understanding ability through the focused figure. In this paper, we introduce the propagation mechanism of important words over the corresponding sentences. This propagation mechanism can select candidate sentences appropriate to explain the focused figure.

Keywords: Figure explanation, Weight propagation, Paper reading.

1 Introduction

Todays, we can obtain a lot of information easily and instantly through the Internet. Of course, this phenomena is similarly observed in research and development fields. Scientific/technical papers take an important role when researchers/investigators should progress their creative studies successfully: in the initial step, to grasp the investigation direction and scope in the research field; in the progress step, to understand the research objective, the approach and method, and the experimental results and discussions from interesting paper with a view to making up its own research viewpoints; and in the final step, to classify many related papers analytically into citation-oriented references in order to prepare the research paper.

Extraction of figure-specific explanation sentences is a kind of summary generation. Many methods about automatic summary generation are based on extraction means [1,2]. The method calculates the importance of sentence from some information, and generates summaries to use the information smartly. There

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 230–239, 2010.

are some studies that extract important sentences by using machine learning. Kupiec, et al. assumed the extraction of important sentences as a statistical classification problem and created a function that calculates the probability of whether a sentence is important by use of training data made analytically on the basis of Bayesian classifier [3]. In addition, Lin created summaries using decision tree learning [4]. These learning means need many training data: Hirao, et al. extracted important sentences using Support Vector Machine that has good generalization ability [5]. These methods can create summaries or extract important sentences statically, but cannot extract dynamically from user input. Figure-specific explanation sentence is dependent on figures, and the traditional methods cannot be well useful.

How to read and understand research-related papers is too strongly dependent on the information management in the research phase: fast reading to extract the important feature certainly; careful reading to know the description contents in detail; and pin-point reading to reconfirm the already-read content. Our objective in this paper is to develop a smart paper-reading function. The research idea is to support figure-oriented explanation means: namely, figures used in scientific/technical papers almost partly explain important concepts, procedural methods, experimental environments/results and so on illustratively, and also resource knowledge related directly to the paper description can be grasped, corresponding to the interests of researchers/developers. Traditionally, some methods such as summary composition, topic extraction and so on have been investigated with respect to the support for smart paper understanding. These research subjects presented some useful effects with help of natural language processing. Of course, topic extraction or extraction of important word is looked upon as one of basic technical methods for summary generation. These extraction methods of topics, important words or co-existed words take an important role to identify a set of sentences related to focused figures/tables. In particular, the extraction of co-existed words is useful because the figure/table-specific explanation sentences should be selectively recognized from successively located sentences. The main idea is to propagate the important words, included in directly referred sentences, to the successors and predecessors with co-existence relationship for important words, one by one.

2 Approach

Our objective is to extract well-expressed sentences, related strongly to figures/tables, with a view to grasping the contents of figures/tables effectively. The central idea is to focus on the dependency between sentence and word. The concept of this interdependency was proposed by Karov, et al. in order to get rid of ambiguity attended with word meanings [6]. This dependency between word and sentence means that the similar words are included in the co-related sentences and also that the similar sentences contain the corresponding related words. Generally, the explanation sentences related to figures or tables in academic/technical papers are located collectively after they were firstly referred, but this observation is not always true. In some cases, the detail explanation or related explanation in other places, which are different from the firstly cited place, is often observed.

In our approach, we make use of stepwise propagation means of weights, assigned to important words, for the surrounding/successive sentences: the weight attended to the important word is useful to select the meaningfully related sentences from all sentences in the paper. Moreover, our weight propagation means, based on the interdependency between sentence and word, is suitable to select the appropriate explanation sentences, which are meaningfully co-existed with sentences referred firstly to figures/tables, along the logical structure.

The word propagation process is composed of 2 steps: one is the calculation of word importance; and the other is the update of sentence weight. Until today, many investigations have been reported, which calculate the word importance and extract most important keywords on the basis of the importance. Edmundson proposed a keyword extraction method, using the access words [7]. Kimoto investigated a method to exclude noises from extracted keywords, on the basis of the meaningful relationship between keywords, derived from sentence structure, access words or thesaurus [8]. However, it is difficult to apply these traditionally proposed methods to our objective because our application fields are not fixed to special research scopes with a predefined paper forms, and the amount of data to be preset become too large if we should manage all fields. Luhn proposed another method of keyword extraction based on the frequency of word occurrences [9]. As the frequency-based extraction method is likely to pick up general words, it is necessary to exclude these general words by using *tf-idf* and so on, with a view to making the extraction ratio high. However, this method is useful to distinguish individual important words since the extraction scope is correspondingly limited to application-specific fields. In our objective, it is necessary to develop some advanced methods or approaches in addition to these traditional methods. The weight propagation cannot be completed by only one trial, but must be repeated one by one. We apply the word frequency to estimate the word importance. It is possible to make the importance degrees of general words lower even if we did not use *tf-idf* or others, because we can look upon the important words whose weights are low and whose frequencies are high, as general words. Under such a viewpoint, we address a method to estimate the word importance from the word frequency.

Propagation of Weight: We make use of weight which is assigned individually to each sentence, with a view to selecting appropriate figure/table-specific explanation sentences. It is not sufficient to extract figure/table-specific explanation sentences only by the positional relationships, derived from the corresponding paragraphs. To recover this insufficient process, we introduce the weight as a calculative value and distinguish the useful sentences on the basis of semantic relationships between sentences after having propagated the weight mutually over sentences. Our weight indicates the suitability of its sentence evaluated for figure/table-specific explanation sentence, when the weight propagation process has finished. The weight among sentences is propagated through common words

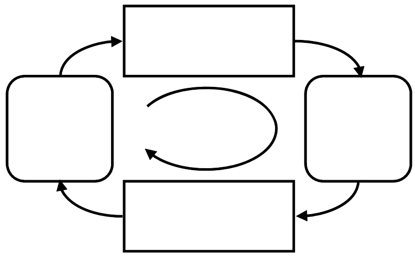


Fig. 1. Propagation Concept

for individual sentences. Figure 1 shows the principle of weight propagation briefly. The importance of word is calculated from the weights of all sentences related inclusively to the word, and also the weight of sentence is derived from the importances of all words included in the sentences. Thus, the weight of sentence is propagated to other sentences one by one, whenever the importance of word is calculated and then the weight of sentence is updated. In our idea the most important viewpoint is to focus on meaningful relationship, but not on the locational relationship. The calculation of word importance and update of sentence weight should not be reflected by the distance from word or sentence.

In our case, the weight propagation is applicable to noun, whose representations are clearly identified or not changeable in comparison with others. The word importance is calculated from all nouns in the target paper. This importance is a temporary value to be used in weight propagation and is initialized in every propagation. In our weight definition, the importance of word is set from the weight, when the word is counted many times in the sentence of large weight and also included in the sentence of small weight. We assume that our weight propagation assigns a large weight value to the appropriate sentence as the figure/table-specific explanation sentence. In this case, the initial value is set on the basis of sentences referred to focused figures or tables. Generally, in the location where figures or tables are described the figures or tables are firstly referred. Thus, the sentence directly referred to figure/table-specific explanation sentences is looked upon as the initial part of word propagation.

3 Method

The flow chart for extracting figure/table-specific explanation sentences is shown in Figure 1. First, the system calculates the initial values of weights for each sentence without being dependent on positional relationship among sentences. Next, the calculation of word importance and update of sentence weight are repeated until user-specified number. An importance of word is calculated from weights of sentences, and the weight of sentence is calculated from importance of word, so a weight is propagated from a sentence to another sentence. Weights of appropriate sentences for figure/table-specific explanation become greater by

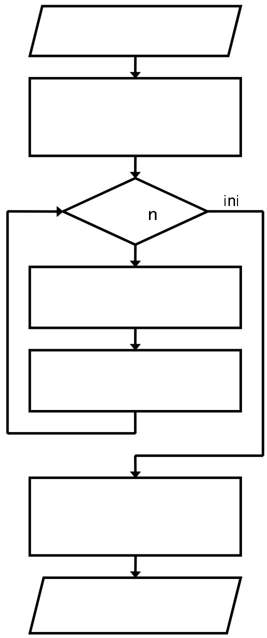


Fig. 2. Processing flow

repeating the propagation of weight. Finally, the system ranks in the order based on the weight and extracts from the upper ones.

3.1 Calculation of Initial Weight

Initial value of sentence weight is set on the basis of relative positions between sentences. This is based on the idea that the sentences describing a figure or table are located near by a figure/table reference sentence. First, the system looks for figure/table reference sentences. Next, for each sentence s_l, the initial value of weight $Weight_0(s_l)$ is calculated by using the following definition.

$$Weight_0(s_l) = \alpha \sum_{r \in R_f} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{(l-r)^2}{2}\right\} \tag{1}$$

Equation 1 includes a normal distribution formula whose average is an index of figure/table reference sentence r and whose standard deviation is 1. l represents an index of calculating sentence, and s_l is the sentence. If there are multiple figure/table reference sentences, R_f has multiple elements, and the weight is summed up. α is a normalization factor defined as follows.

$$\alpha = \frac{1}{\sum_{l \in L_s} \sum_{r \in R_f} \frac{1}{\sqrt{2\pi}} \exp -\frac{(l-r)^2}{2}} \tag{2}$$

L_s is the set of indices of all sentences. α is the inverse number of the sum of weights.

3.2 Calculation of Word Importance

An importance of a word is calculated from the weights of sentences that include the word. For each word w_l, the importance $Importance_p(w_l)$ is defined as follows.

$$Importance_p(w_l) = \frac{1}{|S_{w_l}|} \sum_{s \in S_{w_l}} Weight_{p-1}(s) \tag{3}$$

S_{w_l} is a set of sentences containing the word w_l. p represents the number of propagation. The sum of the weights of the sentences in S_{w_l} is divided by the number of the sentences in S_{w_l}. In this way, the importance of common words that appear throughout the paper is held down.

3.3 Update of Sentence Weight

A weight of sentence is updated based on the idea that semantically similar statements share many words. The weight of a figure/table-specific explanation sentence is increased by the sentences which are composed of important words and the sentences whose weights are greater. The weight of a sentence $Weight_p(s_l)$ is updated by the definition as follows.

$$Weight_p(s_l) = \beta \left\{ Weight_{p-1}(s_l) + \gamma \sum_{w \in W_{s_l}} Importance_p(w) \right\} \tag{4}$$

W_{s_l} is a set of words that compose the sentence s_l. $Weight_p(s_l)$ is the sum of the importance of words that compose the sentence S_l, and the sentence weight in the previous iteration. γ is a coefficient that adjusts a propagation speed. β is a normalization factor defined as follows.

$$\beta = \frac{1}{\sum_{l \in L_s} \left(Weight_{p-1}(s_l) + \gamma \sum_{w \in W_{s_l}} Importance_p(w) \right)} \tag{5}$$

3.4 Extraction of Figure/Table-Specific Explanation Sentences

After the propagation phase, the sentences that have higher weights are extracted as figure/table reference sentences. The procedure for extracting figure/table reference sentences is illustrated in Algorithm 1. This algorithm takes a set of all sentences in research papers/articles S as an input and returns a set of extracted sentences E as an output. First, the sentences in S are sorted in descending order by their weights. Then, the sentences are added to E from the top of S while the condition $l < l_{min}$ is true. Here, l is the total length of the sentences in E and

Algorithm 1. Extraction algorithm

$Sort(S)$;
$l \leftarrow 0$;
$i \leftarrow 0$;
$E \leftarrow \phi$;
while $l < l_{min} and i < |S|$ **do**
  $E \leftarrow E \cup S[i]$;
  $l \leftarrow l + Length(S[i])$;
  $i \leftarrow i + 1$;
end while
$Express(E)$;

l_{min} is the predefined minimum value of the total length of extracted sentences. $Length(S[i])$ is a function that returns the length of the sentence $S[i]$.

In this algorithm, the number of extracted sentences is decided according to the total length of sentences. This is because the amount of information contained in a sentence varies according to its length. Even though the number of extracted sentences is the same number, the amount of information contained in extracted sentences is different according to the total length of the sentences.

4 Prototype System

We constructed a prototype system for extracting figure/table-specific explanation sentences and supporting the paper understanding. When a user points

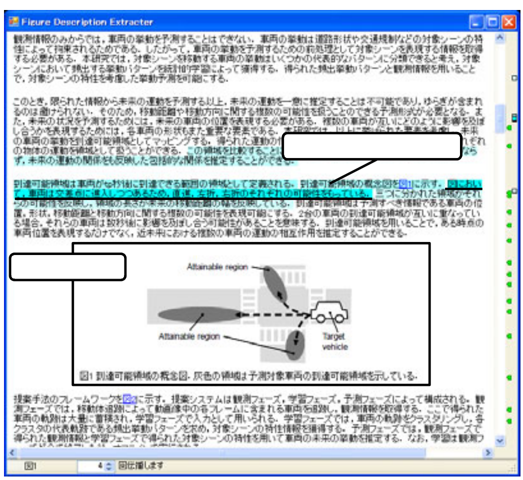


Fig. 3. Main window

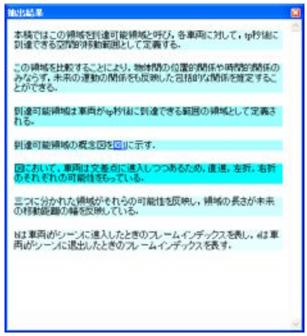


Fig. 4. Extraction window

out a figure or table with a view to understanding the figure/table interpretatively, the system first calculates the weights of sentences and then extracts figure/table-specific explanation sentences. Figures 3 and 4 shows an interface window. The system consists of two windows. Figure 3 is a main window. Papers are displayed on this window, and users can indicate their intentions through this window directly. Of course, the user can change the propagation control by means of an up-down control mechanism. Figure 4 is another operational window which displays the extracted explanation sentences. In this case, these sentences include individually "attainable region" or "moving distance".

5 Experiment

We conducted two experiments to evaluate whether figure/table-specific explanation sentences are extracted successfully so as to be consistent to user understanding objective. The evaluation criterion are *recall* and *precision*. We selected 24 figures from 4 papers. The speed of propagation γ and the number of propagation were set to 0.1 and 4, respectively. We evaluated that a proposal method can extract sentences which assist understanding.

Recall: We pointed out 3 sentences which are required to each figure manually, and regard the sentences as correct answers. We examined how many correct sentences are extracted by the system. Experimental results are shown in Table 1. The system extracted at least one correct sentence. In addition, 3 correct sentences were extracted in 8 cases.

Table 1. Experimental result on recall

The number of correctly extracted sentences	0	1	2	3
The number of cases	0	5	11	8

Precision: We investigated how many sentences related to the focused figure were extracted by our method. We regarded sentences, that include contents relevant to the figures, as correct sentences. Experimental result is shown in Table 2. Ratio of correct sentences was 76.8%. Some of the extracted sentences were not most helpful to understand figures, but many sentences included contents relevant to figures.

Table 2. Experimental result on precision

Figure index	Extracted sentences	Correct sentences	Ratio of correct sentences
1	3	2	66.7%
2	5	4	80.0%
3	6	6	100.0%
4	7	5	71.4%
5	5	3	60.0%
6	5	4	80.0%
7	6	3	50.0%
8	6	5	83.3%
9	6	5	83.3%
10	4	2	50.0%
11	7	6	85.7%
12	7	6	85.7%
13	7	6	85.7%
14	7	6	85.7%
15	7	6	85.7%
16	8	7	87.5%
17	7	4	57.1%
18	9	8	88.9%
19	8	3	37.5%
20	6	6	100.0%
21	6	4	66.7%
22	7	6	85.7%
23	6	4	66.7%
24	6	5	83.3%
Total	151	116	76.8%

6 Conclusion

In this paper, the method for extracting figure/table-specific explanation sentences was proposed. Generally speaking, figures/tables reflect importance contents of papers. Thus, it is better to understand figure/table-specific explanation sentences effectually with respect to paper reading. We proposed the weight propagation means in order to extract the related explanation sentences successfully.

In this method, the result is extracted as a set of sentences. Since it is difficult to understand the extracted sentences only by reading the set, users have to read the sentences around the set. Therefore, it is necessary to improve the method of calculating weights and showing results. Also, the definition of propagation should be reconsidered, and additional parameters for calculating the weight of sentence should be introduced.

From experimental results, we confirmed that figure/table-specific explanation sentences were extracted. Since figures or tables represent important contents in papers, understanding figures/tables lead up to understanding the whole paper. However, it is not clear how a user understand the paper, when the user read figure/table-specific explanation sentences extracted our method. In order to clarify it, we need to conduct other experimentation and to compare our method to other methods with efficiency of understanding papers.

The file format of a paper available in the Internet is often PDF. Users have to convert papers into the expressions with proprietary formats for our prototype system. Conversion takes a lot of troubles; therefore, it is necessary to improve the system so as to handle PDF files directly.

References

1. Mani, I.: Automatic Summarization. John Benjamins Pub. Co., Amsterdam (2001)
2. Radev, D.R., Hovy, E., McKeown, K.: Introduction to the Special Issue on Summarization. Computational Linguistics 28(4), 399–408 (2002)
3. Kupiec, J., Pedersen, J., Chen, F.: A Trainable Document Summarizer. In: SIGIR 1995: Proceedings of the 18th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 68–73 (1995)
4. Lin, C.Y.: Training a Selection Function for Extraction. In: CIKM 1999: Proceedings of the Eighth International Conference on Information and Knowledge Management, pp. 55–62 (1999)
5. Hirao, T., Takeuchi, K., Isozaki, H., Sasaki, Y., Maeda, E.: SVM-Based Multi-Document Summarization Integrating Sentence Extraction with Bunsetsu Elimination. IEICE Transactions on Information and Systems E86-D(9), 399–408 (2003)
6. Karov, Y., Edelman, S.: Similarity-based Word Sense Disambiguation. Computational Linguistics 24(1), 41–59 (1998)
7. Edmundson, H.P.: New Methods in Automatic Extracting. Journal of the ACM 16(2), 264–285 (1969)
8. Kimoto, H.: Automatic Indexing and Evaluation of Keywords for Japanese Newspapers. The Transactions of the Institute of Electronics, Information and Communication Engineers 74, 556–566 (1991)
9. Luhn, H.P.: Statistical Approach to Mechanized Encoding and Searching of Literary Information. IBM Journal of Research and Development 1(4), 309–317 (1957)

Semantic Approach to Image Retrieval Using Statistical Models Based on a Lexical Ontology

Syed Abdullah Fadzli and Rossitza Setchi

Knowledge Engineering Systems Group, Cardiff University, Wales, UK
scesae@cf.ac.uk, setchi@cardiff.ac.uk

Abstract. The increasing amount of digital images available on the Internet has made searching, browsing, and organizing such resources a major challenge. This paper proposes a semantic approach to text-based image retrieval of manually annotated digital images. The approach uses statistical models based on Semantic DNA (SDNA) extracted from the structure of a lexical ontology called *OntoRo*. The approach involves three main techniques: (a) SDNA extraction, (b) word sense disambiguation using statistical models based on the extracted SDNA, and (c) applying semantic similarity measures using SDNA. The experiments performed show that the proposed approach retrieves images based on their conceptual meaning rather than the use of specific keywords in their annotations.

Keywords: Lexical Ontology, Word sense disambiguation, Semantic similarity.

1 Introduction

The widespread use of the internet has made available vast quantities of digital images, which need to be effectively indexed in order to be used in image retrieval applications. Traditional text-based image retrieval is predominantly based on processing the words in the annotations accompanying the images [1]. The availability of structured lexical ontologies such as Princeton's *WordNet* [2] and *OntoRo* 1.7 based on the Roget's Thesaurus [3,4,5], offers opportunities for improving traditional text-based search by adding explicit semantics to the images.

This paper focuses on exploratory concept-based search, which aims at collecting a set of images that provide inspiration, better understanding of a domain or clarification of a concept. "Relaxing, refreshing and peaceful" and "fresh and trendy" are examples of such exploratory queries. The paper aims at developing an approach that uses text analysis to semantically extract the contextual meaning of annotations and then perform conceptual searches by employing semantic similarity measures.

The paper is organized as follows: Section 2 reviews current research in the area of semantic similarity. Sections 3 and 4 present the method developed and its experimental evaluation. Section 5 concludes the paper.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 240–250, 2010.

2 Related Work

Several methods have been proposed for automatic image tagging and annotations. Sigurbjornsson and Van Zwol [6] use co-occurrence to recommend new tags that complement user-defined tags in Flickr. Garg and Weber [7] and Takashita et al. [8] extend this approach by personalizing the recommendation based on user history of tagging or Web browsing behaviour. These and other studies in automatic image annotation aim to support users in annotating images. This paper proposes an image retrieval method that can be applied on top of any of the annotation methods mentioned above. The method extracts the semantic elements of image annotations based on semantic indexing and semantic similarity measures.

Measures of semantic similarity have a long tradition in cognitive sciences and especially in psychology [9]. The approaches developed can be divided into two groups: distance based and corpus based approaches.

Distance-based approaches measure the semantic similarity by considering the distance between the words in a lexicon or a knowledge base. For example, Rada et al. [10] use the length of the shortest path connecting any two words as a measure of the semantic similarity between them. Leacock and Chodorow [11] count the path length in nodes rather than links, and adjust the result by taking into account the maximum depth of the taxonomy. Hirst and St-Onge [12] rely on the path length as well as on the number of changes of direction in the path; the changes are defined in terms of the *WordNet* semantic relations.

Corpus based approaches measure similarity between words by considering not only taxonomic information but also auxiliary information, such as word co-occurrence frequencies or other statistical information extracted from an available corpus. For example, Resnik [13] brings together taxonomy and corpus to propose a similarity measure which is based on the idea that the information shared by two concepts is determined by the information content of their nearest common ancestor in the taxonomy. Jiang and Conrath [14] and Lin [15] propose approaches based on edge weight by adding the information content, abstracted from a corpus, as a decision factor. The similarity between any two words is then computed as the sum of the edge weights along the shortest path linking the words.

The results reported in these studies show that approaches which use a corpus outperform distance-based approaches in terms of the correlation between the manual and computerized measuring of the semantic similarity between a pair of words. All studies mentioned above use *WordNet* as a lexical ontology. The use of ontologies to define high-level concept properties is proposed to improve the results of image retrieval systems [16].

Previous work closely related to the proposed method includes ontology-based concept indexing of images [4]. Apart from the actual methods used to identify and rank the concepts, the main differences to the work are that (i) they are dealing with free text that surrounds images in web pages, while this paper deals with user annotated images in image stock collections, (ii) they only consider the co-occurrences of concept numbers, while this paper considers the whole chain of the ontology's hierarchy.

3 Proposed Method

The approach proposed in this paper is a corpus-based approach, which uses a lexical ontology (*OntoRo* 1.7) and information content from a domain specific corpus (the *fotoLibra* collection). *OntoRo* is a lexical ontology created by Tang [3] while the extraction of SDNA from it was originated from Fadzli and Setchi [5]. Fig.1 shows the hierarchical structure of one of the meanings of the word *bird* in Roget's (bird as a symbol of *love*). According to *OntoRo*, the word *bird* has 13 different senses, including *bird* as an animal, as a traveller, as a symbol of feminism and as a symbol of love. Based on this premise, each sense of a word can have a unique SDNA. The SDNA of the example in Fig.1 is *6-37-83-887-1-5*; it is composed of the class number, section number, subsection number, head number, part-of-speech type and paragraph number.

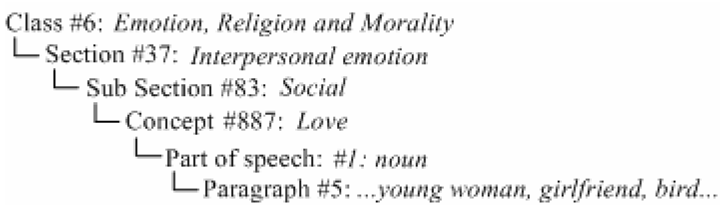


Fig. 1. Hierarchical structure of the word *bird* as a symbol of love

This paper proposes a semantic approach to image retrieval which uses a statistical method applied on a collection of images. As shown in Fig. 2, the approach involves Semantic Image Indexing (SII) and Semantic Image Search (SIS).

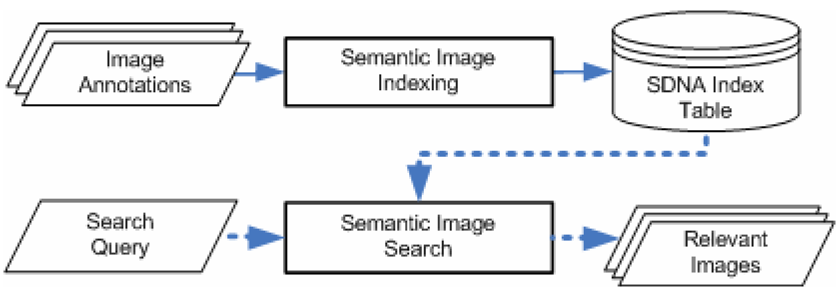


Fig. 2. Conceptual diagram showing the process flow of Semantic Image Indexing and Semantic Image Search

In the SII phase, the annotations of all images in the collection are analysed to produce an SDNA Index Table. Next, in the SIS phase, the SDNA Index is used to retrieve images related to search query. The next sections explain the two phases in detail.

3.1 Semantic Image Indexing (SII)

The SII phase starts with parsing all annotations in the collection into unique words, which are then compared against *OntoRo*. Words which do not exist in the lexical ontology are ignored. This phase consists of two steps: SDNA extraction and word sense disambiguation. At this stage, a set of SDNAs is produced (referred to as an image SDNA set); each SDNA in the set representing one possible meaning.

For example, the words "*leisure, holiday, beautiful, peace*"[1] are the keywords used to annotate the image shown in Fig. 3. These words have 6, 8, 12 and 10 SDNA strings respectively, producing a set of 36 SDNAs as shown in Table 1. Then, the most significant SDNA of each word in the annotation is selected as representing the contextual meaning of the image. The selection is based on the SDNA level co-occurrences within the whole image SDNA set. The process is repeated for all images in the collection, producing a SDNA Index further used in the SIS phase.

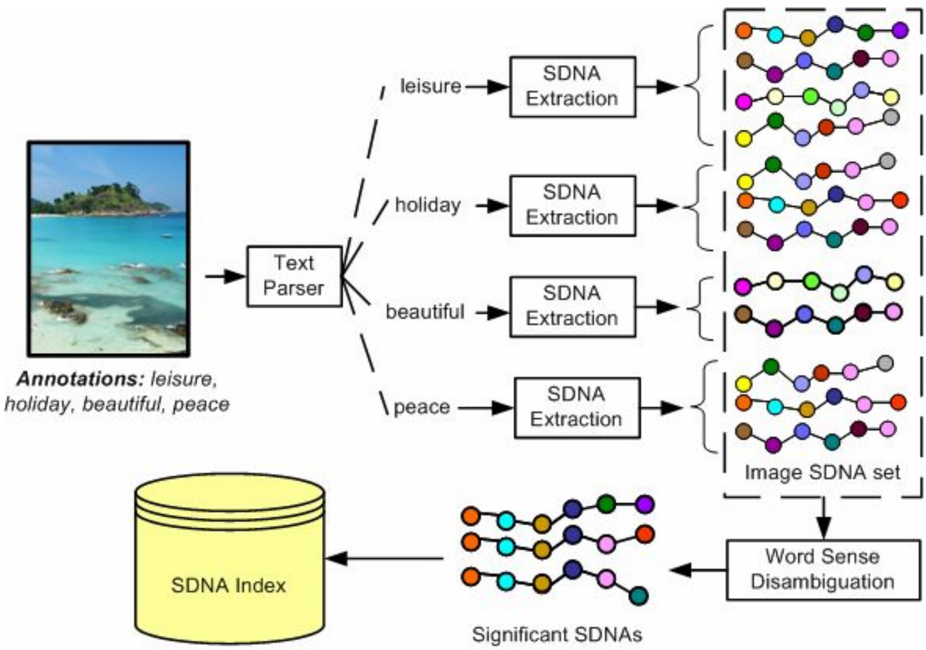


Fig. 3. Semantic Image Indexing

The selection of the most significant SDNA is based on the following empirical observations: (a) every word in the same *OntoRo*'s POS group (5^{th} level of SDNA) and paragraph (6^{th} level of SDNA) conveys the same idea in the same context; (b) words that express the same idea tend to belong to similar hierarchies and therefore share similar SDNA; (c) the most significant SDNA of a word corresponds to the

[1] In this example, an annotation of 4 words is used for clarity. In real implementations, images with short annotations are considered badly annotated and are excluded from the collection.

Table 1. List of SDNAs extracted from the annotation “*leisure, holiday, beautiful, peace*”

w_1: *Leisure*			
ID	SDNA	*freq(w_pd_q)*	*P(c)*
w_1d_1	***6-36-78-826-2-1***	**6.0000**	**0.1601**
w_1d_2	*3-15-48-376-2-1*	3.3333	0.0900
w_1d_3	*6-36-79-841-2-1*	2.0000	0.3310
w_1d_4	*6-36-79-843-2-1*	2.0000	0.0746
w_1d_5	*6-37-83-887-2-3*	0.6667	0.1864
w_1d_6	*4-25-58-575-2-1*	0.0000	0.0000
Total SDNA: 6			

w_2: *Holiday*			
ID	SDNA	*freq(w_pd_q)*	*P(c)*
w_2d_1	***6-36-78-837-1-1***	**9.0000**	**0.3899**
w_2d_2	*1-7-24-145-1-3*	7.1667	0.1093
w_2d_3	*5-28-65-683-1-1*	7.1667	0.0798
w_2d_4	*5-28-65-681-1-1*	7.1667	0.0720
w_2d_5	*6-36-78-837-2-1*	5.3333	0.3899
w_2d_6	*5-28-65-683-2-1*	3.5000	0.0798
w_2d_7	*2-9-33-190-1-1*	0.5000	0.0845
w_2d_8	*5-32-71-756-1-2*	0.5000	0.0652
Total SDNA: 8			

w_3: *Beautiful*			
ID	SDNA	*freq(w_pd_q)*	*P(c)*
w_3d_1	***6-36-78-826-1-1***	**9.6667**	**0.1601**
w_3d_2	*6-36-78-837-1-1*	9.0000	0.3899
w_3d_3	*1-7-24-145-1-3*	7.1667	0.1093
w_3d_4	*5-28-65-683-1-1*	7.1667	0.0798
w_3d_5	*5-28-65-681-1-1*	7.1667	0.0720
w_3d_6	*5-28-65-677-1-1*	2.1667	0.0827
w_3d_7	*5-28-65-679-1-1*	2.1667	0.0770
w_3d_8	*2-12-41-278-1-1*	1.1667	0.0573
w_3d_9	*6-39-95-988-1-9*	0.6667	0.1330
w_3d_{10}	*1-6-21-108-1-2*	0.5000	0.1546
w_3d_{11}	*1-6-22-137-1-2*	0.5000	0.0945
w_3d_{12}	*5-31-70-753-1-1*	0.5000	0.0246
Total SDNA: 12			

w_4: *Peace*			
ID	SDNA	*freq(w_pd_q)*	*P(c)*
w_4d_1	***6-36-78-826-1-1***	**9.6667**	**0.1601**
w_4d_2	*3-15-48-376-1-3*	3.3333	0.0900
w_4d_3	*3-15-48-399-1-1*	2.0000	0.0300
w_4d_4	*2-12-40-266-1-2*	1.1667	0.1261
w_4d_5	*5-30-69-730-1-2*	0.5000	0.1293
w_4d_6	*1-2-8-24-1-1*	0.5000	0.0820
w_4d_7	*1-4-13-60-1-1*	0.5000	0.0559
w_4d_8	*5-29-68-710-1-1*	0.5000	0.0492
w_4d_9	*5-29-68-717-1-1*	0.5000	0.0261
w_4d_{10}	*5-29-68-719-1-2*	0.5000	0.0194
Total SDNA: 10			

sense of that word in a particular context; and (d) the comparison of the SDNAs of any set of words can determine the extent to which they share information.

The proposed technique analyzes the information shared between the SDNAs of all image annotations by calculating the occurrences of SDNA levels within the image SDNA set. The shared information is captured using a λ function (1), where x and y denote different SDNAs in an SDNA set:

$$\lambda(x, y) = \begin{cases} 1, \text{if } x = y \\ 0, \text{if } x \neq y \end{cases} \tag{1}$$

The sum of the distinct[2] occurrences of each SDNA of every word, denoted by *freq*, is calculated using the following algorithm:

```
REPEAT for every SDNA in image SDNA set
  k := 1
  freq := 0
  REPEAT while k <= 6
    count := 0
    x := SDNA from level 1 to level k
    REPEAT for every word
      REPEAT for every SDNA in word
```

[2] Defined is the number of words with identical SDNAs, regardless of the number of identical SDNAs per word.

```
9 :           y := SDNA in word from level 1 to level k
10:           IF λ(x, y) = 1 THEN
11:             count := count + 1
12:             GO TO next word
13:           END IF
14:         END REPEAT
15:       END REPEAT
16:       freq := freq + count x k / 6
17:       k := k + 1
18:     END REPEAT
19:   END REPEAT
```

The algorithm considers 6 SDNA levels (see line 4). Once a match is found, indicated by $\lambda(x, y) = 1$ in line 10, the algorithm continues to the next word (line 12), ignoring the other identical SDNAs belonging to the same word. Thus only one distinct occurrence is counted per word.

In Table 1, w_1d_3 has 4 occurrences of $k=1$ and $k=2$ in w_1, w_2, w_3 and w_4 therefore $freq(w_1d_3)$= (4×1/6) + (4×2/6) = 2.0000. On the other hand, w_1d_1 has 4 occurrences of $k=1$, $k=2$ and $k=3$ in w_1, w_2, w_3 and w_4, and 3 occurrences of $k=4$ in w_1, w_3 and w_4. Therefore $freq(w_1d_1)$ = (4×1/6) + (4×2/6) + (4×3/6) + (3×4/6) = 6.0000. Table 1 contains all 36 SDNAs extracted for the four words used in the example. The SDNAs with the highest frequency (in bold) are the significant SDNAs. In the case of several SDNAs with the same highest frequency value, the most significant SDNA is the one with highest *P(c)*, which is obtained by the number of images where concept *c* appears (from the SDNA extraction process in step 1), calculated over the total number of images in a collection *I*:

$$P(c) = \frac{|\{i : c \in i\}|}{|I|} \tag{2}$$

3.2 Semantic Image Search (SIS)

Fig.4 illustrates the process flow in the SIS phase. First the search query provided by the user is parsed and analysed to produce significant query SDNAs. Then, the similarity between them and the SDNA strings from the SDNA Index table is measured, based on their information content as explained below.

Resnik [13] suggests that the information content of a concept *c* can be measured by the negative logarithm of its likelihood in a piece of information:

$$IC(c) = \log^{-1}(P(c)) \tag{3}$$

where *P(c)* is the probability of encountering an instance of concept *c* in an image annotation. *P(c)* is calculated as the number of images in which concept *c* appears, over the total number of images in the collection. The semantic similarity between the query SDNAs and a set of SDNAs in the SDNA Index is calculated using (1) as the sum of the co-occurrences' frequencies multiplied by *IC(c)*[3].

---

[3] *c* is the concept number of an SDNA obtained from the 4th level of the SDNA.

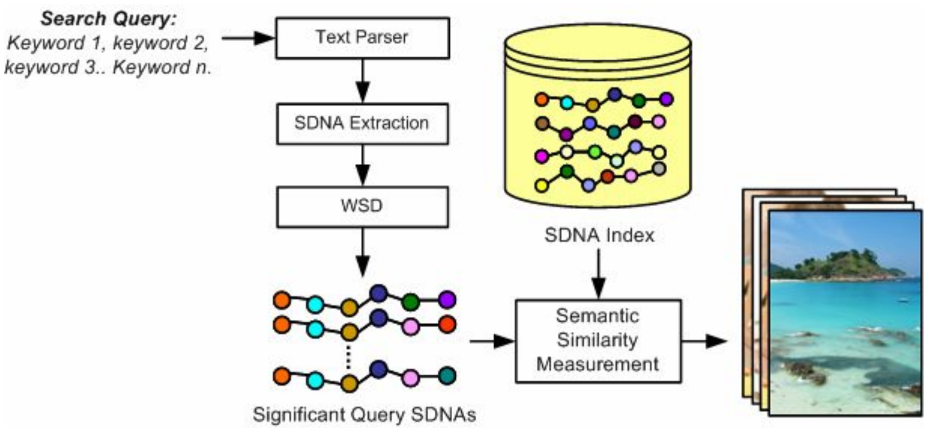


Fig. 4. Semantic Image Search

Let α be a set of significant query SDNAs, β represent a set of SDNAs for an image in a SDNA Index, then the semantic similarity *Sim(α, β)* is calculated using the following algorithm:

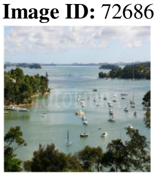
```
1 :  REPEAT for every SDNAs in α
2 :    k := 6
3 :    Sim := 0
4 :    REPEAT while k >= 1
5 :      count := 0
6 :      x := SDNA in α from level 1 to level k
7 :        REPEAT for every SDNAs in β
8 :          y := SDNA in β from level 1 to level k
9 :          IF λ(x,y) := 1 THEN
10:            Sim := Sim + (k/6 x IC(SDNA β))
11:            GO TO next SDNA in β
12:          END IF
13:        END REPEAT
14:      k := k - 1
15:    END REPEAT
16:  END REPEAT
```

Sim(α,β) determines the similarity between the significant query SDNAs and the SDNA set representing the individual images in the collection. Images with higher *Sim(α,β)* are considered more semantically relevant to the search query. *IC(c)* in line 10 is calculated using the concept number of SDNA from the SDNA Index. Different from the *freq* calculation, the SDNAs comparison in *Sim* starts with the highest level of the SDNA, followed by lower levels (shown in line 2 and 4).

4 Evaluation

So far there is no standard data set used in image retrieval community. Traditionally, Corel Stock Photos have been used widely for CBIR evaluation. However, most of researchers use different subsets of the 800 CDs collection, making it very hard to compare the performance [17]. In addition, Corel is copyrighted and no longer sold, so it is certainly not easily accessible. Caltech256 and the database of University of Washington (groundtruth) are among free image database that are available publicly, but only have a small number of images, in a limited number of categories. The images are also poorly annotated with the average of 5 words per image.

Table 2. Semantic Image Indexing sample result from the *fotoLibra* image collection

Image	*Annotations*	*Significant SDNA*	*Related Words*
Image ID: 72686 	**Annotations:** Opua; Bay Of Islands; Northland; North Island; New Zealand; NZ; Coast; Sea; Island; Yacht; Yachts; Boat; Boats; Leisure; Holiday; Vacation; Travel; Beauty; Beautiful; Calm; Serene; Serenity; Peace; Peaceful; Landscape; Landscapes; Summer 2005	*5-28-65-683-2-1* *2-12-40-266-2-2* *6-36-78-826-1-1* *2-12-40-269-3-1* *6-36-78-826-2-1* *2-12-40-269-3-2* *2-12-40-269-3-4* *2-12-40-269-2-1* *2-12-40-265-1-1*	holiday calm leisure yacht beautiful coast boat sea travel
Image ID: 167828 	**Annotations:** Lathkill; Dale; Derbyshire; Peak; District; River; Couple; Lying; Relaxing; Resting; Recreation; Holiday; Lush; Green; Water; Flow; Trees; Nature; Tranquil; Warm; Sunny; Mc0334; Mandy; Collins	*6-36-78-826-2-1* *6-36-78-837-1-1* *6-36-78-837-1-4* *3-14-46-348-1-1* *5-28-65-683-2-1* *3-15-48-379-2-2*	tranquil recreation green dale relaxing warm

Research collaboration with an online image library website, *fotoLibra* [18], has provided this study with 161,402 digital images. All images are manually annotated by the image owners. The proposed approach was tested with all 161,402 digital images annotated with an average of 20.8 words per image. 59.8% of the words used in the annotations exist in *OntoRo* 1.7 producing 670,898 significant SDNAs. Based on empirical evidence, two threshold are applied: (a) the minimum number of words in an image annotation should be greater than *4*; and (b) a threshold of *freq / MAX freq > 0.15* is used in extracting and selecting significant SDNAs to avoid SDNAs with weak concept disambiguation influencing the WSD process of other words.

The experiment result shows that 92.2% of all images produce at least 1 significant SDNA (the remaining 7.8% images are not indexed due to bad annotation which produces no significant SDNA). Table 2 shows two results obtained during the indexing process. Using the SDNA Index table produced, the proposed Semantic Image Indexing and Semantic Image Search techniques have been implemented to evaluate the image retrieval results using an experimental search query. For the purpose of this paper, four pairs of keywords have been chosen as a search query: (i) soft and vulnerable, (ii) refreshing and peaceful, (iii) arrogant and pride and (iv) tough and aggressive. Fig. 5 shows the top 5 results for the experimental queries sorted according to the *sim(α,β)* values. Most of the top 5 images retrieved do not contain the query words in their tags, but yet they were retrieved according to the semantic context of the image.

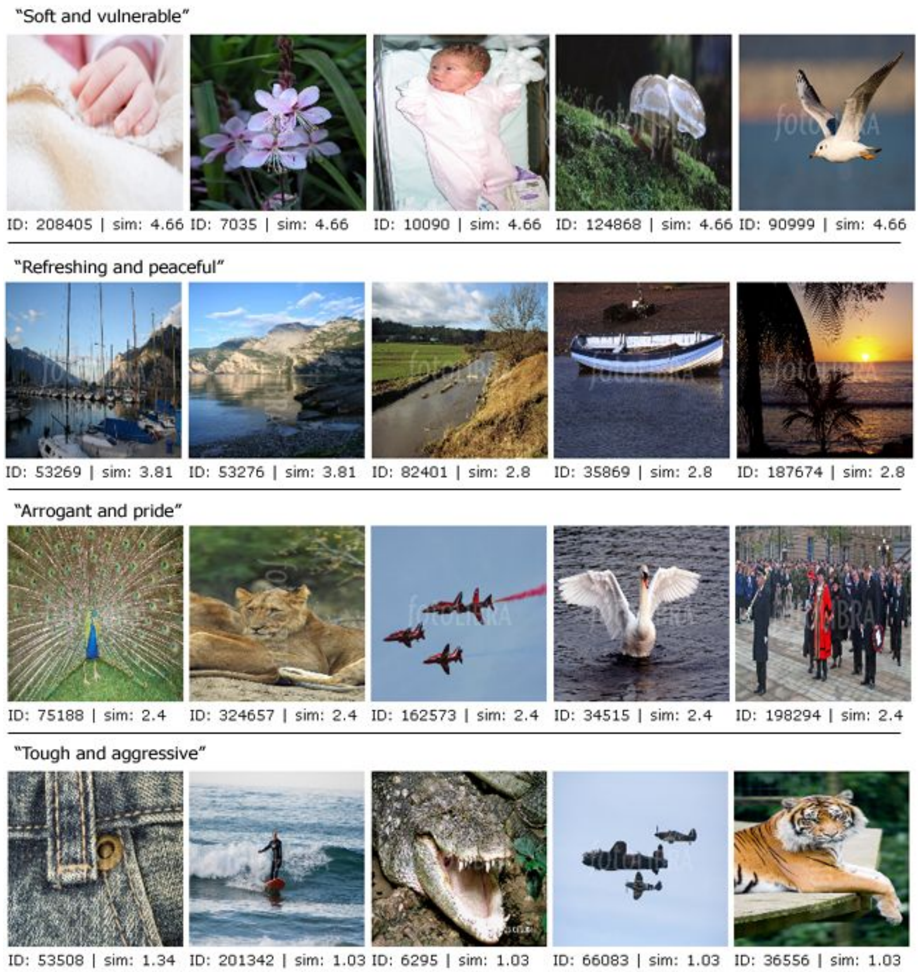


Fig. 5. Search results using Semantic Image Search

For example, the first image retrieved using *soft and vulnerable* as a query was tagged with *Baby, Child, Infant, Hand, Fingers, Blanket, Comfort, Gentle and Warmth*. While the first image retrieved by the *arrogant and pride* query was tagged with *Peacock, Fancy Birds, Fine Plumage, Display*. This shows the ability of the proposed approach to retrieve images based on their conceptual meaning rather than the use of specific keywords in their annotations. Naturally, these images would not be retrieved using keyword-based search because the terms are not contained in the annotation.

5 Conclusion

The proposed approach is designed to improve image retrieval using a semantic code called a Semantic DNA (SDNA). It is a key element of both the semantic word sense disambiguation technique and semantic similarity measure developed in this research. The approach uses a general-purpose lexical ontology and a large collection of annotated digital images.

The approach is tested with 161,402 annotated digital images. The Semantic Index produced contains 670,898 significant SDNAs, which have been disambiguated using an algorithm which computes the co-occurrences of the SDNA strings generated for each image in the collection. Another contribution of this paper is the semantic similarity measure developed, which uses information content calculated using a corpus containing the annotations of the images within the collection. Both the semantic word sense disambiguation technique and semantic similarity measure developed capitalize on the availability of a large lexical ontology with a well established hierarchical structure and conceptual grouping.

The experiments produced a diversity of images which show the ability of the proposed approach to retrieve images based on their conceptual meaning rather than the mention of specific keywords in their annotations. The experimental results also show that *OntoRo* 1.7 is a useful lexical ontology source for SDNA extraction although the proposed approach can be applied using any existing lexical ontology or taxonomy with a consistent and well-organized hierarchical structure.

Future work includes considering named entities and including more hierarchical levels of the lexical ontology. There are altogether 8 levels in Roget's hierarchy, but only 6 levels are used in this study. It is believed that the use of the lowest level of the Roget's hierarchy, i.e. the *sub-paragraph*, could improve the precision of both the semantic word sense disambiguation technique and semantic similarity measure proposed. In the experiment, a large number of the annotations were not processed as they could not be found in *OntoRo* 1.7. This number could be substantially reduced by employing named entity recognition. Further work also includes semantic retrieval of other multimedia content including audio and video.

Acknowledgments. Special thanks to all members of the Knowledge Engineering Systems group for their valuable help with this research; the Ministry of Higher Education of Malaysia and Universiti Darul Iman Malaysia for their sponsorship.

References

1. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Image Retrieval: Ideas, Influences, and Trends of the New Age. ACM Comput. Surv., 1–60 (2008)
2. Fellbaum, C.: WordNet: An Electronic Lexical Database. The MIT Press, Cambridge (1998)

3. Tang, Q.: Knowledge Management Using Machine Learning, Natural Language Processing and Ontology. Cardiff University, Cardiff (2006)
4. Setchi, R., Tang, Q., Bouchard, C.: Ontology-Based Concept Indexing of Images. In: Velásquez, J.D., Ríos, S.A., Howlett, R.J., Jain, L.C. (eds.) Knowledge-Based and Intelligent Information and Engineering Systems. LNCS, vol. 5711, pp. 293–300. Springer, Heidelberg (2009)
5. Fadzli, S.A., Setchi, R.: Semantic Approach to Text-based Image Retrieval Using a Lexical Ontology. In: Proc. of the Knowledge Engineering and Emotion Research Int. Conf., Paris, France, March 2-4, pp. 866–876 (2010)
6. Sigurbjornsson, B., van Zwol, R.: Flickr tag recommendation based on collective knowledge. In: WWW 2008: Proceeding of the 17th International Conference on World Wide Web, pp. 327–336. ACM Press, New York (2008)
7. Garg, N., Weber, I.: Personalized, interactive tag recommendation for flickr. In: RecSys 2008: Proceedings of the 2008 ACM Conference on Recommender Systems, pp. 67–74. ACM, New York (2008)
8. Takashita, T., Itokawa, T., Kitasuka, T., Aritsugi, M.: Tag Recommendation for Flickr Using Web Browsing Behavior. In: Taniar, D., Gervasi, O., Murgante, B., Pardede, E., Apduhan, B.O. (eds.) ICCSA 2010. LNCS, vol. 6017, pp. 412–421. Springer, Heidelberg (2010)
9. Goldstone, R.L., Son, J.: Similarity. In: Holyoak, K., Morrison, R. (eds.) Cambridge Handbook of Thinking and Reasoning, pp. 13–36. Cambridge University Press, Cambridge (2005)
10. Rada, R., Mili, H., Bicknell, E., Blettner, M.: Development and Application of a Metric on Semantic Nets. IEEE Transactions on Systems, Man and Cybernetics 19(1), 17–30 (1989)
11. Leacock, C., Chodorow, M.: Combining Local Context and WordNet Similarity for Word Sense Identification. In: Fellbaum, C. (ed.) WordNet: An Electronic Lexical Database, pp. 265–283. The MIT Press, Cambridge (1998)
12. Hirst, G., St-Onge, D.: Lexical Chains as Representations of Context for the Detection and Correction of Malapropisms. In: Fellbaum, C. (ed.) WordNet: An Electronic Lexical Database, pp. 305–332. The MIT Press, Cambridge (1998)
13. Resnik, P.: Using Information Content to Evaluate Semantic Similarity in a Taxonomy. In: 14th Int Joint Conf. on AI, Montreal, Canada, pp. 448–453 (1995)
14. Jiang, J.J., Conrath, D.W.: Semantic Similarity Based on Corpus Statistics and Lexical Taxonomy. In: 10th International Conference on Research on Computational Linguistics, ACL, Taiwan (1997)
15. Lin, D.: An Information Theoretic Definition of Similarity. In: International Conference on Machine Learning, Madison, WI (1998)
16. Liu, Y., Zhang, D., Lu, G., Ma, W.Y.: A survey of content-based image retrieval with high-level semantics. Pattern Recognition 40(1), 262–282 (2007)
17. Mueller, H., Marchand-Maillet, S., Pun, T.: The truth about Corel—evaluation in image retrieval. In: Lew, M., Sebe, N., Eakins, J.P. (eds.) CIVR 2002. LNCS, vol. 2383, pp. 38–50. Springer, Heidelberg (2002)
18. fotoLIBRA, `http://www.fotolibra.com`

IEC-Based Motion Retrieval System Using Laban Movement Analysis

Yuki Wakayama, Seiji Okajima, Shigeru Takano, and Yoshihiro Okada

Graduate School of ISEE, Kyushu University
744, Motooka, Nishi-ku, Fukuoka, 819-0395 Japan
{yuki.wakayama,seiji.okajima,takano,okada}@i.kyushu-u.ac.jp

Abstract. This paper proposes a motion retrieval system using Interactive Evolutionary Computation (IEC) based on Genetic Algorithm (GA) and motion features defined based on Laban Movement Analysis (LMA) used for the similarity calculation of motions in the system. The proposed IEC-based motion retrieval system allows the user to retrieve motions similar to his/her required motions easily and intuitively only through the evaluation repeatedly performed by scoring satisfaction points to retrieved motions without entering any search queries. The authors newly define LMA-based motion features to represent them as genes of GA used for the similarity calculation in the system. This paper also clarify that the LMA-based motion features are available as similarity features of motions by showing results of analyzing them using SOM visualization.

Keywords: Motion Retrieval, IEC, GA, Laban Movement Analysis.

1 Introduction

Recently, 3D CG animations have become in great demand for movie and video game industries. For the creation of 3D CG animations, character design is very important factor but very hard work. Especially its motion design is very laborious work. To solve this problem, our research group has already proposed a motion generation and editing system using Interactive Evolutionary Computation (IEC) based on Genetic Algorithm (GA) [1] that allows us to generate required motions easily and intuitively. However, since the system employs GA for IEC, it needs several existing motion data represented as genes used for the initial generation of GA. The user has to prepare several motion data those are similar to his/her required motions. To prepare such motion data, the easiest way is to retrieve those from a motion database. Hence, in this paper, we propose new motion retrieval system using Interactive Evolutionary Computation (IEC). This system allows the user to retrieve motions similar to his/her required motions easily and intuitively only through the evaluation repeatedly performed by scoring satisfaction points to retrieved motions without entering any search queries. IEC method of the system is based on Genetic Algorithm (GA), that is a heuristic algorithm to search optimized solution proposed by John Henry

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 251–260, 2010.

Holland in 1975, so that motion data should be represented as genes practically used as similarity features for the similarity calculation in the system. To extract motion features, we newly define them as mathematical expressions using Laban Movement Analysis (LMA), which is created by Rudolf Laban as a movement analysis system for the dance based on relationships between human body movements and emotions, because not only the idea of LMA is intuitively understandable for us but also motion features specified in LMA are possible to be represented as mathematical expressions. In this paper, we also clarify that the LMA-based motion features are available for the similarity calculation in the system by showing results of analyzing them using SOM visualization.

The reminder of this paper is organized as follows: First of all, we introduce IEC method based on GA after describing related works. Next, a feature extraction method for motion data and gene representation of motions are explained. After that, our proposed motion retrieval system is explained. In the last section, we conclude the paper.

2 Related Work

Our research purpose is to provide a motion retrieval system having an intuitive interface that makes it possible to retrieve motion data interactively and easily. For the motion retrieval, there are some researches. Müller et al. proposed content-based retrieval of motion capture data by using various kinds of qualitative features describing geometric relations [2]. Liu et al. proposed content-based motion retrieval algorithm by partitioning the motion database and constructing a motion index tree based on a hierarchical motion description [3]. Our proposed motion retrieval system uses our own defined LMA-based motion features. Fangtsou et al. proposed a feature extraction method of motions by using LMA [5]. However, this method does not use Shape feature of LMA. Our defined motion features also include Shape features. Yu et al. proposed a motion retrieval system which allows the user to retrieve motions via Labanotation [4]. This system requests the user to prepare motion data for the queries. Our proposed system does not request any search queries because the system employs IEC method. IEC [6] proposed the interactive calculation method that the user evaluates target data interactively, and finally the system outputs optimized solution based on its evaluated values. The remarkable point where IEC is useful is that a necessitated operation is only the evaluation against data by the user. The data is optimized based on the user's subjective evaluation. So, the system can consider requirements of the user. There are some experimental systems of IEC researches. Ando, et al. proposed a music composition support system for the classical music using IEC [7]. Faffi, et al. proposed a design system for Microelectro-mechanical Systems (MEMS) using IEC [8]. Nishino, et al. proposed an integral 3D-CG contents system based on IEC [9]. By their proposed IEC framework, it is possible to create various attributed 3D-CG contents. Usually, IEC method is based on GA. There is a system [10] that generates some various walk motions using GA. However, there is not any motion data retrieval system using IEC that retrieves and presents motion data according to the user requirement from a motion database. In this paper, we propose such a motion retrieval system using IEC method based on GA.

3 IEC Method Based on GA

IEC is a general term for methods of evolutionary computation that use human interactive evaluation to obtain optimized solutions. In IEC method, first of all, a system presents some candidate solutions to the user, and then the user evaluate them by giving a numerical score depending on his/her requirement. After that, the system again presents some solutions more suitable for the user requirement solved by a certain algorithm like GA. After several trials of this operation, the user obtains his/her most desirable solutions. In this way, since IEC method is intuitive and useful to deal with problems depending on human feelings, we decided to employ IEC method based on GA for our motion retrieval system. In IEC method based on GA, the system encodes target data (individual) as genes, and prepares a number of them as current generation. There are various gene coding methods according to types of problems and it is not necessary to obey biological gene behavior. Next, the user evaluates each data by the scoring in an interactive manner and then fitness values of genes are calculated according to the corresponding evaluated scores. After that, the system applies a gene operation, which is any combination of *Selection*, *Crossover* and *Mutation*, to each genes to adjust them to their fitness values. As a result of this operation, new set of genes is generated as next generation. This generation is replaced as current generation. The evaluation operation by the user will be performed repeatedly until his/her desirable solutions are obtained.

4 Motion Features Using Laban Movement Analysis

As described above, we have been developing a motion retrieval system using IEC method based on GA. To use GA, it is necessary to represent motions as their corresponding genes. For that, we newly define motion features as mathematical expressions based on the idea of LMA.

4.1 Motion File Format

First of all, we introduce motion file formats. There are some typical motion data formats. For example, BVH file format is employed by Biovision Co., Ltd. and ASF-AMC file format is employed by Acclaim Co., Ltd. In our system, we use BVH file format because it is supported by a lot of commercial 3D-CG animation software like Alias Motion Builder, 3dsMAX Character studio, and Poser. This file format consists of two sections, skeleton information (HIERARCHY section) and motion information (MOTION section). The HIERARCHY section defines an initial pose of a skeleton that includes bone lengths as offset values. The

MOTION section defines time series data about sequential poses of a skeleton in a motion.

4.2 Laban Movement Analysis

In Laban Movement Analysis (LMA), human body movement is explained by features of *Effort* and *Shape*. Each feature has two opposite forms which are *Fighting Form* and *Indulging Form*. Fighting Form means a strong, linear and sudden movement, and Indulging Form means a weak, spiral and sustained movement.

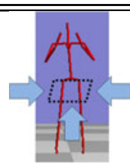
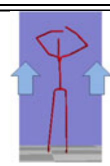
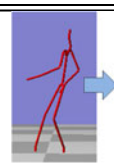
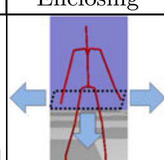
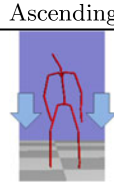
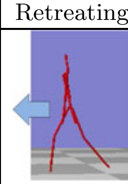
Effort. *Effort* is a mechanical feature of human movement. Effort has three elements which are *Weight*, *Space* and *Time* elements, each of which has two opposite forms (Table. 1). What these elements mean are as follows.

- Weight: Energy and Speed of movement.
- Space: Spatial Pattern of movement.
- Time: Temporal Alternation of movement.

Shape. *Shape* is a shape feature of the whole body movement. Shape has three elements which are *Table plane*, *Door plane* and *Wheel plane*. Each of them also has two opposite forms (Table. 1). Shape feature means spread and movement of body silhouette projected on each of the following three planes.

- Table plane: Spread of body silhouette projected on the transverse plane.
- Door plane: Spread of body silhouette projected on the frontal plane.
- Wheel plane: Movement of body silhouette projected on the sagittal plane.

Table 1. Effort and Shape elements

	Weight	Space	Time	Table Plane	Door Plane	Wheel Plane
Fighting Form	Strong	Direct	Sudden	 Enclosing	 Ascending	 Retreating
Indulging Form	Light	Indirect	Sustained	 Spreading	 Descending	 Advancing

4.3 LMA-Based Motion Features

To extract body movement features from motion data, we define them as mathematical expressions according to the idea of motion features specified in LMA. In our system, we focus on end-effectors of a human body to extract its features, i.e., its root, left hand, right hand, left foot and right foot.

Feature extraction method for Effort is as follows.

1. Weight
   Weight element in LMA represents active emotion derived from the energy and speed of movement. To extract this feature, we focus on speeds of end-effectors in a motion.

   Let F be the number of motion frames and $v_n(f)$ be the speed of an end-effector n in a motion frame f.

   Then, we calculate Weight feature $Weight_n$ of an end-effector n by the next equation.
$$Weight_n = \frac{1}{F}\sum_{f=1}^{F}|v_n(f)| \tag{1}$$
2. Space
   Space element in LMA represents concentrated or unconcentrated emotion derived from the trajectory of movement. To extract this feature, we focus on distributions of speed vectors of end-effectors in a motion and define Space feature value as a norm of a covariance matrix of all speed vectors about each end-effector in a motion.

   Let $V(=[V_1^n V_2^n V_3^n])$ be a speed vector in $\mathbb{R}^3$ and $\mu_i(=E(V_i^n))$ be the mean of V_i^n about an end-effector n.

   Then, we calculate Space feature $Space_n$ as a norm of a covariance matrix A_n of a speed vector of an end-effector n by the following equations. In the practical calculation, each of V_1^n, V_2^n and V_3^n means a vector about the complete frames in a motion.
$$A_n = \begin{bmatrix} E[(V_1^n-\mu_1^n)(V_1^n-\mu_1^n)] & \cdots & E[(V_1^n-\mu_1^n)(V_3^n-\mu_3^n)] \\ \vdots & \ddots & \vdots \\ E[(V_3^n-\mu_3^n)(V_1^n-\mu_1^n)] & \cdots & E[(V_3^n-\mu_3^n)(V_3^n-\mu_3^n)] \end{bmatrix} \tag{2}$$
$$Space_n = ||A_n|| = \max_{1\le j\le 3}\sum_{i=1}^{3}|a_{ij}^n| \tag{3}$$
3. Time
   Time element represents tension emotion derived from sudden or sustained movement. To extract this feature, we calculate the acceleration of a motion.

   Let F be the number of motion frames and $a_n(f)$ be the acceleration of an end-effector n in a motion frame f.

   Then, we calculate Time feature $Time_n$ of an end-effector n by the next equation.
$$Time_n = \frac{1}{F}\sum_{f=1}^{F}|\frac{d}{df}a_n(f)| \tag{4}$$

As for the feature of Shape, we use the mean about all frames of RMS (Root Mean Square) of distances between each end-effector and the root (Center of Mass) of a skeleton in each motion frame.

Let F be the number of motion frames, N be the number of end-effectors and $P(n, f)$ be the coordinate value of an end-effector n in a motion frame F. Then we calculate each Plane feature by the following equations.

$$TablePlane = \frac{1}{F}\sum_{f=1}^{F}\sqrt{\frac{1}{N}\sum_{n=1}^{N}(P_x(n,f) - P_x(root,f))^2} \tag{5}$$

$$DoorPlane = \frac{1}{F}\sum_{f=1}^{F}\sqrt{\frac{1}{N}\sum_{n=1}^{N}(P_y(n,f) - P_y(root,f))^2} \tag{6}$$

$$WheelPlane = \frac{1}{F}\sum_{f=1}^{F}\sqrt{\frac{1}{N}\sum_{n=1}^{N}(P_z(n,f) - P_z(root,f))^2} \tag{7}$$

4.4 Gene Representation

We represent motions as their corresponding genes using the LMA-based motion features. As for each of the three types of Effort features, we employ the maximum among corresponding feature values of all end-effectors rather than the mean of them because in this case we obtain better results of the motion similarity analysis using SOM visualization. Therefore, each gene consists of six chromosomes as shown in Fig.1. On the other hand, the allele is represented as a motion index number assigned in a motion database because each motion is allocated a unique index number. When the number of motions in the database is N, the allele is represented as an integer value from 0 to $N-1$.

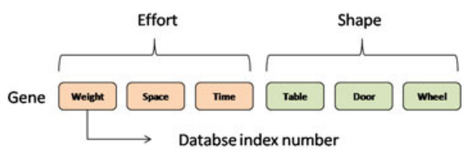
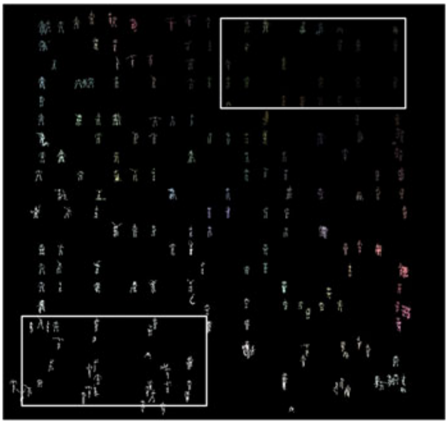


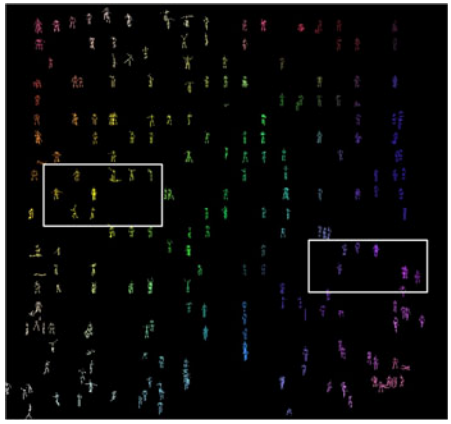
Fig. 1. Gene representation using LMA-based features

4.5 Visualization and Analysis

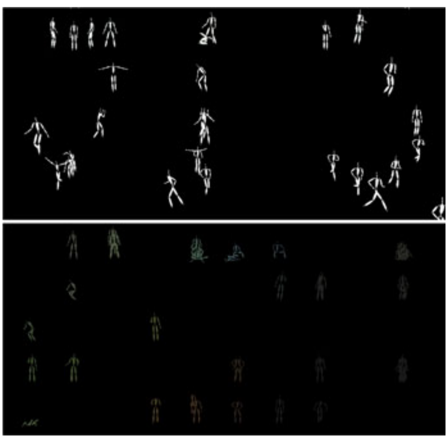
To analyze effectiveness of our defined LMA-based motion features for the motion data retrieval, we apply Self-Organizing Maps (SOM) visualization to motion data using their LMA-based features as the feature vectors of SOM. Using SOM layout, similar feature data are located in the same area and it arranges each data in grid, and thus SOM is useful for analyzing similarities among data records



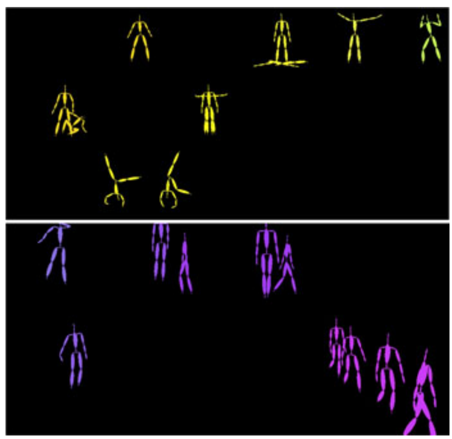
(a) SOM layout of motion database colored by Effort feature: Red assigned to the Weight feature value and green assigned to the Space feature value and blue assigned to the Time feature value.



(b) SOM layout of motion database colored by Shape feature: Red assigned to the TablePlane feature value and green assigned to the DoorPlane feature value and blue assigned to the WheelPlane feature value.



(c) Zoom-in figures of the two regions within rectangular lines in (a).



(d) Zoom-in figures of the two regions within rectangular lines in (b).

Fig. 2. SOM layout of motion database colored by Effort (a) and Shape (b) feature values. (c) and (d) are zoom-in figures of the regions within rectangular lines in (a) and (b).

of a database. Fig.2 shows SOM layout of our motion database including 296 motions. Each motion is colored according to its Effort and Shape features.

The color gradation in Fig.2(a) illustrates there are positive correlations between Effort feature values. Besides, this color gradation indicates emotions expressed in human movements become more active with the color gradient from black at top-right to white at bottom-left. Actually, as shown in Fig.2(c), bottom-left motions become more active compared to top-right motions. By contrast, the color gradation in Fig.2(b) illustrates there are poor correlations between Shape feature values. Consequently, motions are divided into similar shape motion groups clearly. For example, motions such as cartwheel, open-arms or something are drawn yellow in Fig.2(d) (upper) which are zoom-in figures of the regions within rectangular lines in Fig.2(b). This means these motions have high TablePlane feature value and DoorPlane feature value. This is intuitively correct. Similarly, motions including mainly walk motions are drawn blue or purple in Fig.2(d) (lower). This means these motions have high TablePlane feature value and WheelPlane feature value. This is also intuitively correct. These observations may clarify that our proposed LMA-based motion features introduced in the previous section 3 are available as similarity features for motion data.

5 Motion Retrieval System

In this section, we explain our proposed IEC-based motion retrieval system. Fig.3 and Fig.4 show the overview and a screen snapshot of the motion retrieval system respectively. As the preprocessing, the system creates a database of LMA-based features from a motion database. In this process, index numbers of motions are assigned to each LMA-based features and the gene is represented as combination of index numbers. The allele is represented as an index number of a motion. When the user runs the system, it randomly generates genes and retrieves the corresponding nine motions appeared on a screen. The user evaluates each of

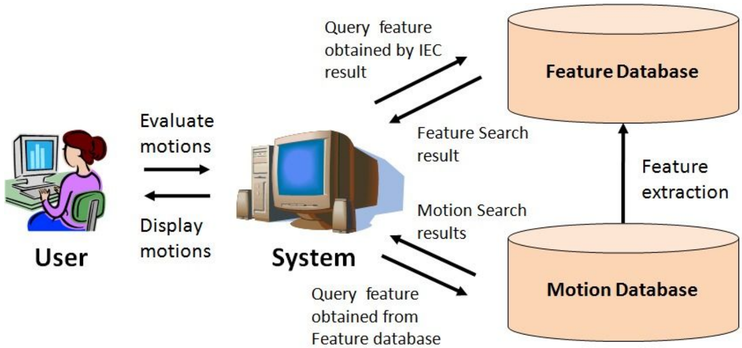


Fig. 3. Overview of motion retrieval system

Fig. 4. Screenshot of motion retrieval system

these motions by three stage scoring, i.e., good, normal and bad. This evaluation is performed only by mouse clicks on thumbnails of the motions. After the evaluation, the system automatically applies GA operations, i.e., selection, crossover and mutation, to the genes in order to generate the next generation. And then, the system searches motion data having LMA-based features similar to the features of the newly generated genes in order to presents them to the user as his/her more desirable motions. In this system, we choose the cosine similarity as a gene similarity measure.

After several trials of the above evaluation process, the user will obtain his/her most desirable motions without any difficult operations. The number of trials cannot be specified because it depends on types of motions the user wants and on the number of motion data in a database. However, we can show that an actual time spent for one GA operation is less than ten milli-seconds and search time to present next generation is around 80 milli-seconds in the case of 300 motion data of a database. So, the user manipulates the system without feeling any impatience.

6 Concluding Remarks

In this paper, we proposed the motion retrieval system using IEC based on GA and motion features defined based on LMA. The proposed IEC-based motion retrieval system allows the user to retrieve motions similar to his/her required motions easily and intuitively only through the interactive operation to evaluate retrieved motions without any difficult operations. For the motion similarity calculation of the system, we newly defined LMA-based motion features and clarified that those features are available as similarity features by showing results of analyzing them using SOM visualization.

As future works, we will evaluate our proposed motion retrieval system by asking several subjects to practically use it and consulting them. We will improve GUI of the system to make it more useful. We also have a plan to provide the proposed system as one of the web services.

References

1. Wakayama, Y., Takano, S., Okada, Y., Nishino, H.: Motion Generation System Using Interactive Evolutionary Computation and Signal Processing. In: Proc. of 2009 International Conference on Network-Based Information Systems (NBiS 2009), pp. 492–498. IEEE CS Press, Los Alamitos (2009)
2. Müller, M., Röder, T., Clausen, M.: Efficient content-based retrieval of motion capture data. In: Proc. of ACM SIGGRAPH 2005, pp. 677–685 (2005)
3. Liu, F., Zhuang, Y., Wu, F., Pan, Y.: 3D motion retrieval with motion index tree. Journal of Computer Vision and Image Understanding 92(2-3), 265–284 (2003)
4. Yu, T., Shen, X., Li, Q., Geng, W.: Motion retrieval based on movement notation language. Journal of Computer Animation and Virtual Worlds 16(3-4), 273–282 (2005)
5. Fangtsou, C.T., Huang, W.S.: Analysis and Diagnosis of Human Body Movement Efforts Based on LMA. In: Proc. of 2009 International Conference on Business And Information, BAI 2009 (2009)
6. Takagi, H.: Interactive Evolutionary Computation: Fusion of the Capacities of EC Optimization and Human Evaluation. Proc. of the IEEE 89(9), 1275–1296 (2001)
7. Ando, D., Dahlstedt, P., Nordahl, M.G., Iba, H.: Computer Aided Composition for Contemporary Classical Music by means of Interactive GP. Journal of the Society for Art and Science 4(2), 77–87 (2005)
8. Kamalian, R., Zhang, Y., Takagi, H., Agogino, A.M.: Reduced human fatigue interactive evolutionary computation for micromachine design. In: Proc. of 2005 International Conference on Machine Learning and Cybernetics, pp. 5666–5671 (2005)
9. Nishino, H., Aoki, K., Takagi, H., Kagawa, T., Utsumiya, K.: A synthesized 3DCG contents generator using IEC framework. In: IEEE International Conference on Systems, Man and Cybernetics, pp. 5719–5724 (2004)
10. Lim, I.S., Thalmann, D.: Pro-actively Interactive Evolution for Computer Animation. In: Proc. of Eurographics Workshop on Animation and Simulation, pp. 45–52 (1999)
11. Kohonen, T.: Self-Organizing Maps. Springer, Japan (1996)

Automatic Composition of Personalized Appreciation Route Based on Semantic Relationship between Exhibits in Museum

Chihiro Maehara[1], Kotaro Yatsugi[2], Daewoong Kim[3], and Taketoshi Ushiama[3]

[1] School of Design, Kyushu University
[2] Graduate School of Design, Kyushu University
[3] Faculty of Design, Kyushu University

Abstract. Recently, many studies have been reported on ranking and recommendation of a variety of digital contents on the Internet. On the other hand, exhibitions in museums can be considered as actual contents in the real world, so we think that some techniques of ranking and recommendation can be applicable to appreciate exhibitions in a museum effectively and efficiently for a visitor based on his/her interests. Today, information devices as exhibition guide systems to support visitor's appreciation are introduced in many museums. In this paper, we propose an approach for composing personalized appreciation route matched to user's interest and situation, using the information device. Our system supports a user to appreciate an exhibition well, and arouses user's interest by composing a route based on semantic relationship between exhibits.

Keywords: Personalized route composition, museum, HITS algorithm, exhibition guide system.

1 Introduction

Recently, there are a variety of digital contents such as web page, digital photograph, and video on the Internet and the number of them is increasing rapidly. In order to find suitable contents for a user, many studies have been reported on recommendation of digital contents. Appropriate techniques for recommendation are different depending on the characteristics of targeted contents. On the other hand, various contents exist not only on the Internet but also in the real world. For example, exhibitions in a museum can be considered as an environment to browse actual contents such as pictures, statues and so on.

Today, many museums introduce mobile information devices as guide system to support visitors to appreciate an exhibition. For example, the National Museum of Nature and Science in Tokyo[1] introduced sound guide systems which support four kinds of languages. Such mobile information devices provide visitors detailed explanations of exhibits which they are interested in. This helps the visitors to understand the exhibition well. Conventionally visitors in a museum basically appreciate exhibition along the route composed by curators. Such

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 261–270, 2010.

routes are well organized because they are composed based on domain knowledge. However, there is a problem that it might be inappropriate to show the same route to all visitors because their interest and knowledge on the exhibition are different in each visitor.

In this paper, we propose an approach for composing of a personalized appreciation route suitable for visitor's interests and situation by considering semantic relationship between exhibits with a mobile information device. Our goal is to support a visitor to appreciate exhibition well, and arouse his/her interests by making his/her understand the culture which surround each exhibit.

This paper is composed of as followings: Section 2 describes related works. Section 3 proposes the approach of our system. Section 4 describes the semantic relationship between exhibits. Section 5 describes the details of our system. Section 6 describes the composition of personalized appreciation route. Section 7 evaluates our system by experiment and consideration. Section 8 describes conclusion and future works.

2 Related Works

ubiNext[2] is a museum guide system which supports visitor's active learning experience in a museum through the Internet services. It recommends some exhibits for the next appreciation to a visitor based on the visitor's interests such as gGoghh and gthe impressionistsh, the evaluations on the exhibitions by the visitor, and appreciation history of the visitor. Our approach proposed in this paper is different from ubiNext in terms of composing personalized appreciation route matched to the individual visitor.

Koyanagi et al.[3] developed the guide system for the guest of the Hakone Open-Air Museum. It was designed for representing the most recommended path which enables the visitor to appreciate the maximum number of sculptures within the given time interval. Since the combinations of actual paths in a museum are limited, they apply one by one inquiry method to seek an optimum solution. Our system composes an appreciation route based on the structure of semantic relationships between exhibits rather than the layout of exhibits.

3 Approach

As follows, we assume the situation that our system would be used. First, a visitor receives a mobile information device at the information desk in a museum, and inputs one or more subjects which the visitor is interested in and scheduled staying time(Fig.1).

Then, the system automatically composes a personalized appreciation route sufficient to the information given by the visitor, and shows it. Since the route depends on visitor's interest and scheduled staying time, exhibits included in it might be different in each visitor. For example, if a visitor is interested in a picture, the route will be mainly composed by pictures. Moreover, in order

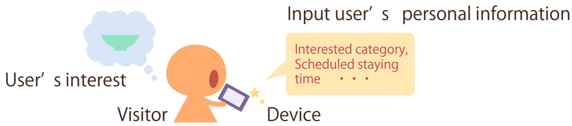


Fig. 1. Input visitor's personal information

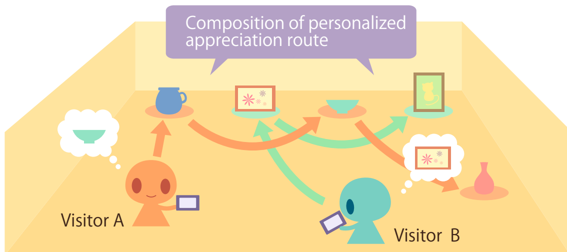


Fig. 2. Showing of personalized appreciation route

to appreciate the whole of exhibition in scheduled staying time, the number of exhibits will be limited(Fig.2).

In this paper, a personalized appreciation route is composed based on semantic relationships between exhibits, and the exhibits are related to others in terms of cultural influence. Features are extracted from the structure of semantic relationships, and some selected exhibits are recommended to user.

4 Relationship between Exhibits

In a museum, we can find various semantic relationships between exhibits such as shape, object drawn in it, usage and so on. Many conventional systems also use those relationships for recommendations. Typical relationships used in them are age, region, artist, and technique on the exhibits. These relationships are representative indicators showing characteristic features of the exhibits.

On the other hand, we relate exhibits with cultural influence between them. For example, suppose that two exhibits A and B are similar in their theme, technique and so on, and A was produced after B. Then, when the exhibit A can be thought as if it progresses technically or develops culturally by adding new elements more than B, we define that A is influenced by B. This cultural influence is shown by the arrow between exhibits(Fig.3). These relationships are derived based on the keywords included in explanation sentences of the exhibits.

Moreover, such influences are not limited between the exhibits in museum. For example, there are some exhibits may be influenced by culture of foreign country. So we introduce relationships between exhibits and conceptual entities such as culture. Using such relationships we can relate exhibits in different field such as picture and earthenware(Fig.4).

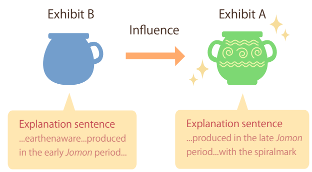


Fig. 3. Influence between exhibits

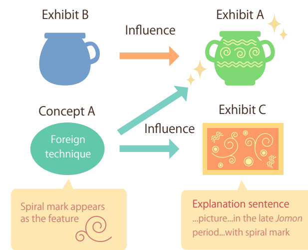


Fig. 4. Influence from conceptual entity

An example of personalized appreciation route for a user who is interested in earthenware is shown in Figure 5. In this route, not only the exhibits A and B which are in the same category that the user is interested in, but also the exhibit C that was produced by the same influence as A is included in the route. The meaning of the pattern drawn on the exhibit A is hard to understand for the user in case the user appreciates only earthenware, but the exhibit C, which has a close relationship with the exhibit A, supports the user to understand it. Moreover, for the user who is interested in earthenware, the exhibit, which is influenced by the same culture, may arouse user's interest in another category in which the user has little interest.

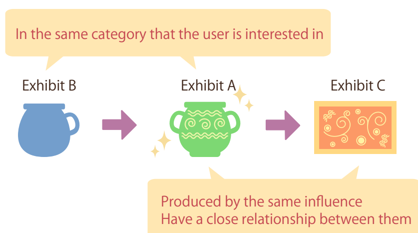


Fig. 5. Example of personalized appreciation route

Thus, by considering cultural influence between exhibits, our system can recommend not only the exhibits in category in which a user is interested, but also the exhibits having close relationships with them though the user is not much interested in them. As a result, our system enable a user to understand the culture surrounding exhibits, and to arouse user's interest.

5 Extraction of Features Using Relationships between Exhibits

This section describes how to extract features of exhibits. Firstly a directed graph is composed. Each node in the graph represents individual exhibit, and

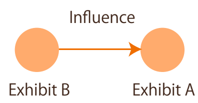


Fig. 6. Example of influence between exhibits

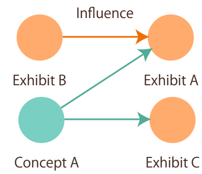


Fig. 7. Example of influence between exhibits and concepts

each edge represents an influence between a pair of exhibits. Secondly the HITS algorithm[4] is applied on the graph, and features of the exhibits are obtained. They are used for composing personalized appreciation routes. HITS is an algorithm for ranking web pages based on the link structure of a target set of web pages. Using HITS, we can obtain the hub and authority score on each Web page in a target set. When we regard the relationships between web page and link on the HITS algorithm as the influences between exhibits, we can think that a high hub score expresses an important exhibit or concept which influences many exhibits, and an exhibit which has a high authority score integrates many important influences. Our system composes an appreciation route by using hub and authority scores as importance and integration scores of exhibits respectively.

First, relationships between exhibits are set by curators. A directed graph is composed based on it. Each node in the graph represents an exhibit, and each edge represents the influence between exhibits. Figure6 shows an example of graph in which the exhibit A is influenced by B. Figure7 shows an example of graph in which the exhibits A and C are influenced by the concept A.

Next, exhibits which the user is interested in are selected from all the exhibits, and define them as the root set R. All the children nodes which have links from any node included in R, and all the parents nodes which have links to any node included in R are searched, and define them as the base sub-graph G(Fig.8). For an exhibit v, the importance score $a(v)$ and the integration score $h(v)$ are calculated by the following formula with all the nodes included G.

$$a(v) = \sum_{\omega, v \rightarrow \omega} h(\omega) \tag{1}$$

$$h(v) = \sum_{\omega, \omega \rightarrow v} a(\omega) \tag{2}$$

The importance score of an exhibit shows how strongly it influences others, and the score is defined as the sum of the integration scores of exhibits influenced by the exhibit(Fig.9). On the other hand, the integration score of an exhibit shows how it is influenced by other exhibits having high importance score, and the score is defined as the sum of the importance scores of exhibits which influence the exhibit(Fig.10). These scores converge by normalization and repetition of the formula (1) and (2).

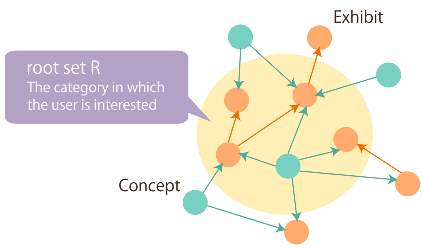


Fig. 8. Example of base sub-graph

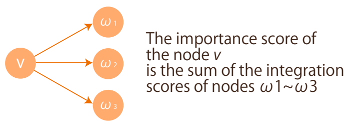


Fig. 9. Illustration of expression(1)

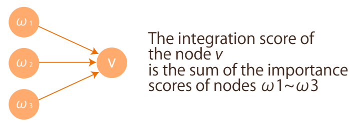


Fig. 10. Illustration of expression(2)

6 Composition of Personalized Appreciation Route

An exhibit whose integration score is high integrates important influences on a sub-graph G. By recommending such exhibits, the personalized appreciation route is suitable for learning how the exhibits that the user is interested in are influenced and evolved. Our approach uses only integration scores for composing a personalized appreciation route. Importance scores are used for calculating integration scores.

In order to compose a personalized appreciation route, all the nodes representing concepts are removed from sub-graph G, and save only the nodes representing actual exhibits which exist in the museum. Next, in order to fit the length of appreciation route to the user's scheduled staying time, appropriate numbers of nodes are selected in descending order of integration score. Finally, the selected exhibits are listed in the chronological order(Fig.11). This is because cultural influences are basically given from old one to new one, and in

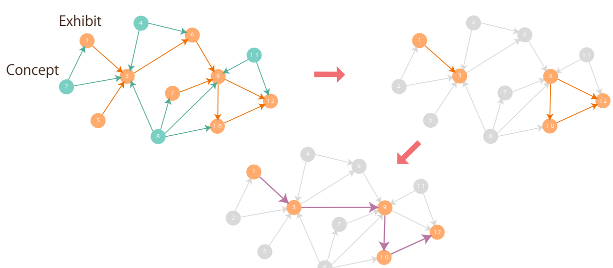


Fig. 11. Example of composition of appreciation route

order to understand evolution of culture, it is reasonable to appreciate exhibits in the chronological order. Through these operations, in case the user's scheduled staying time is short, appreciation route may be composed mainly by the characteristic exhibits.

7 Evaluation

7.1 Experiment

We did an experiment to evaluate our technique. We composed a personalized appreciation route which fits to user's interest and scheduled staying time using our technique. We used representative art works in Japanese art history[6] instead of exhibits in a museum. We supposed that the user is interested in "Buddhism picture" and tried to compose appreciation route which makes user understand them deeply.

First,16 Buddhism pictures are consisted in as the root set R. Next, the base sub-graph G was composed of 40 exhibits including 14 pictures and crafts which have close relationships between Buddhism pictures, and 10 nodes expressing concepts taken from explanation sentences. In this experiment, these operations were worked by hand. On this sub-graph, the importance scores and integration scores were calculated by a prototype system implemented with the PHP programming language. The result is shown in Figure12.

Node num	Category	Age	a(v)	h(v)
1	craft	Asuka	0.001649	0.001934
2	craft	Asuka	0.001649	0.001934
3	Buddhism picture	Asuka	0.018793	0.004154
4	Buddhism picture	Nara	0.001436	0.006867
5	craft	Nara	0	0.000981
6	craft	Nara	0	0.002259
7	Buddhism picture	Nara	0.010136	0.003621
8	Buddhism picture	early Heian	0.010136	0.025736
9	Buddhism picture	early Heian	0.010136	0.025736
10	Buddhism picture	early Heian	0.010136	0.025736
11	Buddhism picture	early Heian	0.057165	0.025736
12	craft	early Heian	0.015887	0.023674
13	fable picture	late Heian	0.033935	0.043054
14	Buddhism picture	late Heian	0.031113	0.088042
15	picture scroll	late Heian	0.033935	0.039101
16	craft	late Heian	8.71E−05	4.24E−02
17	craft	late Heian	0	0.006946
18	Buddhism picture	late Heian	0.033935	0.082733
19	Buddhism picture	late Heian	0.026301	0.090257
20	picture scroll	late Heian	5.09E−08	8.33E−02

Node num	Category	Age	a(v)	h(v)
21	Buddhism picture	Kamakura	0.021211	0.061706
22	Buddhism picture	Kamakura	0.021211	0.054349
23	Buddhism picture	Kamakura	6.38E−05	0.054349
24	Buddhism picture	Kamakura	0.026301	0.068991
25	Buddhism picture	Kamakura	6.38E−05	0.068991
26	ink painting	Kamakura	8.71E−05	0.026355
27	Buddhism picture	Muromachi	0	0.000191
28	ink painting	Muromachi	0	0.000261
29	picture scroll	Muromachi	0	1.02E−07
30	craft	Muromachi	0	0.000261
31	concept		0.007246	0
32	concept		0.004487	0
33	concept		0.003911	0
34	concept		0.116263	0
35	concept		0.182165	0
36	concept		0.190947	0
37	concept		0.129201	0.001278
38	concept		0.000238	0.00E+00
39	concept		0.000174	0
40	concept		5.09E−08	0.039101

Fig. 12. Importance scores and Integration scores for the test set

On the basis of this result, a personalized appreciation route was composed by extracting exhibits having high integration scores and aligning in chronological order. The appreciation route composed of 16 exhibits is shown in Figure13, and the route composed of 11 exhibits is shown in Figure14. In this experiment, the number of exhibits included in appreciation route is limited experimentally, so it is not based on research on the time visitors stay in museum.

Node num	Category	Age	h(v)
8	Buddhism picture	early Heian	0.025736
9	Buddhism picture	early Heian	0.025736
10	Buddhism picture	early Heian	0.025736
13	fable picture	late Heian	0.043054
14	Buddhism picture	late Heian	0.088042
15	picture scroll	late Heian	0.039101
16	craft	late Heian	0.042355
18	Buddhism picture	late Heian	0.082733
19	Buddhism picture	late Heian	0.090257
20	picture scroll	late Heian	0.083311
21	Buddhism picture	Kamakura	0.061706
22	Buddhism picture	Kamakura	0.054349
23	Buddhism picture	Kamakura	0.054349
24	Buddhism picture	Kamakura	0.068991
25	Buddhism picture	Kamakura	0.068991
26	ink painting	Kamakura	0.026355

Fig. 13. Composed of 16 exhibits

Node num	Category	Age	h(v)
13	fable picture	late Heian	0.043054
14	Buddhism picture	late Heian	0.088042
16	craft	late Heian	0.042355
18	Buddhism picture	late Heian	0.082733
19	Buddhism picture	late Heian	0.090257
20	picture scroll	late Heian	0.083311
21	Buddhism picture	Kamakura	0.061706
22	Buddhism picture	Kamakura	0.054349
23	Buddhism picture	Kamakura	0.054349
24	Buddhism picture	Kamakura	0.068991
25	Buddhism picture	Kamakura	0.068991

Fig. 14. Composed of 11 exhibits

7.2 Discussion

This section discuss our technique with result of the experiment. Figure15 shows a part of the base sub-graph composed in the experiment.

As shown in Figure12, Buddhism pictures with high integration scores are the nodes 19, 14, and 18. Those pictures are influenced by concepts with high importance scores, and they are the nodes 36 and 35. The node 36 expresses the change of culture from Chinese style to Japanese style in late of the *Heian* era. The node 35 expresses the appearance of new kind of Buddhism in late of the *Heian* era. Japanese Buddhism pictures were influenced by these concepts in drawing technique[7].

In this experiment, Buddhism pictures naturally had high integration scores because they were set as root set. Therefore, we mention exhibits having high integration scores though they are in other fields. The node 20 is a picture scroll which is not drawn as Buddhism picture for religion, but it was drawn on the theme of Buddhism. For that reason, it was influenced by the nodes 36 and 35, and they came to have high integration scores. Moreover, the nodes 13, 16, and

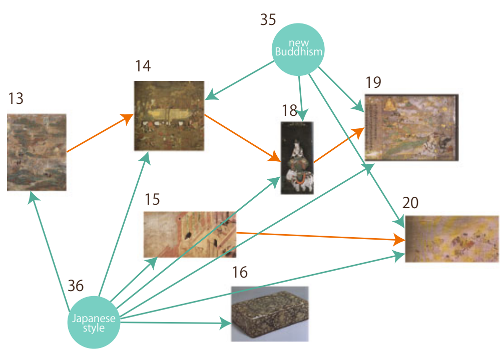


Fig. 15. A part of the base sub-graph

15 are also influenced by the node 36, and have changed in drawing technique as well as Buddhism pictures. Therefore, it can be thought that by considering cultural influence between exhibits, not only the exhibits in the field that the user is interested in but also the exhibits having close relationships with them though the user is not much interested in it can be recommended.

Secondly, we mention the selection of exhibits for composition of appreciation route using integration scores. In this experiment, when the exhibits with high integration scores are selected simply, the exhibits included in the route concentrate in the exhibits produced in late of the *Heian* era to the *Kamakura* era. This is because Buddhism pictures were actively produced in those era. Since they evolved in various ways, and a lot of them exist today, they have high importance scores and integration scores. However, it can be thought that some user want to appreciate the whole exhibition which is not limited to the exhibits produced in particular age. As the solution of such requirement, improvement in composition of base sub-graph, and recommendation using importance scores and integration scores may connect to the composition of appreciation route matched to the individual user well. Moreover, by comparing of Figure13 with Figure14, when the user's scheduled staying time is short, it may be composed mainly by the representative exhibits as described in Section 7.

Thirdly, we mention the order of presentation of the exhibits. In Section 6, we mentioned that it is natural to appreciate exhibits in chronological order. However, tracing links between nodes may compose better appreciation routes than by using it. In Figure 15, when we compose appreciation route by chronological order, it is composed from the nodes 14 to 15. However, since there is a direct relationship between the nodes 14 and 18, it is better to trace this link to compose a meaningful route. Thus, we plan to consider the composition of appreciation route based on strength and semantics of relationships between exhibits.

Finally, we mention the problem of the spatial layout of the exhibits in a museum. Appreciation routes are composed using semantic relationships between exhibits, so our approach may not work well depending on the spatial layout of the exhibits in a museum. It is necessary to lose the distance between the route composed using semantic relationships between exhibits and the route which the user actually traces, and it will be solved as the problem of its interface. Consideration of the restriction for the spatial layout of the exhibits in a museum is one of important future works.

8 Conclusion

In this paper, we proposed an approach for composing a personalized appreciation route matched to user's interest with the information device. By considering cultural influence between exhibits, our system can recommend not only the exhibits in the field that the user is interested in, but also the exhibits having close relationships with them though the user have been not much interested. As a result, it can be thought that our system makes user understand the culture surrounding individual exhibit, and arouses user's interest.

As future works, recommendation using the importance and integration score for the composition of appreciation route matched to the individual user well, improvement of the order of presentation of exhibits, evaluation of exhibition by user, and application of browsing history of user are considered. In addition, development of the system which automatically takes relationships between exhibits from explanation sentence and inspection of the effect by subjective experiments with prototype system are necessary.

References

1. National Museum of Nature and Science, Tokyo, http://www.kahaku.go.jp/
2. Masuoka, A., Fukaya, T., Takahashi, T., Takahashi, M., Ito, S.: ubiNEXT: A New Approach to Support Visitor's Learning Experience in Museums. In: CHCI International (2005)
3. Koyanagi, F., Kon, T., Higashiyama, A.: The Recommended Path Indication System in Hakone Open-Air Art Museum with Time Designation. The Journal of the Faculty of Science and Technology Seikei University 43(2)
4. Kleinberg, J.: Authoritative sources in a hyperlinked environment. In: Proc. 9th Ann. ACM-SIAM Symp. on Discrete Algorithms, pp. 668–677. ACM Press, New York (1998)
5. Baldi, P., Frasconi, P., Smyth, P., Masahiro, Mizuta, Minami, H., Komiya, Y.: Modeling the Internet and the Web: Probabilistic Methods and Algorithms. Morikita Publishing Co., Ltd. (2007)
6. Tsuji, N.: Nihon bijutsushi. Bijutsu Shuppan Holdings Co., Ltd. (2003)
7. Bukkyou kaiga - Wikipedia, http://ja.wikipedia.org/wiki/%E4%BB%8F%E6%95%99%E7%B5%B5%E7%94%BB

Design and Implementation of a Context-Aware Guide Application for Mobile Users Based on Machine Learning

Yuichi Omori, Yuki Nonaka, and Mikio Hasegawa

Tokyo University of Science,
1-14-6, Kudankita, Chiyoda-ku, Tokyo, Japan
{omori,nonaka}@haselab.ee.kagu.tus.ac.jp,
hasegawa@ee.kagu.tus.ac.jp
http://haselab.ee.kagu.tus.ac.jp/

Abstract. This paper proposes a design and implementation of a context-aware application system to guide mobile users about their interesting spots (e.g. restaurants, stores, sightseeing spots) appropriately. A machine learning algorithm enables adaptive recommendation of spots for the mobile users based on their real-time context such as preference, location, weather, time, etc. Our proposed guide system recommends context-aware information for any users by switching two kinds of recommendation algorithms according to the number of user's training data. By experiments using our implemented system in real environments, we confirm that our implemented system correctly works on the off-the-shelf mobile phones having a built-in GPS module and show that it recommends useful information for the mobile users according to their context.

Keywords: Context-Aware, Recommendation, Mobile, Machine Learning.

1 Introduction

Various recommendation systems have been developed and commercialized. Recently, some of them play an important role of serious business tools in E-Commerce [1]. For example, "movielens" developed by GroupLens Research [2], which is a web application to recommend favorite movies for the users by making user's profile., is widely used by Internet users. Various mail order systems, such as amazon.co.jp [3], recommend the commodity information to the users based on the buying history of other users whose profile is similar to the user. However these systems do not use real world context information (e.g. weather, location and etc.), but only use user's preference as the static context information.

In the field of ubiquitous computing, there are various researches on applications which recommend information according to real world information with mobile terminals. A user navigation system: CoCo [4] deals with user's preference and situation as the context information. CoCo infers user's situation (e.g. standing, walking) from various context information collected by sensors, which

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 271–279, 2010.

the user wears, and informs suitable information to the mobile users. However, in this system, the users have to input their preference data manually to get satisfactory recommendation. Blue Mall proposed in Ref. [5] is a context-aware application to notify the mobile users about advertisement of nearby shopping stores based on the location of the user estimated by the Bluetooth RSSI. However, this system does not use user's preference, so its advertisement may contain unnecessary information and may annoy the user. A system called Bookmark Handover [6] is a context aware reminding application. It reminds mobile users about events or places, which they had been registered before based on their context such as location, timing, etc. Although this system sends information based on user's preference and situation, the users have to input all preferable contents manually.

As an approach which may solve issues described above, learning algorithms have been introduced into context-aware applications to make automatic and adaptive decisions according to the context information. A system called Synapse [7] is a context-aware service platform to provide appropriate services by decisions of service provision based on machine learning in home network environment. This system learns relations between the services provided by home appliances and the user's contexts (e.g. preference, life style) that are collected by the sensor networks, and automatically provides appropriate services according to the user's context information.

Machine learning algorithms have been applied also for automatic context-aware recommendation of appropriate information according to the user's context and preference. The context-aware application to recommend the movies based on the Bayesian networks [8] learns user's preference model by regarding the user's context and the features of the movie as stochastic variables, which are the nodes in the network. This system can recommend adequate information for the users by optimizing the network model's parameters based on the feedbacks from the users. The recommendation system proposed in Ref. [9], introduces three kinds of contexts: user's preference, situation, and network environment. It uses a Bayes classification machine and constructs user's preference model to decide useful contents according to the contexts. These applications [8], [9] can recommend context-aware information for the users by constructing their preference models based on machine learning algorithms. However, these recommendation systems based on machine learning algorithms need sufficient amount of training data for each user and each context to make user's preference model. The problem becomes so serious for the users who use the system for the first time or in a short period. Particularly in recommendation systems for a local area, the problem occurs frequently.

In this paper, we propose context-aware guide application for a sight-seeing spot. Our proposed guide system solves such an issue and enables context-aware recommendation for any users by switching two kinds of recommendation algorithms according to the number of user's training data. In this paper, we show the design and the implementation of the context-aware guide system called Kagurazaka Explorer, which is available for the off-the-shelf mobile phones. This system guides mobile users about Kagurazaka street in Tokyo, Japan. We evaluate effectiveness of this proposed system by testing it in real environment.

The rest of this paper is organized as follows. In section 2, we describe about our proposed guide system. In section 3, we describe the detailed design of our proposed guide system. In section 4, we show the implementation on the server and evaluate the system in section 5. We conclude the paper on section 6.

2 A Context-Aware Guide System Based on a Machine Learning Algorithm

We propose a context-aware application to provide specific regional information to the mobile users. This application appropriately guides the users according to their contexts (e.g. user's preference and situation). The most appropriate information which should be provided to the users is decided by a learning algorithm with the feedbacks from the user. The proposed system is supposed to be used to provide specific information of some specialized area (e.g. commercial avenues, sightseeing spots) to the mobile users. In such a system, there may be two kinds of users. The first users are people who use this system for long term (e.g. residents in the area, repeat customer, and so on). The second users are people who use this system for a first time or in short term (e.g. tourists). In this system, the learning algorithm learns the user's preference model and the customer model based on training data which can be collected by the user's feedbacks. By using these models, this system automatically selects appropriate information for the target user. For the first users, the proposed system learns the user's own preference based on his/her training data. This method is almost the same as the method proposed in Ref. [10], and we call this Context-Dependent Preference Modeling (CDPM). For the second users, the proposed system learns customer models based on other users' training data in the same class. We call this method Context-Dependent Customer Modeling (CDCM). The proposed

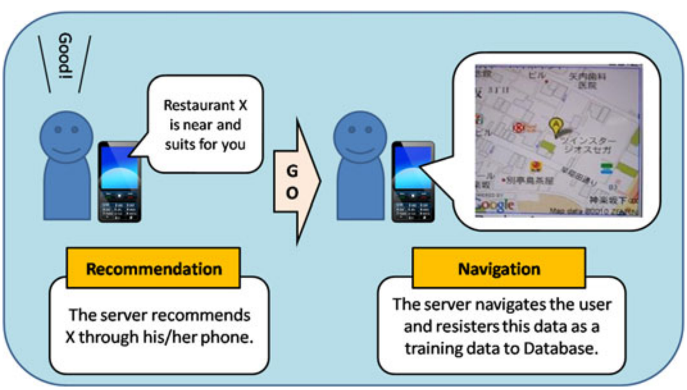


Fig. 1. Concept of our proposed guide system

system can recommend spots according to user context information for any users by switching these two kinds of recommendation algorithms according to the number of user's training data.

Fig.1 shows the overall concept of our proposed guide system. There are two phases: recommendation phase and navigation phase. In the recommendation phase, this system recommends information for the mobile users according to their context (e.g. location, weather). In the navigation phase, the system sends the detailed map of the spot selected by the user to navigate them. At this time, this system regards this spot as the user's favorite in this context and registers this data to the database as a new training data for machine learning algorithms automatically. Thus this system collects training data and constructs a model based on such data to recommend context-aware information.

3 Design of a Context-Aware Guide System

There are at least, two possible design methods to implement the learning algorithm into the system. The first one is to put all functional components on the mobile terminal, and the second is to put functions of learning and content delivery on the server in network side. The advantage of the former is reducing the lower loads on the server side. The latter's advantages are the large amount of training data collected from many terminals and lower loads on the mobile terminal. Our proposed system needs all users' training data for the CDCM. Therefore, we focus on the second method, i.e. keeping all training data and learning function in the network side so that a large amount of training data can be used.

3.1 Recommendation Algorithm

This system recommends appropriate spots for the users according to their context by switching two kinds of methods (i.e. CDPM, CDCM) according to the number of user's training data. For both methods, we use the separation technique to classify spots or users into two classes, appropriate ones to make recommendation and others, according to the context data. As the classification algorithm, we employ Support Vector Machine (SVM), which is shown effective for context-aware recommendation in Ref. [10]. The SVM improves the generalization performance by maximizing the margin between training data and the separating hyper plane.

In case of the CDPM, the features of the spot (e.g. meat, sweets, history) and the user's situations (e.g. accompany, weather) are used for the feature parameters (Fig.2 (1)). This means that this method models the user's preference based on the user's own training data and classifies the contents into appropriate ones and others. It recommends the contents which are regarded as appropriate ones for the user.

On the other hand, in the CDCM, the user's attribute (e.g. sex, age) and user's situation are used for the feature parameters (Fig. 2 (2)). This method models customers based on all users' training data which has already been collected

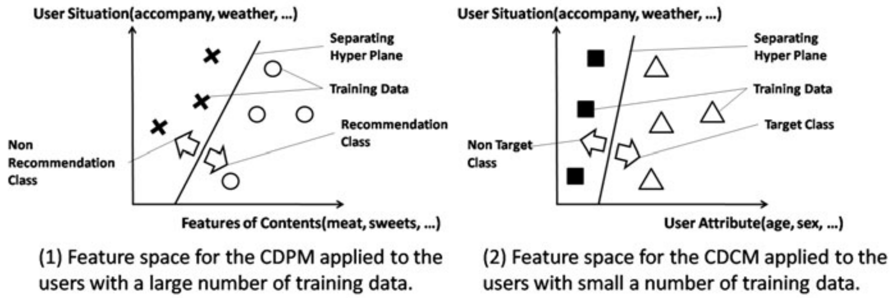


Fig. 2. Feature space for the context-based learning

and classifies the users into target or non target classes by constructed customer models. It recommends the content for the users who are regarded as the target by the customer model.

3.2 Components and Network Environment of the Proposed System

Our proposed guide system consists of 4 main components: the mobile terminals, a context aware decision server (CADS), a database(DB) for storing training data and learning parameters, and a parameter optimizer. All components are connected to each other as shown in Fig. 3. The mobile terminals are used to receive interesting information and to upload user's context data which is collected by the sensors equipped to the mobile terminals. The CADS has two functions. The first one is to decide useful contents based on the learning algorithm and to deliver them to the mobile terminals. The second one is to make new training data based on context data uploaded from the mobile terminals and to register it to the DB. The parameter optimizer optimizes parameters of the learning

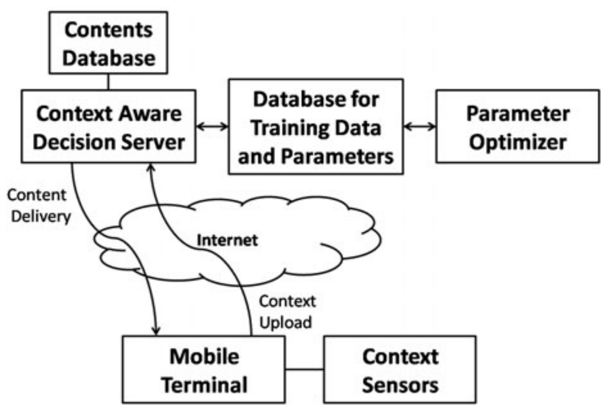


Fig. 3. Proposed guide system

model using the learning algorithm based on the training data, and updates the learning parameters preserved in the DB. By this system, the mobile terminals can get context-aware recommendation with a smart learning algorithm only by uploading context information.

3.3 Implementation of a Context-Aware Guide Application: Kagurazaka Explorer

We have implemented this guide system on the server, which can be used via cellular phone networks, by off-the-shelf mobile phones as the mobile terminals. This section shows the detailed implementation method and adopted techniques. The mobile terminal used in this implementation is the 3G mobile phones which are generally available in Japan provided by main three operators. They are equipped with GPS, and can access to the Internet services by web browsers. The CADS has a global IP address and the mobile terminals can access it by HTTP. We have prepared contents information about restaurants, sightseeing spots, and souvenir stores at Kagurazaka street neighboring to Kagurazaka campus, Tokyo University of Science in Tokyo, Japan, so we call this guide system Kagurazaka Explorer. Table 1 shows elements of feature vectors in Kagurazaka Explorer. The feature vectors for the CDPM are composed by 18 dimensions, and that of the CDCM are composed by 12 dimensions. These elements are used to charcterize the real spots. Values in each dimension are normalized between 0 and 1. As the kernel function for the SVM, we use the RBF kernel defined by

$$K(\mathrm{x}_i, \mathrm{x}_j) = exp\left(-\frac{\| \mathrm{x}_i - \mathrm{x}_j \|^2}{\sigma^2}\right), \tag{1}$$

where the kernel parameter σ is optimized by cross-validation in this system.

Fig. 4 shows the screen shots of the mobile phone connecting to the Kagurazaka Explorer web page. Fig. 4 (1) shows a main screen of Kagurazaka Explorer. From this screen, the user can start search of the recommended spots. After getting the position of the user by built-in GPS module, the user has to input the number of accompanies manually in the screen shown in Fig. 4 (2). Although the system requires manual input here, it leads much improvement on the rate of correct recommendation as shown in Ref. [11]. After sending

Table 1. Composition of Feature Vectors for the CDPM and the CDCM

Features of Contents for CDPM (8 dimensions)	User Situation for CDPM and CDCM (10 dimensions)	User Attribute for CDCM (2 dimensions)
Genre of Restaurant (Japanese, Western, Chinese, Asian, Liquor, Cafe, Noodle) Budget	Season Time Week Position Accompany Weather	Sex Age

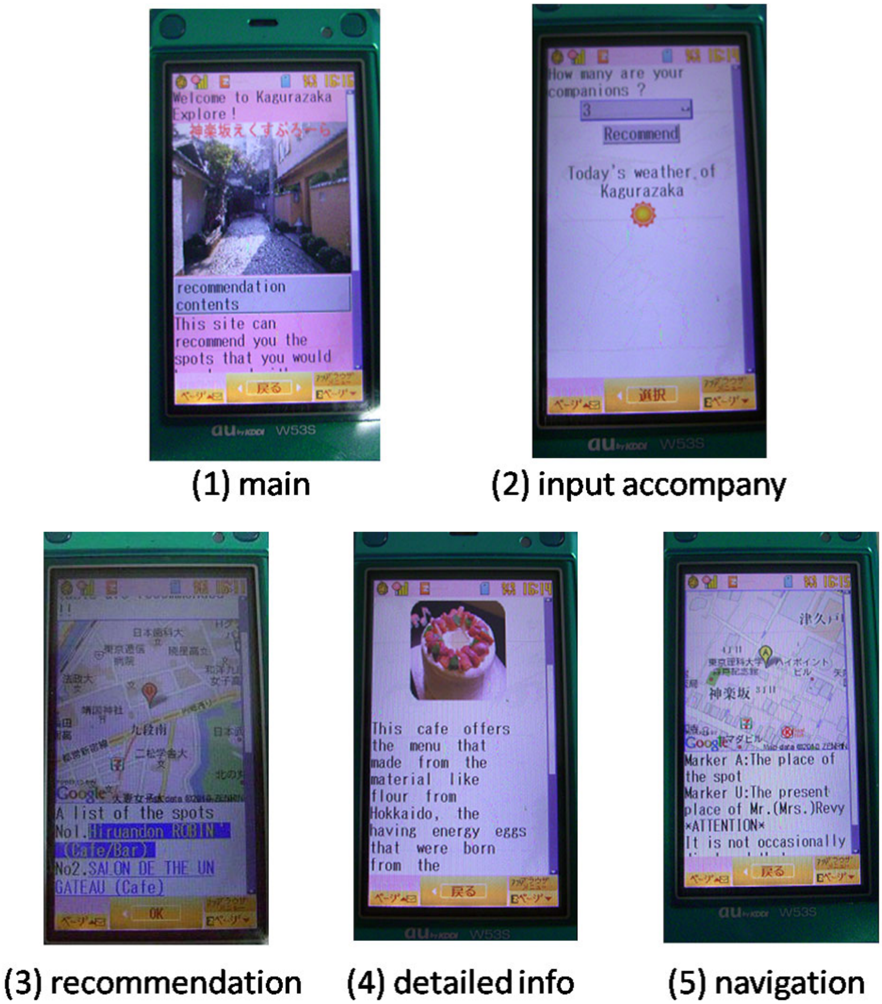
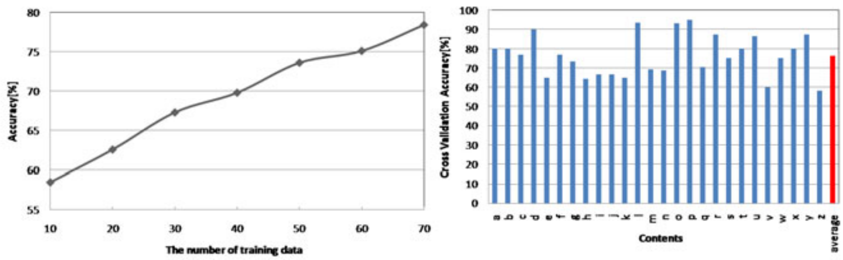


Fig. 4. Screen shots of the implemented system

such context information, he/she can get context-aware recommendation (Fig. 4 (3)). The user can access to the detailed information by selecting one in the list (Fig. 4 (4)). When farther information including the map shown Fig. 4 (5) is required by the user by another one click, this spot is regarded as a favorite one for the user and the training data on the server is updated. Fig. 4 (5) shows the screen of the detailed map to navigate the user to the spot.

4 Evaluation of Kagurazaka Explorer

We evaluate Kagurazaka Explorer with two types of algorithms, the CDPM and the CDCM, respectively. The recommendation method based on the CDPM is



(1) Relation between the number of training data and the rate of correct recommendation on unknown inputs in the CDPM.

(2) Rate of correct recommendation in the CDCM.

Fig. 5. Evaluation of learning methods

evaluated by 10 subjects, who make 90 training data to construct the user's preference model. The rate of correct recommendation on 20 unknown data is evaluated by changing the number of training data. Fig.5 (1) shows the relation between the number of training data, and the average of the correct recommendation rate on unknown inputs. In the models which use 10 training data, the rate of correct recommendation becomes lower than 60%, while the 60 training data models perform almost 75% on average. As the number of training data grows, the rate of correct recommendation marks higher.

The recommendation method with the CDCM is also evaluated. We use the training data provided by 10 subjects, who use this system in the Kagurazaka street. About 1000 training data is generated for each subject by using this system in various situations (for a rainy day, a fine day, at various positions, time and accompany). Fig.5 (2) shows the rate of correct recommendation by 5-fold cross validation in the CDCM. We use the spots with more than 20 training data for the CDCM. The correct recommendation rate becomes 76% on average. In this case, 76% in the CDCM is higher than the rate of the CDPM when the number of training data is less than 60.

According to these results, we confirm that the recommendation using the CDCM should be applied for the users have less than 60 training data, and the recommendation using the CDPM should be applied for the users have more than 60 training data. In this case we should regard 60 as threshold to switch these two methods.

5 Conclusion

This paper has shown the design and implementation of a context-aware application system to guide mobile users. We confirm that our implemented system works on the off-the-shelf mobile phones by experiments and that it recommends appropriate information for the mobile users by switching two kinds of recommendation algorithms according to the number of user's training data.

Other context-aware applications based on the machine learning which are introduced in the section 1 are not evaluated in real environment. We have proposed the context-aware application which works in a specialized local area. We have introduced and evaluated our implemented system in the real shopping street.

Our results indicate that the rate of correct recommendation in our system improves by increasing the number of the training data. It means, however, the number of training data was not enough in the experiment. Therefore, as a futurework, we will experiment our system using the larger number of subjects and improve the accuracy and the easiness of this system with considering real use by the general users.

Acknowledgments

The authors would like to thank to Kagurazaka Street shopping mall association and Prof. Hangai who is the department head of Division of TUS-Kagurazaka Future Network Research, Tokyo University of Science and all people who have cooperated with this research.

References

1. Ben Schafer, J., et al.: Recommender systems in e-commerce Subsequences. In: Proc. of ACM Conference on Electronic Commerce, pp. 158–166 (1999)
2. GroupLens Research, movielens, `http://movielens.umn.edu/login` (April 2010)
3. Amazon.com, amazon.com, `http://www.amazon.com/` (April 2010)
4. Hayashi, T., et al.: Design and Implementation of a Context-Aware Content Delivery Service Using Tiny Mobile Sensors. Technical Report of IEICE 104(691), 149–154 (2005)
5. Sanchez, J., et al.: Blue Mall: A Bluetooth-based Advertisement System for Commercial Areas. In: Proc. ACM Workshop on Performance Monitoring and Measurement of Heterogenious Wireless and Wired Networks, pp. 17–22 (2008)
6. Tran, H., et al.: Design and Implementation of Bookmark Handover - A Context-aware Reminding Application for Mobile Users. In: Proc. of WPMC, pp. 1227–1231 (2006)
7. Si, H., et al.: Design and Implementation of a Context-Aware Service Platform Synapse. In: Proc. of the IEICE General Conference, B-19-3 (2005)
8. Ono, C., et al.: A Context-Aware Movie Preference Model Using a Bayesian Network for Recommendation and Promotion. In: Conati, C., McCoy, K., Paliouras, G. (eds.) UM 2007. LNCS (LNAI), vol. 4511, pp. 247–257. Springer, Heidelberg (2007)
9. Yu, Z., et al.: Supporting Context-Aware Media Recommendations for Smart Phones. Pervasive Computing 5, 68–75 (2006)
10. Oku, K., et al.: Context-Aware Recommendation System Based on Context-Dependent User Preference Modeling. Proc. of Information Processing 48(SIG11) (TOD34), 162–176 (2007)
11. Omori, Y., et al.: Design and Implementation of a Context-Aware Recommendation System Based on Machine Learning for Mobile Users. Technical Report of IEICE 109(382), 117–122 (2010)

Adaptive Traffic Signal Control Based on Vehicle Route Sharing by Wireless Communication

Hiroyasu Ezawa[1] and Naoto Mukai[2]

[1] Dept. of Electrical Engineering, Graduate School of Engineering,
Tokyo University of Science
Kudankita, Chiyoda-ku, Tokyo, 102-0073, Japan
j4309617@ed.kagu.tus.ac.jp

[2] Dept. of Electrical Engineering, Faculty of Engineering, Division 1,
Tokyo University of Science
Kudankita, Chiyoda-ku, Tokyo, 102-0073, Japan
mukai@ee.kagu.tus.ac.jp

Abstract. In these years, a problem of traffic congestion is a growing concern due to increasing vehicle holders in urban areas. One of key issues to ease the traffic congestion is a traffic signal control. In this paper, we propose a control system for traffic signals by a vehicle route sharing. The vehicle route sharing is to share position and path information (i.e., probe data) among vehicles. The shared information is used for calculating a criteria of traffic congestion called expected traffic congestion. The parameters of traffic signals (i.e., cycle, split, and offset) are optimized on the basis of the expected traffic congestion. We performed a multi-agent simulation to evaluate the effectiveness of our proposed system. The results show that our proposed method can ease traffic congestion effectively compared with other traditional methods.

1 Introduction

Intelligent Transportation System (ITS) is a system to solve various traffic issues by using computers, electronics, and advanced sensing technologies. One of the important themes of the ITS is to ease traffic congestion. The problem of the traffic congestion is becoming more serious due to increasing vehicle holders in urban areas. A traffic signal control is one of the approaches to solve the traffic congestion problem. A traditional control method for traffic signals is a pattern selection, i.e., a suitable signal pattern which is selected from some preformed patterns. However, such method is inadaptable to a dynamic change of traffic situations. Thus, adaptive methods for the traffic signal control and traffic simulators are reported by numerous researchers [1,2,3,4,5].

In these years, control methods using probe data (i.e., real-time position and speed of vehicles) for traffic signals are getting a lot of attention. For example, in [6], a delay time at a red light is estimated by using probe data instead of vehicle detectors, and traffic signal parameters (i.e., cycle, split, and offset) are optimized to minimize the estimated delay time. Moreover, in [7,8,9], an on-demand traffic

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 280–289, 2010.

control scheme called ADS(Advanced Demand Signals scheme) and its extended version are proposed. In the schemes, traffic signals are controlled by real-time vehicle (or pedestrian) positions without traditional signal parameters, and the indication of traffic signals depends on the evaluation of demand (i.e., request for keep indication or change indication).

In this paper, we propose a new control system for traffic signals by vehicle route sharing (i.e., vehicles share their traveling routes by wireless communication). Vehicles sends their scheduled routes to traffic control agents at fixed intervals, and the shared routes is used for calculating Expected Traffic Congestion (ETC for short) which is proposed by Yamashita et al. [10]. The ETC is a criteria of congestion degree, and the value is high if the overlap paths in the shared routes are large. The traffic control agents optimize the parameters of traffic signals (i.e., cycle, split, and offset) by the ETC. We propose two types of controllers for the optimization: cycle and split controller (CS-controller), and offset controller (O-controller). At the end of the paper, we will show that our method outperforms traditional methods by simulation experiments based on a multi-agent model.

The remainder of this paper is organized as follows. Section 2 defines the expected traffic congestion. Section 3 describes the simulation model of traffic flows. Section 4 presents our traffic control method based on the expected traffic congestion. Section 5 shows experimental results and discusses it. Finally, Section 6 offers conclusions and future works.

2 Definition of Expected Traffic Congestion

In this section, we show a congestion degree called Expected Traffic Congestion (ETC for short) which is proposed by Yamashita [10]. The ETC represents the amount of link traffics according to the shared routes (a path from present position to destination) among vehicles. In [10], the ETC was used to find shorter paths by avoiding traffic congestion. On the other hand, in this paper, the ETC is used to optimize the control parameters of traffic signals (i.e., cycle, split, and offset).

2.1 Representation of Path

In order to calculate the ETC, all vehicles send their paths periodically to traffic control agents which optimize the parameter of traffic signals, and the agent calculates ETC from the collected paths. A path of a vehicle v_i is represented by a sequence of links as Equation 1. For simplicity, we assume that each vehicle always selects the path of the shortest distance (not the shortest time).

$$L_i = (l(s_0, s_1), l(s_1, s_2), l(s_2, s_3), l(s_4, s_5)) \tag{1}$$

2.2 Expected Traffic Congestion

Here, we explain how to calculate the ETC according to collected paths as follows. First, a pass weight (PW for short) of link $l \in L_i$ is calculated as Equation

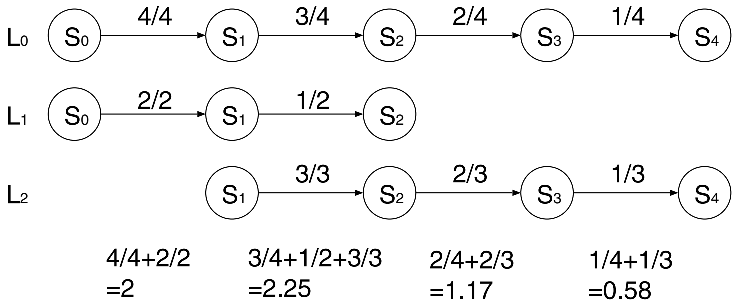


Fig. 1. An example of expected traffic congestion

2 where $|L_i|$ is the path size of L_i, and $num(l)$ is the sequence number of link l in the path of L_i (i.e., $1 \leq num(l) \leq |L_i|$). Moreover, we introduce a new parameter $\alpha(\geq 1)$, and the α is enable to overestimate (or underestimate) the PW according to time-series. Consequently, the PW represents a congestion degree for one vehicle, and the weight of links near present position is estimated high than far links. For example, in the path of Equation 1, a pass weight for $l(s_2, s_3)$ is $2/4$.

$$PW_i(l) = \left(\frac{num(l)}{|L_i|}\right)^{\alpha} \tag{2}$$

Next, the ETC of link l is calculated as Equation 3. The expected traffic congestion represents the sum of the PW for all vehicles. For example, in Figure 1, there are three collected paths (L_0,L_1, and L_2). The expected traffic congestion of links ($l(s_0, s_1)$,$l(s_1, s_2)$,$l(s_2, s_3)$, and $l(s_3, s_4)$) are calculated (subject to $\alpha = 1$) as follows. A link $l(s_0, s_1)$ is contained in the paths L_0 and L_1, thus, the ETC of $l(s_0, s_1)$ is the sum of $PW_0(s_0, s_1) = 4/4$ and $PW_1(s_0, s_1) = 2/2$. Moreover, a link $l(s_1, s_2)$ is contained in the all paths (L_0, L_1 and L_2), thus, the ETC of $l(s_1, s_2)$ is the sum of $PW_0(s_1, s_2) = 3/4$, $PW_1(s_1, s_2) = 1/2$, and $PW_1(s_1, s_2) = 3/3$.

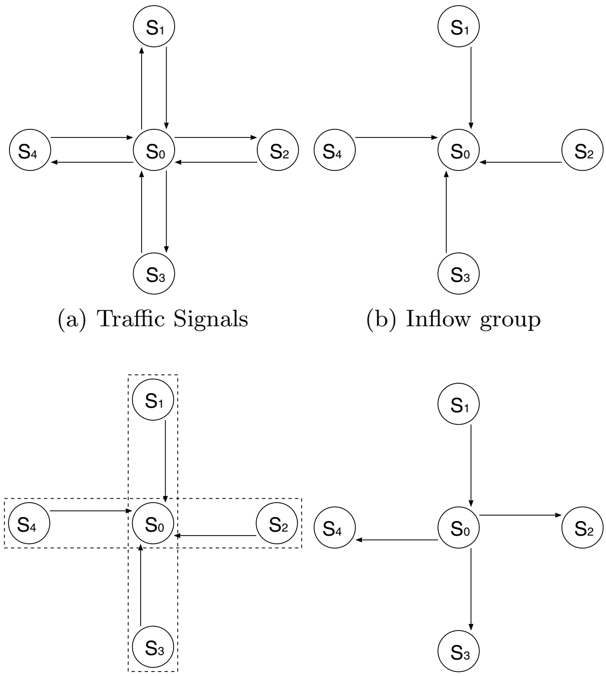
$$ETC(l) = \sum_i PW_i(l) \tag{3}$$

3 Model of Traffic Network

This section describes the model of traffic network which includes a road network and traffic signals for a traffic simulation.

3.1 Formulation of Road Network

Figure 2(a) shows an example of a road network. There are five traffic signals $(s_0, \cdots, s_4)$ and links between traffic signals ($l(s_0, s_1)$ represents a link from s_0



(a) Traffic Signals (b) Inflow group

(c) Synchronous pair (d) Cooperation pair

Fig. 2. A model of traffic network

to s_1). We define three types of link set: inflow group, synchronous pair, and cooperation pair as follows.

Inflow Group. An inflow group is a set of inflow links into a center traffic signal. The all passing vehicles on the inflow links are influenced the center traffic signals. For example, in Figure 2(b), a set of links ($l(s_1, s_0)$, $l(s_2, s_0)$, $l(s_3, s_0)$, and $l(s_4, s_0)$) is an inflow group into a center signal s_0.

Synchronous Pair. A synchronous pair is a pair of adjacent links in the opposite direction on the straight road in the inflow group. The signal indications of the pair are synchronized with each other. For example, in Figure 2(c), a pair of $l(s_4, s_0)$ and $l(s_2, s_0)$ is a synchronous pair.

Cooperation Pair. A cooperation pair is composed of an inflow link and an outflow link of a center traffic signal. The passing vehicles on inflow link are divided into outflow links after passing the center traffic signal. For example, in Figure 2(d), there are three cooperation pairs ($l(s_1, s_0)$ pairs with $l(s_0, s_2)$, $l(s_0, s_3)$, and $l(s_0, s_4)$).

3.2 Formulation of Traffic Signals

Here, we explain control parameters of traffic signals by using an example shown in Figure 3. There are three control parameters: cycle length, split, and offset as following.

Cycle Length. The cycle length is a time interval that an signal indication loops back. In our model, the value of cycle length is set between 90 seconds and 180 seconds (the range is based on a typical cycle length of traffic signals). For simplicity, in our model, a traffic signal has only red and green indications (i.e., yellow indication is regarded as a part of green indication). In Figure 3, the cycle lengths for s_0 and s_1 are 120 seconds.

Split. The split is a ratio of the cycle length which represents a time interval of each signal indication (red or green indication). In our model, the range of the ratio is set between 30% to 70%. In Figure 3, the splits for s_0 and s_1 are 50% to 50%.

Offset. The offset is a time difference for the same indication between two traffic signals of a cooperation pair. In Figure 3, the offset of s_2 is 40 seconds on the basis of s_1. Let consider the case of the network shown in Figure 2. We assume that a vehicle starts from s_0 to s_1 immediately after s_0 changes green indication. If an offset corresponds to transit time from s_0 to s_1, the vehicle does not need to stop at the s_1 because s_1 timely changes green indication. Thus, a suitable offset between traffic signals can reduce the waiting time at a signal.

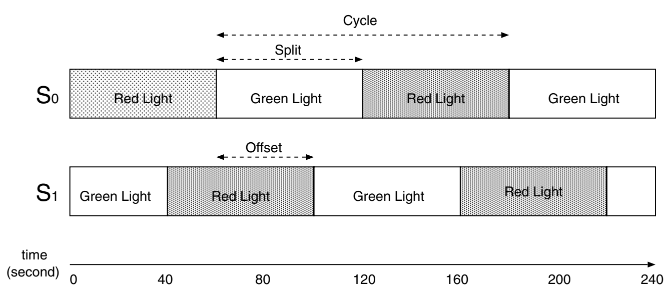


Fig. 3. Control parameters of traffic signal

4 Traffic Signal Control Based on Expected Traffic Congestion

Each traffic control agent has two traffic controllers: Cycle-Split controller (CS-controller) and Offset controller (O-controller). The controllers optimize the control parameters of traffic signals (cycle length, split, and offset) on the basis of the expected traffic congestion at fixed intervals.

4.1 CS-Controller

The CS-controller optimizes cycle length and split of traffic signals at the end of cycle. First, the CS-controller calculates the average expected traffic congestion of links in a inflow group IG as Equation 4. The value represents the amount of traffic of a center signal in the inflow group. Thus, the CS-controller increases the cycle length by a value β if the value is more than limit θ_1. On the other hand, the CS-controller decreases the cycle length by β if the value is less than the limit. In the case of heavy traffic, long cycle enables to reduce the amount of acceleration-deceleration time in front of the signal.

$$AETC(IG) = \frac{\sum_{l \in IG} ETC(l)}{|IG|} \tag{4}$$

After the cycle control, the CS-controller calculates the total expected traffic congestion of links of a synchronous pair SP as Equation 5. The values of synchronous pairs in a inflow groups are compared with each other. The difference of the values represents the bias of traffic amounts. Thus, the CS-controller shifts the split ratio by γ from small synchronous pair to large synchronous pair if the difference is higher than limit θ_2.

$$TETC(SP) = \sum_{l \in SP} ETC(l) \tag{5}$$

4.2 O-Controller

First, O-controller filters cooperation pairs, and selects some cooperation pairs whose expected traffic congestion is more than limit θ_3. Second, the O-controller calculates the ratio of expected traffic congestion of a inflow link to outflow links in the cooperation pair. If the ratio is higher than threshold θ_4, O-controller activates the offset cooperation (i.e., setting offset value) of the signals of the inflow and outflow links.

For example, there are two traffic signals s_1 and s_2. The offset value of s_2 is set by Equation 6, where $d(s_1, s_2)$ is the distance between s_1 and s_2, and $v(s_1, s_2)$ is the average speed of vehicles between s_1 and s_2. After the offset cooperation, the cycle length and split of s_2 are synchronized with s_1 to keep the offset cooperation. This way enables to reduce the amount of stop at the s_2 for vehicles which start from s_1.

$$OFFSET(s_2) = \frac{d(s_1, s_2)}{v(s_1, s_2)} \tag{6}$$

5 Simulation Experiments

We developed a traffic simulator shown in Figure 4 by Artisoc which is a multi-agent simulator, and compared our traffic signal controller with other traditional controllers on the simulator. Here, we report our experimental results.

5.1 Traffic Simulator

Figure 5 illustrates the road network for our experiments. The road network consists of 62 intersections, and the size of the network is 5×5 km (80×80 pixel). Vehicles come from the outside of the network, and move to their destinations. There are two types of destination distribution. The one is a uniform distribution: all intersections are uniformly selected from the network as destinations. The other is a biased distribution: limited intersections are selected from the network as destinations. The inflow rate of vehicles is set to 50%,33%,25%, and 20% (i.e., a vehicle occurs once per 2 simulation steps in 50%). The speed of vehicles depends on the distance between vehicles, and is classified into four stages (i.e., four types of speed) as shown in Table 1. In order to avoid a crash with a vehicle, the stage is down if the distance between vehicles is less than the current speed. In contrast, if there is no obstacles in front of the vehicle, the stage is up. A default parameter setting for our simulator is summarized in Table 2. We enumerate control methods for comparison in our experiments as follows.

Fixed Controller (FC). The cycle length and split are fixed for all of traffic signals in advance. In this experiment, the cycle length is set to 120 seconds per cycle, and the split ratio is set to 1 : 1.

CS-controller using Availability (CS-A). The cycle length and split are calculated based on green time availability (GTA), which is the rate of the number of passing vehicles for the capacity per cycle.

CS-controller using Queue (CS-Q). The cycle length and split are calculated based on the number of vehicles in the waiting queue in front of traffic signals.

CS-controller using ETC (CS-ETC). The cycle length and split are calculated based on ETC.

CS-controller and O-controller using ETC (CSO-ETC). All control parameters (i.e., cycle length, split, and offset) are calculated based on ETC. O-controller is given priority over CS-controller if the O-controller is activated.

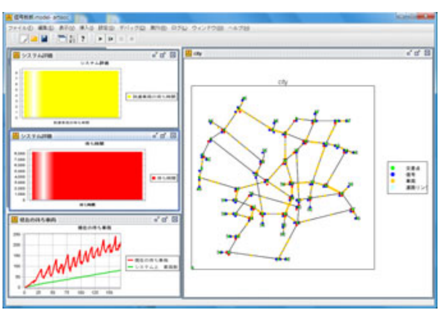


Fig. 4. A screenshot of traffic simulator

Fig. 5. A road network at Chiyoda-ku, Japan

Table 1. Speed stage of vehicles

Stage	Speed(km/h)	Speed(pixel/step)
0	0	0.0
1	11	0.2
2	27	0.5
3	54	1.0

Table 2. Parameter setting for simulator

Parameter	Value
θ_1	cycle length
θ_2	10
θ_3	cycle length or 40
θ_4	70%
α	1
β	5
γ	5%

5.2 Experimental Results

Figure 6 shows the comparison of traffic controllers on the destination distribution (The inflow rate is set to 33%). In the figure, the left bar is the uniform distribution, and the right bar is the biased distribution. The average delay time represents the waiting time of vehicles in front of traffic signals[1]. The result indicates that CS-A and CS-Q can reduce the delay time compared to FC, and CS-A is superior to CS-Q. The advantage of CS-A over CS-Q is that the cycle length and split are updated based on the previous tendency of passing vehicles, but CS-Q consider present situation of vehicles in the waiting queue (i.e., it takes a short view). Moreover, we found that our proposed controllers, CS-ETC and CSO-ETC, outperform other controllers, and CSO-ETC is more effective in the case of biased distribution. In biased distribution, there is a great frequency of traffic congestion, thus, O-controller is capable to provide relief to the traffic congestion in restricted area in the traffic network.

Figure 7 shows the comparison of traffic controllers on the inflow rate (the destination distribution is biased distribution). It is obvious that the waiting time is increased in response to the inflow rate. We can see that the controllers with

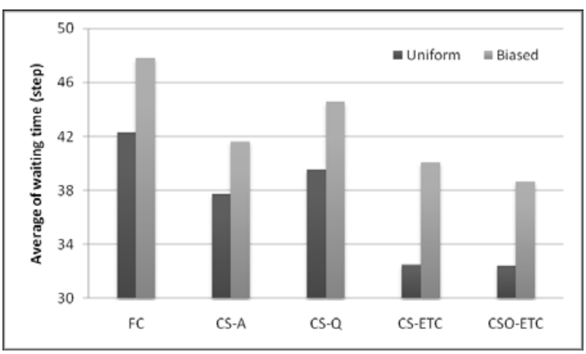


Fig. 6. Comparison of traffic controllers on the destination distribution (uniform and biased distributions)

[1] The unit of the average delay time is one step (45 seconds).

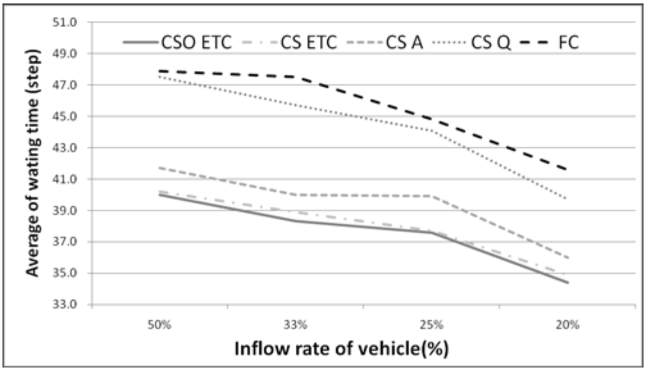


Fig. 7. Comparison of traffic controllers on the inflow rate (50%,33%,25%, and 20%)

ETC (CSO-ETC and CS-ETC) are superior to other controllers in all inflow rate patterns. Moreover, CSO-ETC slightly improves the waiting time of CS-ETC. O-controller may have a bad influence on outside traffic flows from the center signal because the cycle length and offset are forcibly synchronized. However, the bad influence can be avoid by CS-controller to some degree.

6 Conclusion

In these years, a traffic congestion in urban city is becoming a serious problem. Therefore, in this paper, we focused on traffic signal control based on expected traffic congestion, which is a congestion barometer, in dynamic and decentralized manner. The ETC can be calculated from the shared routes of vehicles by wireless communication. Moreover, we proposed cycle-split and offset controllers which optimize control parameters for traffic signals on the basis of the ETC. We performed simulation experiments which compared five signal controllers (FC, CS-A, CS-Q, CS-ETC, and CSO-ETC). The experimental results show that our methods (CS-ETC and CSO-ETC) outperform other traditional strategies.

Acknowledgment

We gratefully acknowledge the advice and criticism of Prof. Naoyuki Uchida of Tokyo University of Science, Japan.

References

1. Miyanishi, Y., Miyamoto, E., Maekawa, S.: A proposal of traffic signal control using its technology. IPSJ SIG Technical Reports 2001(83), 53–60 (2001)
2. Sun, X., Kusakabe, T.: A study on traffic signal control for prevention of traffic congestion. IPSJ SIG Technical Reports 2005(89), 31–34 (2005)

3. Dresner, K., Stone, P.: Multiagent traffic management: A reservation-based intersection control mechanism. In: The Third International Joint Conference on Autonomous Agents and Multiagent Systems, pp. 530–537 (2004)
4. Chen, R.-S., Chen, D.-K., Lin, S.-Y.: Actam: Cooperative multi-agent system architecture for urban traffic signal control. IEICE Transactions on Information and Systems 88(1), 119–126 (2005)
5. Jyunichi, K., Yasuo, K., Souichi, T., Yoshitaka, K., Taisuke, S.: Applying genetic programming with substructure discovery to a traffic signal control problem. Transactions of the Japanese Society for Artificial Intelligence 22(2), 127–139 (2007)
6. Hanabusa, H., Iijima, M., Horiguchi, R.: Development of delay estimation method using real time probe data for adaptive signal control algorithm. In: Proceedings of ITS Symposium 2009, pp. 207–212 (2009)
7. Kato, Y., Hasegawa, T.: Traffic signals controlled by vehicles: Paradigm shift of traffic signals. IEICE Technical Reports. ITS 100(284), 67–71 (2000)
8. Aso, T., Hasegawa, T.: Full automated advanvced demand signals ii scheme. In: Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems, pp. 522–527 (2009)
9. Aso, T., Hasegawa, T.: Traffic signal control schemes for the ubiquitous sensor network period. In: Proceedings of ITS Symposium 2009, pp. 67–72 (2009)
10. Yamashita, T., Kurumatani, K., Nakashita, H.: Approach to smooth traffic flow by a cooperative car navigation system. Transactions of Information Processing Society of Japan 49(1), 177–188 (2008)

A System to Share Arrangements for Daily Tasks and Life Events on the Web

Hitomi Sato[1], Akira Hattori[2], and Haruo Hayami[1,2]

[1] Graduate School of Kanagawa Institute of Technology Course of Information and Computer Science, Japan
[2] Kanagawa Institute of Technology Department of Information Media, Japan

Abstract. Abstract. It is important to make plans and preparations for daily tasks and life events, for example when moving house. A well-considered arrangement makes it possible to efficiently finish such tasks and events. It would also undoubtedly be helpful for many people. Therefore, we propose a web-based system for sharing arrangements for these tasks and events in order to increase productivity in advance of carrying them out. Each user of our system can customize shared arrangements and use them together with a scheduling functionality. The arrangements are linked to events in the scheduler, and users can check their progress as a workflow. This function is useful to make arrangements taking users' different schedules into consideration. We discuss advantage of our system.

Keywords: arrangement sharing, knowledge management, workflow.

1 Introduction

Various methods are used to complete daily tasks and life events, such as those required when moving to a new home. In general, many people make arrangements, i.e., plans and preparations, to carry out such tasks. A well-considered arrangement makes it possible to efficiently complete these tasks and events. This is helpful for the people involved.

We discuss here some arrangements for daily tasks and life events. These tasks contain some actions, which are classified into two categories. The first category is for tasks that need to be conducted in a sequence that is optimum for the arrangement, and the second one is for tasks that can be done in any desired order. We use a situation involving moving house to explain them. In Fig. 1, a sequence of actions is shown: "Leaving a contact address," "Taking the trash out," "Disconnecting the electricity," "Disconnecting the gas," "Disconnecting the water," "Tidying up and leaving," and "Returning the keys." Some of them can be done in any desired order and even in parallel with other tasks. For example, someone who is planning a move can contact a public utility company to disconnect the gas during a break at his/her office.

We also discuss how to make arrangements. In general people generally make arrangements using their own experience as a guide, but they can also do that based on arrangements made by others. Arrangements made by others are also helpful for

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 290–297, 2010.

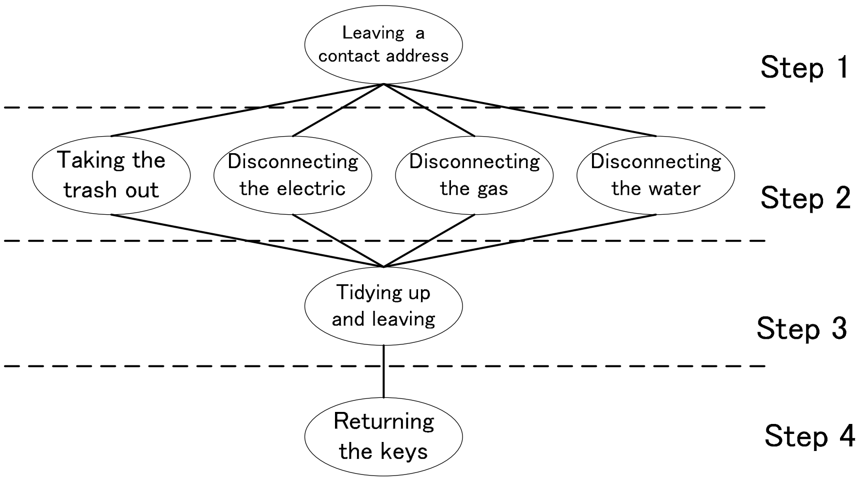


Fig. 1. Arrangement for moving house

someone to use to brush up his/her own arrangements. However, the details of others' arrangements are most likely specific to that person. That is why user would have to select and customize an arrangement made by someone else that matches his/her own characteristics.

Arrangements made by others are divided into three categories. First, we can ask our friends and colleagues directly about their arrangements. Second, we can observe their actions in order to learn from them. Third, we can obtain useful information for our arrangements from sources such as books, magazines, and blogs. There are many media we can use as resources, but to do that, we have to check various kinds of information.

Therefore, the purpose of this study is to make it possible to share arrangements with others. Our system selects an arrangement from others' arrangements matching the user's characteristics. The arrangements are evaluated by users who have used them. Another feature of our system is that it provides users with a scheduling functionality. This function is useful for making arrangements taking the user's schedule into consideration. We developed a prototype system.

2 Related Work

Workflow management systems define, create, execute and monitor arrangements, and support users' planning. However, they have no function to share arrangements with other groups.

It is important to increase members' motivation and the quality of their management know-how, when sharing knowledge in an office organization [1]. Most knowledge-sharing systems for offices are based on a top-down approach, in which managers at the offices provide motivation and know-how to their staff. In contrast, our proposed system is based on a bottom-up approach. This approach gives direct benefits to members, and they can improve arrangements in cooperation with each other.

RESTER2 [2] shares similarities with our proposed system. It is a ToDo management system with an ontology-based model. We will compare our system to RESTER2 in Section 4.3.

Remember the Milk [3] is a web-based ToDo management system. It has no function to allow a user to obtain others' knowledge.

Google Calendar [4] is a web-based scheduler. This system's benefits are that friends can share their calendars with each other. However, it cannot be shared with another person's calendar.

3 Our Proposed System

3.1 Outline of the System

Figure 2 illustrates an outline of our proposed system. The system has two databases to manage arrangements: a public database and a private database.

The system manages arrangements using the concept of workflow, which contains work-items to do. It also provides each user with "my page." The page has a private database and a schedule database. Users can make their arrangements from scratch or by customizing a local copy of others' arrangements in the public database. They can also publish the arrangements that they made on the public database.

To use arrangements in a private database, "my page" provides users with a scheduler function. When a user adds an event to his/her scheduler in our system, it attaches

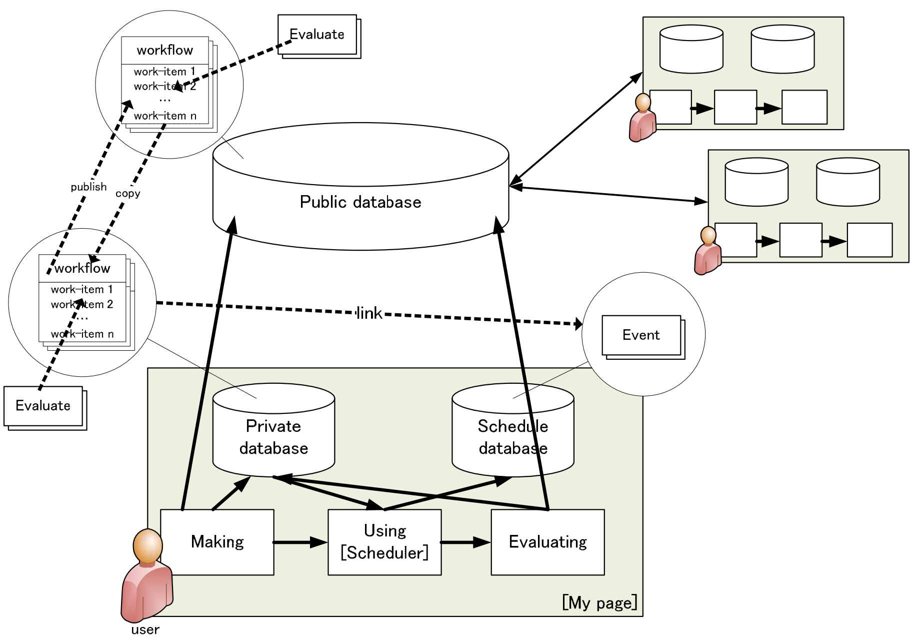


Fig. 2. Outline of our system

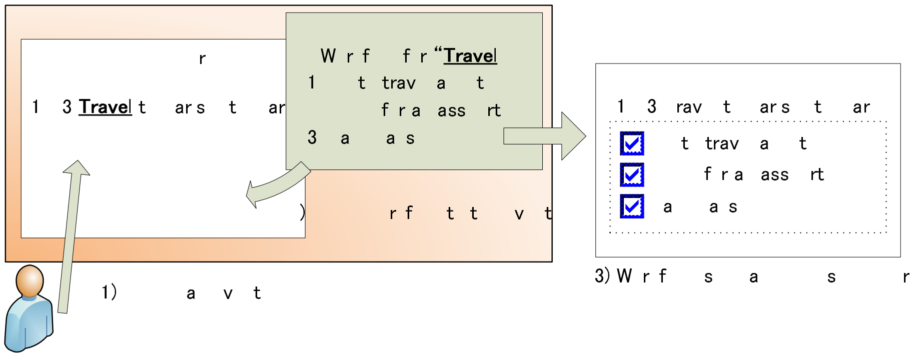


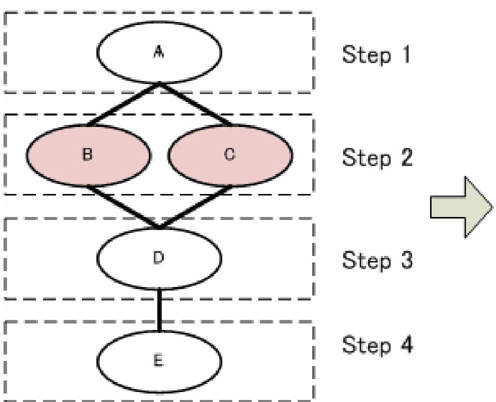
Fig. 3. Adding arrangement in scheduler

a workflow that is appropriate for the event. The user can then check the progress of the workflow using a ToDo list produced by the scheduler according to the corresponding arrangement, as shown in Figure 3. Users can evaluate the arrangements that they have used on a 5-point scale and add some comments.

3.2 Databases for Arrangement Management

To manage arrangements, we use the concepts for "workflow" and "work-item." For example, the workflow corresponding to the arrangement for one's move includes work-items such as "Leaving a contact address" and "Taking the trash out." In our system, a workflow has a set of attributes, which are information items such as who can access, what type, and keywords.

Some work-items should be conducted in the proper sequence, whereas others can be done in any desired order and even in parallel with other work-items. To cope with this feature, our system assigns a number, which we called StepNum, to each work-item. Work-items having the same number can be conducted in any order. For



Num	StepNum	Concrete work
1	1	A
2	2	B
3	2	C
4	3	D
5	4	E

Fig. 4. Data expression of multitask

example, in Figure 4, work-items B and C have the same StepNum, so they can be conducted in any order.

3.3 How to Make a Workflow

In our system, users can input work-items into a form, which is a text area, and handle their order by dragging and dropping the items. However, in general, people often make arrangements based on others' good practice, so our system provides a function to copy public workflows, which are in public databases. With this function, users can make their local copies of published arrangements on their own pages. In addition, users can customize the copied workflow so that it fits them on the page. Many workflow management systems have some standard workflows that can be used commonly. When we try to use the arrangements of daily tasks and life events, standard workflow is better able to customize for each user's environment and conditions.

When the original workflow is updated by the person who made it, the information is also sent to users who have a copy of it. Thus, they can always confirm the latest condition of the original workflow. In general, we are not finding arrangements at random. We always find people like us lifestyle. So this function is useful. We proposed system is keeping the original workflow's ID when we copy others' workflow. Our system is compares original workflow's modified date and copied workflow's creation date. If original workflow has change, our system notifies the user.

Sometime, lifehack and arrangements which media has picked are far from general situation, so they have not played a enough role in making arrangements for daily tasks and life events. Our system provides two functions for finding the arrangements that suit you. One is to allows a user to search for similar users based on their profiles. The profiles include their ages, sexes, jobs and so on, and are set by the users. You can fuzzy search at "similar age" and "metropolitan area". Because of this function, users can find arrangements easier. The other is to allow users to attach tags to workflows and to search for them based on the tags. For example, workflows for an international travel and a domestic travel could have different tags. The tags are granted from workflows' author. Users can select tag list like tag cloud. It is easier tagging and improves search accuracy by organize similar tags.

If users want to share their workflow with others, they can change the corresponding attribute of the workflow to "public." The system also allows users to share their workflow with only the groups they designate.

When we obtain information from TV, books, magazines, and blogs, and use it to make arrangements, we have to look at a lot of information from various media and select the information that is likely to be useful. However, doing that is difficult because it is necessary to choose the information in accordance with each situation. However, our system can locate the workflow of one's choice with the profile recommendation function.

3.4 How to Use Arrangements

Our system is designed to let users carry out arrangements they have made themselves using a scheduler. It has the general functions of schedulers, for example, allowing users to add, edit, delete, and view an event. When users add an event to the scheduler, the system constructs a suggested workflow from the user's own private database and links it to the event. The conjecture strategies are as follows:

1. It is based on case-based reasoning. Our system searches for the past events that are similar to a new event. To do that, it calculates the differences between two events' strings using the Levenshtein Distance. If there is a similar event, our system creates a workflow that is linked to the event.
2. When strategy 1 is not available, our system searches for keywords of a workflow linked to the past or in-progress events. Then, it creates a workflow that has many keywords matching the new event.
3. When strategy 2 is also not available, our system creates a workflow used recently.

3.5 How to Evaluate Arrangements and Evaluation Items

When a user finishes a workflow, he/she can evaluate it with a 5-point scale. The user can also add some comments as the evaluation. Each evaluation result is attached to the corresponding workflow. That is, if the workflow was made by customizing a local copy from a public database, the evaluation result could be attached to both the local workflow in his/her own private database and the original workflow in the public database. The former can be used as a note in the future. but the latter can be helpful for others and for the user who first published the workflow because they can understand the usefulness of the workflow from the evaluations. Moreover, evaluation results from others lead to a more efficient workflow.

According to the YouTube blog [5], most users use the lowest or highest-level rankings in YouTube's 5-star rating system; in fact, the 5-star rating is dominant the ratings, and the medium rating is seldom used. Therefore, we included a clap function in our system as a simple evaluation method.

In addition to obtaining evaluations from users, our system counts the two items below for each workflow as a form of evaluation:

1. Number of workflows copied.
   Effective or popular workflows are copied and used by many users. In our system, users can find out which workflows are used most often by many users.
2. Number of views.
   As stated in 1, effective or poplar workflows are copied by many users. Therefore, such workflows have more traffic. Our system counts the number of views and assists users' browsing of arrangements.

4 Discussion

4.1 User Motivation

It is important to motivate users to publish arrangements and mutually evaluate them. Miura, et al. [6], studied the motivation for users of Yahoo! ANSWER [7], in which users can ask questions and get answers from many users. According to the study, the questioners' motivation is divided into the following three classes:

1. Community revitalization and expressive identity
2. Users can benefit from others' answer
3. To achieve psychological satisfaction (users want to know the daily question's answer)

The functions of our system are intended to produce these types of motivation because it provides a forum for sharing arrangements.

On the other hand, the answerer's motivation is divided into the following four classes:

1. If you give an answer to others based on past experience, then you could get an answer from others for the future
2. To enjoy the merits of providing information in order to receive evaluations and appreciative words
3. For altruistic reasons; the answerer wants to provide information to others
4. Community revitalization and entertainment

We think 1 and 2 will be achieved through the function to publish workflows on the public database of our system. 3 can be achieved with the function to evaluate the workflows that are used. With this function, users can receive evaluations and positive comments about their published workflows. 4 will be solved while using our system.

The answerers' motivation is especially important because it will lead to an increase in the number of published workflows in the public database of our system. In the future, we have to evaluate whether the functionalities of our system are effective to motivate users. This requires an experiment based on a long-term running of our system.

4.2 System Availability

We discuss usage scenarios and need for system.

Our system is helpful to build up the habit of making an arrangement. It is difficult to get own habit from others' arrangements. This is because it links a suggested workflow to the event registered to user's scheduler and he/she can perform work-items included in the workflow correctly. In addition, our system can recommend others' arrangements to users and they can customize the arrangements,

When you have to do something that you have not experienced with before, it is useful to know better way to do that. For example, when traveling aboard for the first time, if you know how persons with experiences made arrangements for the trip, you can prepare for your trip properly.

Our system provides the function to share workflow and along with a scheduler. Thus, the system will make it easy to incorporate others' workflow into users' lifestyles.

4.3 Similar Systems

RESTER2 is a system that is similar to our system. It has a ToDo management function with an ontology of specific office tasks. Users can search for and use one from the ToDos in the process ontology. Tasks supported by it are routine ones and ad-hoc ones, so it is useful for office workers.

On the other hand, the purpose of our system is to improve efficiency of daily tasks and life events. It combines arrangement management and scheduling functions. The combination is one of the key characteristics of our system. Moreover, it manages an arrangements as a workflow consisting of work-items. It is suitable for personal daily tasks as well as specific office tasks, whether long-term or short-term.

In addition, our system gives users opportunities to make many arrangements from the recommendation function based on their profiles. The The connection between using arrangements and the scheduler is one of the key characteristics of our system.

5 Conclusion

In this paper, we proposed a system to share arrangements for daily tasks and a life event on the Web the purpose is to improve productivity for users.

To achieve this purpose, our system incorporates the concepts of workflow to manage arrangements, and provides a scheduler function to make good use of them. These devices lead to effective production and usage of arrangements. In addition, users can evaluate workflows that are used. Furthermore, our system links a proper workflow to an event added to the schedule by a user, and recommends a suitable one based on the user's profile. These functions make it possible for users to choose from many good workflows.

In the future, we plan to develop new input and output methods. It is also necessary to add a the function to deal with scheduling problems when users add multiple tasks in the scheduler and to help them see each task's importance and efficient order. In addition, we need to conduct some experiments to evaluate the performance and efficiency of our system, which includes an evaluation on the quality of published workflows.

References

1. Fukui, M., Sakai, K., Shibazaki, Y., Ohtake, Y., Nakayama, Y.: Practice of Knowhow Sharing in Office Environment using Knowledge and Information Sharing System. In: Information Processing Society of Japan, pp. 13–18 (1998)
2. Saga, R., Okada, N., Kageyama, A., Aoki, S., Tsuji, H.: RESTER2: Ontology based Reusable ToDo Synthesizer. In: IEEE International Conference on Systems, Man & Cybernetics (IEEE/SMC 2007), pp. 214–219 (2007)
3. Remember The Milk, `http://www.rememberthemilk.com/`
4. Google Calendar, `http://www.google.com/calendar/`
5. YouTube Blog: Five Stars Dominate Rating, `http://youtube-global.blogspot.com/2009/09/five-stars-dominate-ratings.html`
6. Miura, A., Kawaura, Y., Jifuku, S., Otaki, N., Okamoto, M.: People who create knowledge-sharing community. In: The 20th Annual Conference of the Japanese Society for Artificial Intelligence, vol. 3D3-1 (2006)
7. Yahoo! ANSWERS, `http://answers.yahoo.com/`

Population Estimation of Internet Forum Community by Posted Article Distribution

Masao Kubo, Keitaro Naruse, Hiroshi Sato, and Takashi Matsubara

[1] National Defense Academy of Japan, yokosuka, kanagawa,239-8686, Japan
{masaok,hsato,matubara}@nda.ac.jp
http://www.nda.ac.jp/cc/cs

[2] University of Aizu, Aizu-Wakamatsu, Fukushima-ken 965-8580 Japan
naruse@u-aizu.ac.jp
http://web-ext.u-aizu.ac.jp/official/index_e.html

Abstract. Traditionally, it seems that there are no typical patterns or regularities on an internet forum and the bulletin board because it is a result of voluntary behaviour of an individual user. However, recent studies show distribution of the number of posted articles per user on them follows as a power law distribution. We think that this regularity of the emergent behaviour is very important for understanding social science and human communication because it is possible to estimate the number of lurkers of each forum thread only using by usually observable data. In this paper, we proposed a set of estimation methods based on preferential attachment which is emerged by reciprocity as a human trait. The effectiveness of our proposed estimation methods is verified by comparing it with indices to be strongly related to the population of its corresponding human behaviour.

Keywords: Information extraction on the web, visualization of internet community, web intelligence, web mining.

1 Introduction

In this paper, a new estimation method of population[3] of internet community in which members are interested in internet forums, bulletin boards, and message board[10], is proposed. Usually, anyone who can read its messages on such internet media also can post. One of coveted information for these user is its population because it is important information to select a forum. However it is nearly impossible to know unless he/she has authority to log into its server. Therefore, the question whether we are able to estimate its population by using only published information is natural and important.

The community is divided into 2 groups, people who posted at least one messages and people who have not posted yet[8][9]. The population of the former group is found by enumerating ID of each message. However, it is difficult to count the population of the later group in the same manner as the former. We think it is possible to estimate the population by using the published information and usually observable data of forum. Here we focus on the regularity of

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 298–307, 2010.

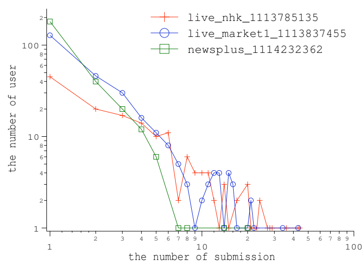


Fig. 1. The frequency distribution of posting activity

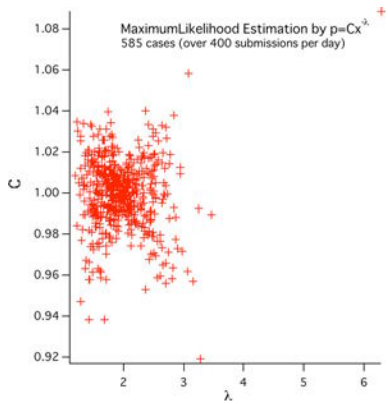


Fig. 2. Result of regression analysis of 584 bulletin boards which include over 400 messages

frequency of the number of posts per member[7].The population of people who posted and the number of messages per person form a line in log-log plot[11][1]. By using this regularity, we propose the estimation methods of the population which includes people who have not posted yet. These methods utilize only posted messages which anyone can usually read.

Note that there are no clear definitions for community of internet forum. Traditionally, people who have posted are admitted as member of its community[9]. This definition is proper but inconvenient because the traditional definition cannot handle candidates for a new poster of forum. Actually the question whether a person who read a forum a few seconds/hours should be included as a member is difficult.Then, in this paper, we define a community as a set of person who definitely submits its messages when enough time is passed. The proposed methods deduce the population under this assumption.Therefore, it may different with the logged data results on server. It may be also a meaningful question how different with the logged information the estimations are.

The remainder of the paper is organized as follows. In the next section, we discuss how reasonable preferential attachment is as a generating mechanism for the regularity of posting activity. In the third section, 2 estimation methods of the population of community are proposed. Finally, we discuss the differences between the estimation and actual data. In this paper, comparison with viewing rate of television suggests reasonable performance of this approach.

2 Characteristics of Posting Activity of Internet Forum

2.1 Data Source

We collected data from http://www.2ch.net/, the biggest Japanese internet bulletin boards system which opens over 700 categories which consists about and

11.7 million people uses. Each category exhibits a few hundred bulletin boards. There are about 100 million page view per day. We collected data from this web site from June 2005 to August 2005, and from December 2006 to March 2007 and total 327,333 bulletin boards are obtained.

2.2 Characteristics of Posting Activity

Let me show a few examples of the frequency of contribution of internet forum. Figure 1 illustrates that of 3 different forums. The x axis means the number of messages per user, the y axis means the number of corresponding user. For example, point (x,y)=(1,45) means that there are 45 person who submit once. In the logarithmic scale, the all of distributions seems to forms a line. Figure 2 shows the result of more statistical analysis, the result of regression analysis of the collected data is shown. 584 bulletin boards which has over 400 submission per day are selected from the collected data and the coefficients of Cx^{λ} are estimated. The 2 coefficients of 584 cases are shown. In this figure, many of them concentrate around (C, λ)=(1,2.0).

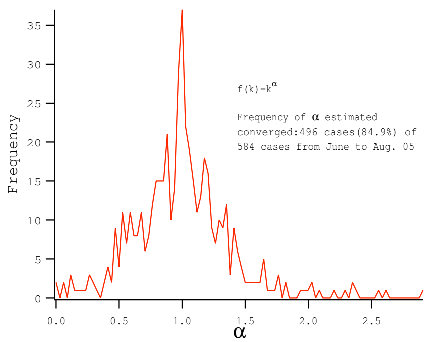


Fig. 3. The coefficint α of Preferential Attachment estimation

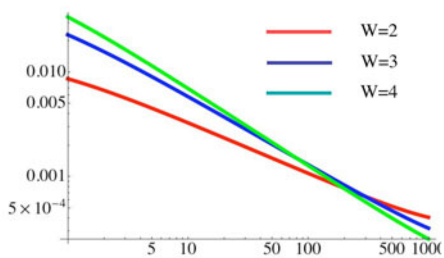


Fig. 4. Example of the proposed model(eq.3)

2.3 Preferential Attachment as a Generating Mechanism of This Power Law Like Trend of Internet Forum Posting Activity

In this paper, we suppose that all member of community submit their messages sooner or later. If their purpose to involve these media is communication, they will submit again when any comment messages by other users occur. Actually, [5] reports that frequency of posting an addition a new message increase by 12% when any comment messages occur.

Therefore, we suppose that Preferential attachment [2] is a good candidate for generating this power law like trend of internet forum because we think this reciprocity promotes similar dynamics as "rich man gets richer" :more frequently posting user submits more because it has much larger chance to obtain comment messages from other user.

If preferential attachment is the generating mechanism of this posting activity, the probability to submit a message of user i , $\Pi(k)$ can be written as

$$\Pi(k) = \frac{k^{\alpha}}{\Sigma_{k'} k'^{\alpha}} \tag{1}$$

where k is the number of messages who post and $\alpha = 1$.

Then, their coefficients α of the collected data are investigated by the method of [6]. The α of 496 cases (84.9%) out of 584 cases is obtained successfully. Figure 3 shows the results of estimation. As you see, there is a high peak at $\alpha = 1$. Therefore we can conclude that preferential attachment is one of main mechanism of this posting activity.

3 The Proposed Method

3.1 The Community Model

So far, we show it is a reasonable idea that Preferential attachment is a good candidate for generating mechanism for this power law like regularity of the posting activity of internet forum. Actually, we can estimate the coefficient α of each set of messages of a forum. However, it is still difficult to estimate of the population of people who have not posted yet. The reasons are as follows. (1) Initial population of a community is not small whereas the number of initial node in preferential attachment is very small. (2) The population of whole of community should be constant whereas the number of node of Preferential attachment is growing constantly. This second assumption may sound queer. Actually, the population of community is changing continuously because someone leaves/joins. However, it is inconvenient when it operates. Then, it seems to be reasonable that population supposing of fixed population assumption is estimated by data collected during sufficiently short term. We suppose that community does not change in a day. We think that this assumption is reasonable and suitable for more long term analysis because it is not too short and too long. By these assumptions, the following mechanism based on Preferential attachment is proposed.

Firstly, we suppose that there is 1 bulletin board and its community. Let N be the population of this community.There is only 1 message posted per time step and the bulletin board includes total t messages posted at time step t. Also, let $n_i(t)$ be the number of messages posted by member i. Then the probability $\Pi(i)$ to post an additional message by member i is

$$\Pi(i) = \frac{wn_i + 1}{\sum_{j=1}^{N}(wn_j + 1)} \tag{2}$$

where w is a weight parameter and let $w \geq 0$. Members who have posted a large number messages submit again with high probability while member who has no messages posted yet (n_i=0) has a chance to submit. If w=0, their posting activity is random because all $\Pi(i)$ is same.

In this community model, the frequency distribution $P(k)$ of the number of messages posted per member is

$$P(k) = \frac{\partial}{\partial k}(1 - \int_0^{(t+\frac{N}{w})\frac{1+\frac{1}{w}}{k+\frac{1}{w}} - \frac{N}{w}} P_{first}(t)dt) = \frac{\Gamma(\frac{(1+w)(N+tw)}{w(1+kw)} - \frac{1}{w})}{(1+kw)\Gamma(\frac{(1+w)(N+tw)}{w(1+kw)})} \tag{3}$$

where $P_{first}(t_i)$ is the probability which it is time step t_i when member i first submitted its message.

$$P_{first}(t_i) = (\prod_{t=0}^{t_i-1}(1 - \frac{1}{wt+N}))\frac{1}{wt_i+N}. \tag{4}$$

Figure 4 shows several example $P(k)$ with different w(=2,3,4).The x axis is the number of posting per member and the y axis indicates the number of corresponding members. Let $N = 1000$ and $t = 1000$. As you see, this community dynamics can generate similar distribution as that of actual data.

3.2 The Proposed Community Population Estimation Methods

3.2.1 The Principle to Estimate the Population When Time Passed Enough

The population N can be found if the parameters of $P(k)$ in eq.3 is estimated by fitting the actual data. However it is not a good idea because this equation includes Γ function which cause large error.

Therefore, other approaches also should be considered. Firstly, if time passed enough from when the bulletin board started, the population is estimated by the limit of eq.3.

Let $Q(k)$ denote fraction of 2 consecutive posting probability. Namely, $Q(k) = \frac{P(k)}{P(k+1)}$. The limit of $Q(k)$ at $t = \infty$ is

$$\lim_{t\to\infty} Q(k) = (\frac{1+(k+1)w}{1+kw})^{\frac{w-1}{w}}. \tag{5}$$

Therefore, given actual data of r_k which is the fraction of population of k time poster and $(k+1)$ poster,

$$r_k = (\frac{1+(k+1)w}{1+kw})^{\frac{w-1}{w}} \tag{6}$$

w can be estimated numerically. Note that the maximum of r_k is

$$\lim_{w\to\infty}(\frac{1+(k+1)w}{1+kw})^{\frac{w-1}{w}} = 1 + 1/k. \tag{7}$$

Therefore, we may need extra caution if the actual value is larger than this bound. Once w is determined, $Post_0$ which is the population of which people have not posted yet can be found by $Post_1$ for example.

$$Post_0 = Post_1 \lim_{t\to\infty} Q(0) = Post_1(1+w)^{\frac{w-1}{w}}. \tag{8}$$

So far, the population estimation method is proposed supposing that it observes when time passed enough.

3.2.2 The Estimation Method When the Observation Time Is Known

The method proposed at the last section can be applied only when times passed enough. In this section, we propose another estimation method at the observation time t.

Let $Post_x(t)$ be the population of which people have posted x times until time step t. Also "$^-$" is introduced for expectation.

The group of 0-time poster gets smaller with probability $\frac{Post_0(t)}{tw+N}$. Therefore the expectation at $t+1$ is as follows.

$$\overline{Post_0}(t+1) = \overline{Post_0}(t) - \frac{\overline{Post_0}(t)}{tw+N} \tag{9}$$

From this recurrence formula,

$$\overline{Post_0}(t) = \frac{N\Gamma(t-\frac{1}{w}+\frac{N}{w})\Gamma(\frac{N}{w})}{\Gamma(t+\frac{N}{w})\Gamma(-\frac{1}{w}+\frac{N}{w})}. \tag{10}$$

In the same manner as $Post_0(t)$, the recurrence formula for $Post_1(t)$ is

$$\overline{Post_1}(t+1) = \overline{Post_1}(t) + \frac{\overline{Post_0}(t)}{tw+N} - \frac{\overline{Post_1}(t)(w+1)}{tw+N}. \tag{11}$$

Then, we get

$$\overline{Post_1}(t) = \frac{t\Gamma(1+\frac{N}{w})\Gamma(-1+t-\frac{1}{w}+\frac{N}{w})}{\Gamma(t+\frac{N}{w})\Gamma(-\frac{1}{w}+\frac{N}{w})}. \tag{12}$$

Generally,$Q_{x/x+1}(t)$, the fraction of x times poster and $(x+1)$ times poster is

$$Q_{x/x+1}(t) = \frac{\overline{Post_x}(t)}{\overline{Post_{x+1}}(t)} = \frac{(x+1)(-1+N+(-(x+1)+t)w)}{(-x+t)(1+xw)}. \tag{13}$$

Therefore, for example, given 3 consecutive population data, for example, $Post_1(t)$, $Post_2(t)$ and $Post_3(t)$, N and w are found by the simultaneous equations.

Let $Post_{x-1}(t)$, $Post_x(t)$, $Post_{x+1}(t)$ be given. In this case, $\hat{w}$, the expectation of w is

$$\hat{w} = Post_{x-1}(t)Post_{x+1}(t)(-1-t+x)(1+x) + Post_x(t)^2(-x+t)/D. \tag{14}$$

And $\hat{N}$, the expectation of N is

$$\begin{aligned}\hat{N} = (xPost_x(t)(t(-t+x)Post_x(t) - (1+x)Post_{x+1}(t))\\ +(1+t-x)Post_{x-1}(t)((-t+x)Post_x(t)\\ +(-2+t)(1+x)Post_{x+1}(t))/D.\end{aligned} \tag{15}$$

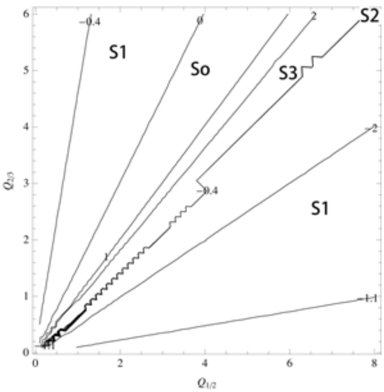


Fig. 5. Map of $\hat{w}$:$Q_{1/2}(t)$,$Q_{2/3}(t)$ are given, $t = 100$

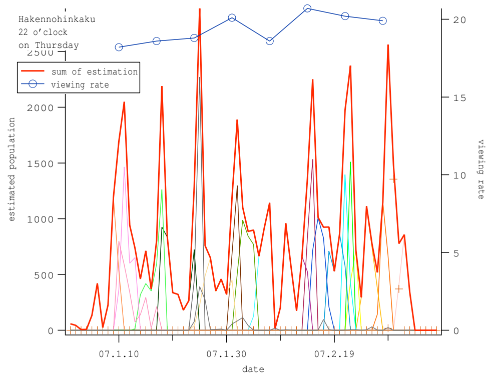


Fig. 6. Example of estimated population and viewing rate of TV program

3.2.3 Analysis of the Proposed Method

In this section, we give an analysis of eq.14DFig.5 shows contour map of $\hat{w}$ when $Q_{1/2}, Q_{2/3}$ are given at $t = 100$. This map can be categorized into the following 4 regions:$\hat{w}$ is positive (**So**),$\hat{w}$ is negative(**S1**),region where $\hat{w}$ is not derived(**S2**), and region around $D = 0$ where $\hat{w}$ is extreme large (**S3**).

The eq.14 does not consider any fluctuation. Therefore, as a matter of fact, it sometimes happens that an estimated $\hat{w}$ is in out of **So** region by fluctuation and then this method can not estimate the population.

Then, we introduce the following additional heuristic estimation rules when the method of eq.14 does not work.

S1 $w < 0$
$Post_0 \leftarrow Post_1$. We suppose there are at least as same size of people as $Post_1$

S2,S3 $Post_0 \leftarrow 0$. We cannot obtain any meaningful information when $D = 0$

S4 $w \gg 1$
$Post_0 \leftarrow \frac{Post_1}{Post_2} Post_1$. This value suggests the Preferential attachment model is not adequate.

where **S4** includes the case which $Q_{1/2}, Q_{2/3}$ are not obtained.

4 Experiment

In the last section, the population estimation methods are proposed.In this section, the proposed method is evaluated by actual data.

Generally speaking, if logged data is not utilized, evaluation of such method is extremely difficult task that this seems to be a bottleneck of social science. In this paper, we evaluate our proposed methods by calculating correlation[4] with indices which has strong relation between the population.For example, the number of votes and related community size have strong relations. In this paper,we

Table 1. # of bulletin boards of each TV drama

Program title	# of bulletin boards
K1	5
H1	24
H2	13
K2	30
K3	6
T1	9
H1	7

Table 2. Comparison with viewing rate: Pearson's product moment correlation coefficient r(eq.17)

	Only Poster Data	Proposed
correlation coefficient	0.7565	0.7987

evaluate our proposed methods by comparison with viewing rate of television drama program and its related bulletin board's community size. Fortunately, viewing rate of television drama program is published from many companies. If the proposed estimation is good, we suppose it makes the correlation stronger.

4.1 Correlation with Viewing Rate and the Estimated Population about TV Drama Related Internet Forum

To evaluate the proposed method, we collect 7 television drama viewing rate and their corresponding bulletin boards (see table 1) from Dec. 2006 to March 2007. Then, the following 2 types of correlations are calculated.

1. correlation between viewing rate and the number of people who have posted. It can get by enumerating ID on bulletin boards.
2. correlation between viewing rate and population of community which includes people have not posted yet. The community size is estimated by the proposed method.

4.1.1 Experiment Detail

It is an advantage of internet forum that people can read/write anytime. Someone accesses a internet forum whereupon a TV drama program starts to be broadcasted while some other people read/write after a few days. Therefore, we suppose that a viewing rate of weekly TV program should be compared with sum of population for a week.First, we estimate a population $\hat{N}$ of a bulletin board for each day from December 2006 to March 2007. We used eq.16 and the heuristics. As table 1 shows, there is not only one bulletin board for each TV program. Next the correlation between a sum of the estimated population for a week of all related bulletin boards and its viewing rate is calculated as follows. Now, let $\hat{N}_x^{day}$ be the population of community $x \in B_d$ on a day. The sum of population for a week $pop_{B_d}^{week}$ is

$$pop_{B_d}^{week} = \sum_{x \in B_d} \sum_{day \in DAY} (\hat{N}_x^{day}) \tag{16}$$

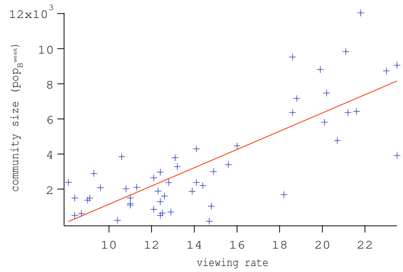


Fig. 7. Correlation between estimated population and viewing rate

where $day \in DAY = \{day1, ..., day7\}$ is each day after broadcast. Pearson product-moment correlation coefficient r is calculated as follows.

$$r = \frac{\frac{1}{M}\sum_i (pop_i - \overline{pop})(v_i - \overline{v})}{\sqrt{\frac{1}{M}\sum_i (pop_i - \overline{pop})^2}\sqrt{\frac{1}{M}\sum_i (v_i - \overline{v})^2}}. \quad (17)$$

We suppose that there are M related bulletin boards.

4.1.2 Results

The estimation results are shown in figure 6,7, and table2. Figure 6 shows an example of estimated daily population transition of internet communities about a TV program. The x axis is day, and the y axis means the sum of population(solid line) and its corresponding viewing rate().The bold line indicates the sum of population of related bulletin boards. Figure 7 plots the pairs, viewing rate and $pop_{B_d}^{week}$. The x axis shows viewing rate and the y axis is $pop_{B_d}^{week}$. As you see, these form a line. Table 2 shows the result of comparison of correlation between population of poster only and the estimated population which includes people who have not posted yet.The former community is obtained by only enumeration of ID of posted messages. By comparison, the former scores high correlation (0.7565) between with these viewing rate. However, the proposed population estimation methods score more high correlation (0.7987) where the number of sample is 51 and the significant level(5%,1%), are about 0.28 and 0.37 respectively.

This does not contradict our expectation. We think that this series of experiments prove the validity of out proposed population estimation methods.

5 Conclusion

In this paper, population estimation methods of community of internet forum and bulletin boards. Traditionally, administrator only know population of community. However, anyone can know the population by using these methods because they only utilize published information and usually observable data, for example, ID and time stamp of messages.

The proposed methods applied the Zipf's law like regularity about the posting activity to population estimation.With small modification of Preferential attachment, the generating model of posting activity of internet forum is proposed. Then, 2 different estimation methods are proposed, the one only uses ID information and another method utilizes ID and time stamp of each messages.

Finally, we evaluate our methods by using correlation between viewing rate of a television program and the estimate population of its corresponding internet community. As a result, the correlation with our estimation methods score slightly higher correlation than the correlation without our estimation methods. We think this suggest the validity of our methods.

This paper is supported by Grant-in-Aid for Scientific Research of Japan (KAKENHI).

References

1. Albert, R., Barabási, A.-L.: Statistical mechanics of complex networks. Reviews of Modern Physics 74(1), 47–97 (2002)
2. Barabási, A.-L., Albert, R., Jeong, H.: Mean-field theory for scale-free random networks. Physica A 272, 173–187 (1999)
3. Bernard, H.R., Killworth, P.D., Johnsen, E.C.: Gene A
4. Gruhl, D., Guha, R., Liben-Nowell, D., Tomkins, A.: Information Diffusion Through Blogspace. In: Proceedings of the 13th International Conference on World Wide Web, pp. 491–501 (2004)
5. Joyce, E., Kraut, R.E.: Predicting Continued Participation in Newsgroups. Journal of Computer-Mediated Communication 11(3), 723–747 (2006)
6. Massen, C.P., Doye, J.P.K.: A self-consistent approach to measure preferential attachment in networks and its application to an inherent structure network. Physica A: Statistical Mechanics and its Applications 377(1), 351–362 (2007)
7. Keitaro, N., Masao, K.: Lognormal Distribution of BBS Articles and its Social and Generative Mechanism. In: Proceedings of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence, pp. 103–112 (2006)
8. Nonnecke, B., Preece, J.: Lurker demographics: counting the silent, Conference on Human Factors in Computing Systems. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 73–80 (2000)
9. Smith, M.A.: Measures and Maps of Usenet. In: From Usenet to Cowebs: Interacting With Social Information Spaces, pp. 47–78 (2003)
10. White, N.: Community Member Roles and Types (2001), http://www.fullcirc.com/community/memberroles.htm
11. http://www.nslij-genetics.org/wli/zipf/

Development of Delay Estimation Method Using Probe Data for Adaptive Signal Control Algorithm

Hisatomo Hanabusa, Morihisa Iijima, and Ryota Horiguchi

i-Transport Lab. Co., Ltd.,
3F, Jimbocho1-4BLDG, Kanda Jimbocho 1-4, Chiyoda-ku, Tokyo, 101-0051, Japan
{hanabusa,iijima,rhoriguchi}@i-transportlab.jp

Abstract. This study describes a real time delay estimation method using probe data with adaptive signal control algorithm, "CARREN" (Control Algorithm Retuning paRameters with self performance EvaluatioN) [1]. So far, CARREN supposed the use of AVI (Automatic Vehicle Identifier) data to estimate the total delay on approaching each intersection. In this paper, the application of CARREN using probe data in place of AVI data is explained subsequently. After the explanation, the real time delay estimation method is validated using traffic simulation. In the conclusion section, the performance of the proposed method will be discussed with respect to the penetration rate of probe vehicles, as well as the issues to be tackled in the future.

Keywords: Adaptive Signal Control, Probe Data, Total Delay Estimation.

1 Introduction

The purpose of this study is to develop a delay estimation method using probe data for a real time signal control algorithm. The conventional traffic signal control systems used in Japan design signal parameters based on the degree of saturation. In this case, the optimal signal parameters are selected from pre-determined sets to manipulate the current traffic conditions. However, the systems can only adapt to limited conditions because of the limited number of pre-determined sets of signal parameters. In addition, a signal control based on the degree of saturation does not guarantee a minimization of traffic congestion cost (=delay) as an original purpose. Therefore, It is important that the delay time is used for the adaptable signal control in real time. For examples, MODERATO (Management by Origin-Destination Related Adaptation for Traffic Optimization) calculates the signal parameters to minimize the estimated delay time using the traffic volume from detectors. However, the estimated delay time is not measured directory. SCOOT (Split, Cycle and Offset Optimising Technique) estimates the queue length using detector data and evaluates by calculating the degree of saturation. SCOOT also does not measure the delay time directory. In these methods, the evaluation indicators based on the degree of saturation are used to decide the signal parameters. To estimate the indicators, the parameters used in the calculation of the degree of saturation such as the saturation flow rate and the travel speed are assumed (not measured in real time).

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 308–317, 2010.

The delay estimation method in this study is applied to the real time signal control algorithm CARREN for the validation in the traffic simulator AVENUE (an Advanced & Visual Evaluator for road Networks in Urban arEas). CARREN uses AVI systems and detectors to evaluate delay. AVI is the device to observe the plate number of the passing vehicles. This system can measure the travel time of the section between AVIs by the plate number matching process. However, the cost of the AVI system such as installation and maintenance is high. Therefore we discuss the possibility application of probe data in place of AVI system.

In recent years, various applications using probe data have been discussed such as the estimation of traffic congestion and vehicle behavior analysis for new technology and services in the field of road traffic. On the other hand, a variety of probe data has been collected and data processing techniques for probe data have been developed. It may be said that probe data is expected to be easily-available by the growing of automotive telematics services. For example, if the precision of probe data is relatively precise, the stop and go status of the vehicle can be discriminated. The method described by Horiguchi, et al.[2] The probe information we assume here has the link-wised linear trajectory data identified by the entry/exit times of each link. In this case, we should consider how to estimate the total delay in real time from the probe data according to the penetration of the probe vehicle.

2 Algorithm of CARREN

CARREN is the real-time signal control algorithm proposed by Koshi, et al. They described that CARREN is used on the premise of observing travel times of all passing vehicles and the traffic volume by AVI systems and detectors. The algorithm

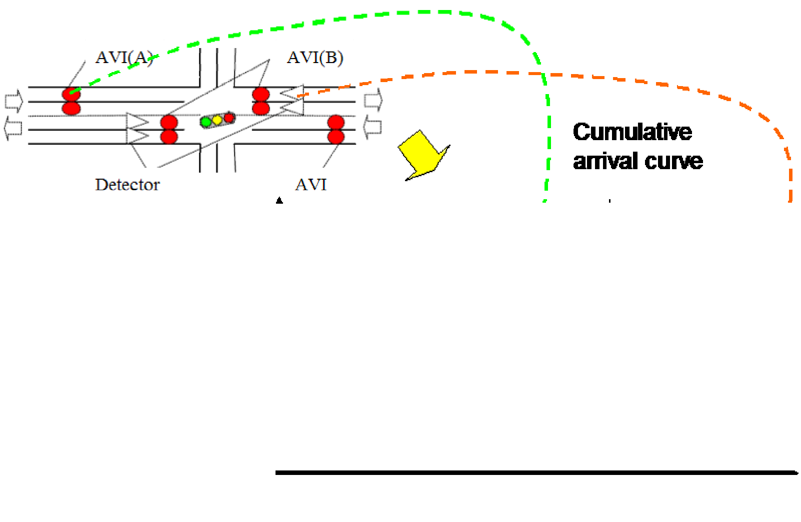


Fig. 1. Estimation of total delay by signal

decides signal parameters (split, offset and cycle) using the total delay time for each stream. Fig.1 shows the image of the total delay evaluation of CARREN. AVI(A) and AVI(B) are located at upstream side and downstream side of the target intersection. The detector is located at the same position of AVI(B). From AVI system, travel times of the passing vehicle between AVI(A) and AVI(B) is observed. The exit time of the section can also be observed from AVI system. From the detector, traffic count with passing times of the vehicle is observed. Using the traffic count data, the cumulative departure curve can be drawn.

Following the observation, the cumulative arrival curve is drawn by the travel times, the exit times and the cumulative departure curve to calculate total delay time. First, the position of each vehicle on the cumulative departure curve (=departure position) is plotted by the exit time. After drawing the cumulative departure curve, the arrival time of each vehicle is calculated by the travel time. Then the cumulative arrival curve is drawn by the arrival positions excluded the free flow travel time. Therefore, the area enclosed by the cumulative arrival curves and the cumulative departure curve in Fig. 1 is the total delay time.

CARREN finds the best combination of split, offset and cycle to minimize the total delay time. Fig. 2 shows the evaluation image of the signal parameter change. To evaluate the parameter change, CARREN estimates the new cumulative curve when the split and the offset are increased or decreased by increments. After the estimation, the relationship between the original total delay time and the estimated total delay time is organized to the matrix of Table 1. f in Table 1 is total delay when the split and the offset are changed ($-\Delta$ or $+\Delta$). For example, if the split is decreased, the arrival cumulative curve is extended and the cumulative departure curve is translated for Δ seconds. On the other hand, if the split is increased, the cumulative arrival curve is shorter than the original one and the cumulative departure curve is translated for $-\Delta$ seconds. Therefore, the area of total delay is decreased.

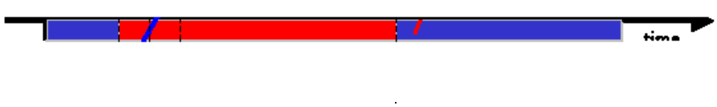


Fig. 2. Evaluation image of signal parameter change using the cumulative curves

Table 1. Matrix of the total delay time by changing the signal timing

		Offset		
		Δ	0	Δ
Split	Δ	f_{--}	f_{-0}	f_{-+}
	0	f_{0-}	$f_{00}=0$	f_{0+}
	Δ	f_{-+}	f_{+0}	f_{++}

In CARREN, the matrices of the total delay time are calculated for all directions of the inflow of each intersection. Fig.3 shows the image of the split adjustment. Basically, signal controller has some signal phases to control the traffic flow from the different directions. In addition, signal controller has to consider the cooperation with the neighbor intersections. Therefore, the total delay times of each signal in the target section are compared to each other, and searched the better combination to meet the requirements. The best combination of signal timing parameters can be determined when the all of *f* are evaluated. The application method of CARREN for practical use described by Asano, et al. [3] [4].

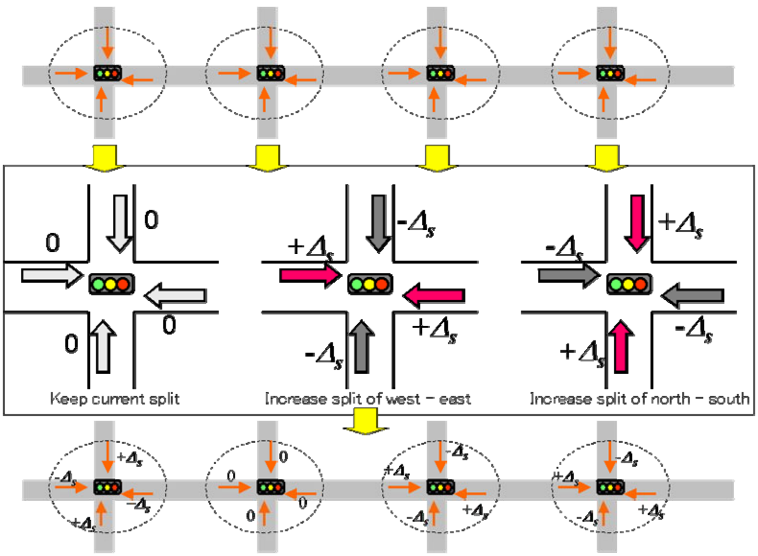


Fig. 3. Split adjustment

3 Delay Estimation Using Probe Data

In the first section on this chapter summarizes the issues for the delay estimation using probe data. The subsequent section explains the assumptions and the procedure of the delay estimation method which we propose.

There would be three issues when we develop the delay estimation method using probe data for CARREN. The first issue is how to draw the cumulative departure curve from the detector data. The cumulative departure curve is very important to identify the order of the departure of probe data. The second issue is how to draw the cumulative arrival curve with the imperfect link traverse data from the probe vehicles. The AVI data used in the original CARREN may provide the perfect information of all vehicles. However, the probe data only provides some samples of link traverses. It is obvious that we may draw the cumulative arrival curve with fairly reasonable accuracy when the penetration rate of probe vehicles is high enough. Therefore, the main issue is how to make the drawing method that is feasible when the probe penetration is low. The third issue is the time lag in the data transmission from probe vehicles to the signal control system. In general, the time lag may lose the accuracy of delay estimation when the traffic conditions are varying time by time. The time lag in data transmission should be considered for the practical examination.

Fig.4 shows the image of the delay estimation method. To estimate the total delay time, the cumulative arrival curve and the cumulative departure curve should be drawn. The cumulative departure curve is drawn by detector data located at downstream side of the intersection. When we assume that a certain percentage of passing vehicles are probe vehicles, the passing time at intersection and the arrival time estimated by travel time consist of a part of cumulative arrival curve arrival curve and cumulative departure curve. At this point, cumulative arrival curve is not drawn.

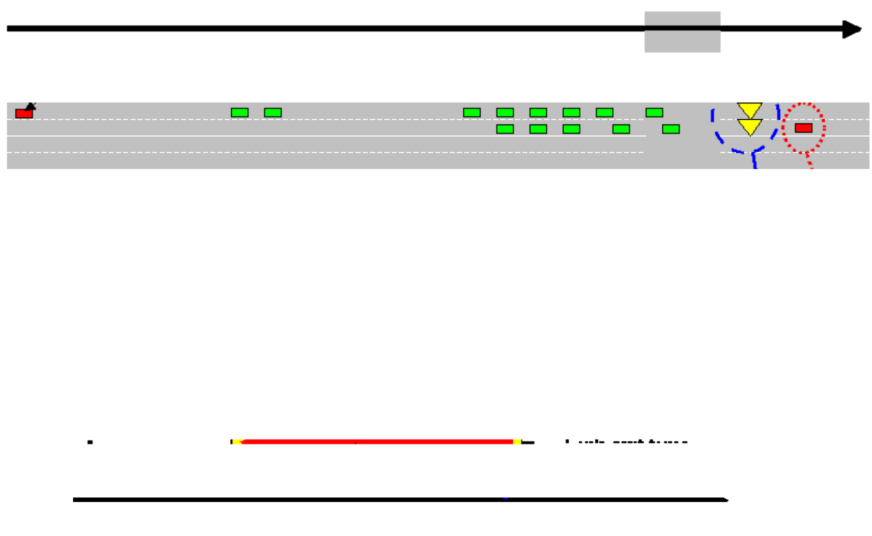


Fig. 4. System for total delay estimation

If probe vehicles are evaluated, the total delay of all passing vehicle is underestimated. Therefore, passing time of all vehicle at intersection is observed using detector (pulse data) to draw departure curve correctly. On the other hand, a part of arrival curve can be drawn by the passing time at intersection and link travel time. This method consists of probe data fusion process and arrival curve estimation process. Since the probe data provides link entry time and link travel time, the cumulative arrival curve can be approximated by interpolating the points of the probe entry times on the cumulative flow diagram. Fig. 5 shows the image of the probe data fusion process and the arrival curve estimation process in total delay estimation method. The following steps describe the process of the total delay estimation.

1) Plot arrival point of probe and departure point of probe on departure curve (departure curve is drawn by passing time of detector) every cycle. Arrival time on arrival curve is excluded free flow travel time.
2) Repeat the step of 1) for specified cycles (parameter of the method) and integrate past arrival curves and past departure curves over the specified cycles. (probe data fusion process)
3) Approximate arrival curve by two lines. (arrival curve estimation process)
4) Estimate total delay by cycle using the approximated arrival curve.
5) Repeat from step 1) to step 4) for every cycle.

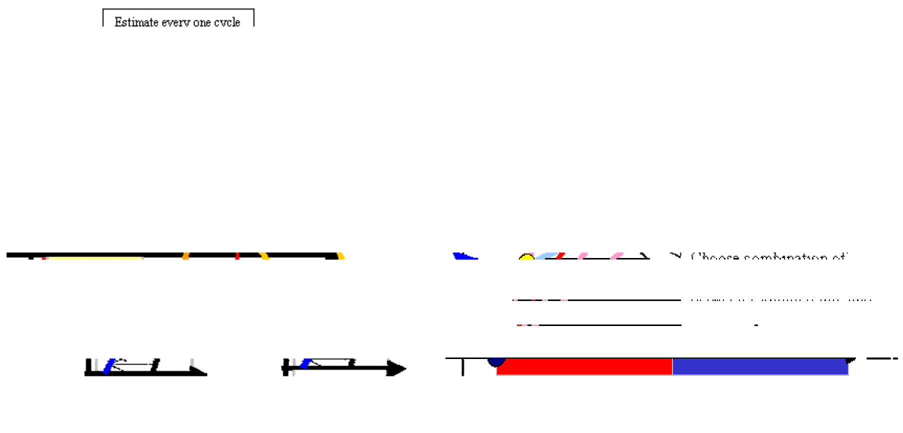


Fig. 5. Fusion of cumulative curve

Here, we simplify the arrival curve with two approximated lines. The reason is that we assume that cumulative arrival curve has two characteristics when traffic is controlled by systematic signal control. The first one is that the first arrival pattern from the upstream intersection can be in saturation flow for a few second to several tens of seconds. The second one is that the second arrival pattern can be in free flow traffic with demand flow rate after the saturated flow (the first arrival pattern). In the situation of saturated flow, the angle of the first approximate line should be equal to saturation flow rate. On the other hand, the angle of the second approximate line depends on the flow rate of demand and it should be smaller than the angle of the first approximate line. Adding to this, the total delay consists of one approximate line is simultaneously evaluated. Then we can have a variety of combination of approximate lines and select the best lines. To select the best approximate lines, RMSE (Root Mean Square Error) between the approximated lines and the arrival points using observed link travel times from probe data is calculated for each lines. Finally, the combination of lines that minimizes the RMSE is selected and is used for the delay estimation.

4 Implementation of the Delay Estimation Method

In this study, we use the traffic simulation model for the validation of the delay estimation method. Therefore the method should be installed as a part of CARREN. Fig. 6 shows the basic procedure of CARREN. *N* is the number of cycle for the data collection (obtain probe data and detector data). The method is installed between the data collection process and the signal timing update process. This procedure is executed every simulation scan (1 second) for each intersection.

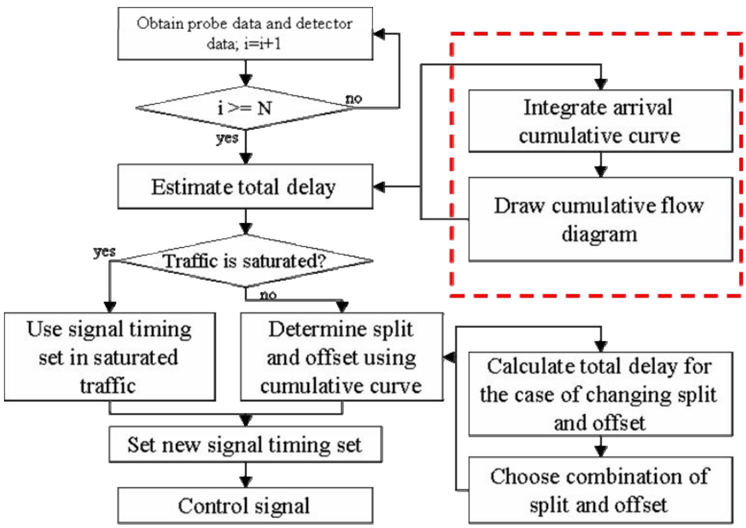


Fig. 6. Procedure of CARREN and the estimation method

5 Test Bed for Validation

As a test bed of this study, we have built a simulation dataset for Yasukuni-dori in Tokyo. Table 2 shows the overview of the test bed (Yasukuni-dori test bed). We surveyed traffic volume and signal timing for each intersection to develop Yasukuni-dori test bed. The travel times of the target section was also observed. This data is used as the input data for the traffic simulator and the validation data for the reproduced traffic situation. The OD traffic volume (Origin-Destination traffic volume) is estimated by the traffic volume data.

The traffic simulator used here is AVENUE. AVENUE employs the 'Hybrid Block Density Method (HBDM)' (Horiguchi et al. 1994). In the HBDM, each link is divided

Table 2. Outline of the Yasukuni-dori test bed

Network size	1.8km * 0.5km (61 links, 32 nodes)
Simulation time	From 7:00 to 19:00 (12 hours)
Total number of demand	127029 vehicles
Schedule of OD traffic volume	Every 10 min.
Type of vehicle	Small vehicle and large vehicle

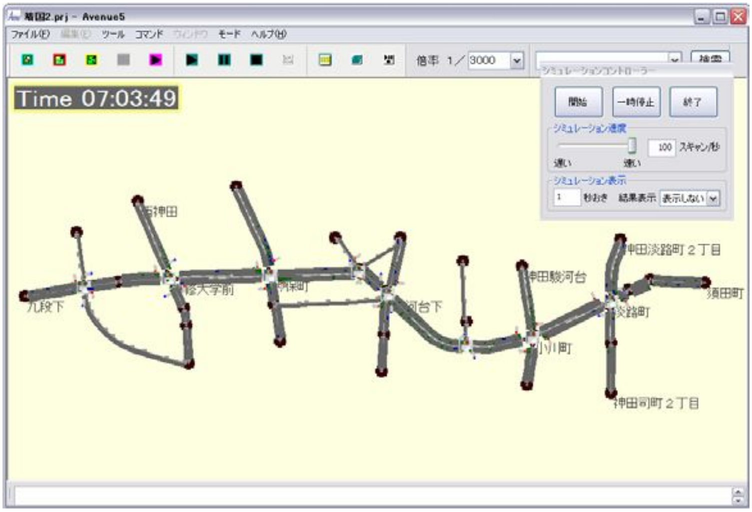


Fig. 7. Test bed on AVENUE

into several blocks. The in/out–flow and the density of each block are revised at every scanning interval base on the flow-conservation law and the Q-K (Capacity and Density) relationships. Fig. 7 shows the display of Yasukuni-dori test bed and AVENUE.

6 Validation of Delay Estimation

For the validation of the delay estimation method, one link section between the intersection A and the intersection B is chosen (the direction is from A to B). Fig. 8

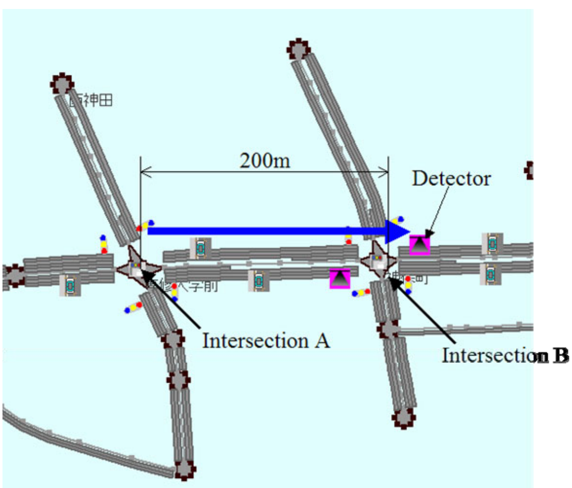


Fig. 8. Section for validation

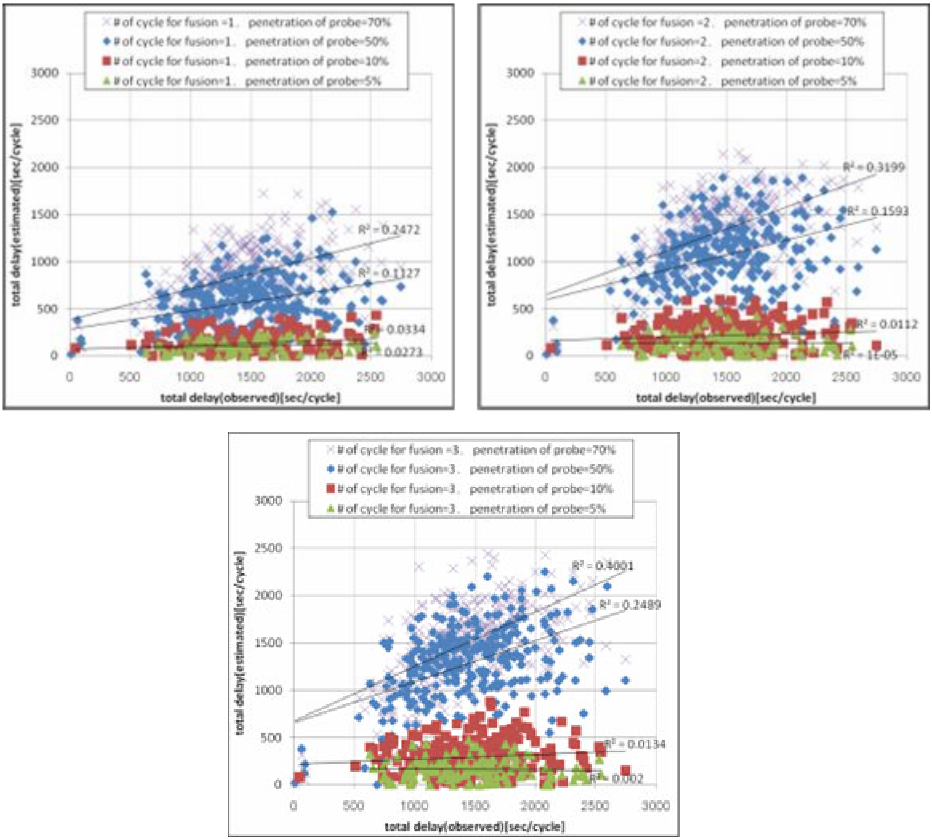


Fig. 9. Estimation result

shows the validation section. The link has 3 lanes with right turn lane. We set a detector on the downstream side of intersection B. As the probe data, the link entry time and the link exit time of probe vehicle are observed.

Fig. 9 shows the comparison result between the total delay estimation method and the total delay time of all vehicles. We simulated for 12 hours (from 7:00 to 19:00) and compared the estimated total delay times from the probe data with the measured total delay times (the total delay time from all vehicles as the true value) in every cycle. We examined 12 simulation cases changing the parameter combinations depending on the probe penetration and the number of cycles for the probe data collection.

The three figures in Fig. 9 describe that the accuracy of the estimated delay times tend to be high when the probe penetration or the number of cycles is high. When the probe penetration was relatively higher as much as 50-70%, the estimated delay times seem to correlate with the measured delay times. On the other hand, the estimated delay times were underestimated when probe penetration was relatively lower as much as 5-10%.

The reason of the underestimate includes that the estimation method tends to choose the approximated arrival curve consisted of one line when the probe penetration is low. If the one-line approximated arrival curve is chosen, the delay time of the first some vehicles in the saturated flow are shortened to the point on the arrival curve (the delay time includes the stop time by signal and the waiting time of the queue). Therefore, the real delay times in saturated flow might not be considered. In addition, the validation result was greatly varied in each case. The reason includes that the estimation method uses all probe information that is collected in few cycles for the total delay estimation. To solute the problem, we think that the data cleansing to exclude the unsuitable samples for the total delay estimation should be installed.

In this study, we simulate the Yasukuni-dori test bed with CARREN and the delay estimation method by AVENUE. The calculation time is less than 1 second per one scan (The scan interval of AVENUE is 1sec). Therefore, we expect the estimation method is adaptable for the real time signal control.

7 Future Works

In this paper, we explained a delay estimation method by using probe, and we developed a test bed of traffic simulation for the validation. Then the validation result of the method is showed in a part of the simulation network by AVENUE. As for the immediate future works, we are going to install some data processes, for example, cleansing process to exclude unsuitableness data. On the other hand, the method will be improved by considering the Kinematic Wave theory that describes the fundamental characteristics of traffic flows. Using the test bed and CARREN, the capability of the application of probe information will be validated considering effect of penetration of probe car and difference from signal timing in real.

References

1. Horiguchi, R., Wada, K.: Effective Probe Data Transmittal with Detection of Congestion Pattern. In: Proceedings of 11th World Congress on Intelligent Transport Systems (2004)
2. Koshi, M., Kuwahara, M., Akahane, H., Ozaki, H., Yasui, K., Nakamura, H., Oguchi, T., Yoshii, T., Horiguchi, R., Oneyama, H.: A Traffic Signal Control Algorithm using ITS Sensing Technologies: Committee of Infrastructure Plannning and Management. JSCE (2002)
3. Asano, M., Nakajima, A., Horiguchi, R., Oneyama, H., Kuwahara, M.: Traffic Signal Control Algorithm Based On Queuing Model Using ITS Sensing Technologies. In: Proceedings of 10th World Congress on Intelligent Transport Systems (2003)
4. Asano, M., Nakajima, A., Horiguchi, R., Oneyama, H., Kuwahara, M., Koshi, M., Akanahe, H.: A Real Time Traffic Signal Control by Self-Evaluating Delay. Japan Society of Traffic Engineering 20(4), 879–886 (2003) (in Japanese)
5. Traffic Simulation System AVENUE, `http://www.i-transportlab.jp/products/avenue`

OVACARE: A Multi-Agent System for Assistance and Health Care

Juan F. De Paz, Sara Rodríguez, Javier Bajo, Juan M. Corchado, and Emilio S. Corchado

University of Salamanca, (Spain)
{fcofds,srg,jbajope,corchado,escorchado}@usal.es

Abstract. This paper presents a case study in which the OVAMAH architecture is applied in order to obtain a multi-agent system (MAS) that can provide assistance and health care for Alzheimer patients. The system makes use of several context-aware technologies that allow it to automatically obtain information from users and the environment in an evenly distributed way, focusing on the characteristics of ubiquity, awareness, intelligence, mobility, etc., all of which are concepts defined by Ambient Intelligence. OVAMAH is made up of a group of related modules that are well-suited for developing systems in other highly volatile environments similar to these. Because the development of this type of system is complex, it is essential to thoroughly analyze the intrinsic characteristics of typical environment applications, and to design all of the system components at a very high level of abstraction.

Keywords: Multi-Agent Systems, Ambient Intelligence, Virtual Organizations.

1 Introduction

This study presents a dependable solution for using a novel architecture in designing and building a system for assistance and health care for Alzheimer patients. The importance of developing new and more reliable ways of providing care and support for the elderly is underscored by this trend, and the creation of secure, unobtrusive and adaptable environments for monitoring and optimizing health care will become vital. Some authors [14] consider that tomorrow's health care institutions will be equipped with intelligent systems capable of interacting with humans. Multi-agent systems and architectures based on intelligent devices have recently been explored as supervision systems for medical care for dependent people. These intelligent systems aim to support patients in all aspects of daily life [7], predicting potential hazardous situations and delivering physical and cognitive support.

Ambient Intelligence based systems aim to improve quality of life, offering more efficient and easy ways to use services and communication tools to interact with other people, systems and environments. Among the general population, those most likely to benefit from the development of these systems are the elderly and dependent persons, whose daily lives, with particular regard to health care, will be most enhanced [8]. Dependent persons can suffer from degenerative diseases, dementia, or loss of cognitive ability.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 318–327, 2010.

Agents and multi-agent systems in dependency environments are becoming a reality, especially in health care. Most agents-based applications are related to the use of this technology in the monitoring of patients, treatment supervision and data mining. [13] present a methodology that facilitates the development of interoperable intelligent software agents for medical applications. [9] propose a system to increase hospital efficiency by using global planning and scheduling techniques.

ALZ-MAS (*ALZheimer Multi-Agent System*)[8] is a distributed multi-agent system designed upon Ambient Intelligence and aimed at enhancing the assistance and health care for Alzheimer patients living in geriatric residences. The main functionalities in the system include reasoning and planning mechanisms [10] that are embedded into deliberative BDI agents, and the use of several context-aware technologies to acquire information from users and their environment.

One of the objectives of MAS is to build systems capable of autonomous and flexible decision-making, and that will cooperate with other systems within a "society" [5]. This "society" must consider characteristics such as distribution, continual evolution and flexibility, all of which allow the members (agents) of the society to enter and exit, to maintain a proper structural organization, and to be executed on different types of devices. All of these characteristics can be incorporated via the open MAS and virtual organization paradigm, which was conceived as a solution for the management, coordination and control of agent performance [12]. The organizations not only find the structural composition of agents (i.e., functions, relationships between roles) and their functional behavior (i.e., agent tasks, plans or services), but they also describe the performance rules for the agents, the dynamic entrance and exit of components, and the dynamic formation of groups of agents[3].

The goal of this study is to present a case study in which the OVAMAH (*Adaptive Virtual Organizations: Mechanisms, Architectures and Tools*) architecture is used to build an open MAS for assistance and health care for Alzheimer patients. OVAMAH is the evolution of architecture THOMAS (*MeTHods, techniques and tools for Open Multi-Agent Systems*) [6][11]. We will propose an application for this architecture and will evaluate its appropriateness for developing an open MAS in a real environment. The first step of this research involves designing the components needed for addressing all the needs and characteristics of a health-care system. The design is based on the GORMAS (*Guidelines for Organization-based Multi-Agent Systems*) [1] methodology, which is specifically geared towards organizations.

This article is organized as follows: section 2 presents the principle characteristics of the architecture and methodologies used; section 3 indicates the MAS that was developed for the actual case study (the health care system), and highlights the characteristics provided by the type of architecture used for its development; and the final section presents some of the conclusions obtained by this research.

2 OVAMAH Outline

The architecture we used is OVAMAH (*Adaptive Virtual Organizations: Mechanisms, Architectures and Tools*). OVAMAH is based on THOMAS [6][11]; as such it is in THOMAS where it was made the necessary modifications so that the system can be used as a model for adaptive virtual organizations. It is made up of a group of related modules that are well-suited for developing systems in other highly volatile environments similar to a shopping mall. It is based on a services oriented approach and primarily focuses on the design of virtual organizations. This design will use a high level of abstraction to determine which components are necessary for addressing all of the needs and characteristics of a assistance system. The architecture is basically formed by a set of services that are modularly structured. It uses the FIPA[1] architecture, expanding its capabilities with respect to the design of the organization, while also expanding the services capacity. OVAMAH has a module with the sole objective of managing organizations that have been introduced into the architecture, and incorporates a new definition of the FIPA Directory Facilitator that is capable of handling services in a much more elaborate way, following the service-oriented architecture directives. The architecture consists of three principle components: *Service Facilitator (SF), Organization Manager Service (OMS)* and *Platform Kernel (PK).* The SF primarily provides a place where autonomous entities can register service descriptions as directory entries. The OMS component is primarily responsible for specifying and administrating its structural components (role, units and norms) and its execution components (participating agents and the roles they play, units that are active at each moment). In order to manage these components, OMS handles the following lists: *UnitList*: maintains the relationship between existing units and the immediately superior units (SuperUnit), objectives and types; *RoleList*: maintains the relationships between existing roles in each unit, which roles the unit inherits and what their attributes are (accessibility, position); *NormList*: maintains the relationship between the system rules; *EntityPlayList*: maintains the relationship between the units that register each agent as a member, as well as the role that they play in the unit. Each virtual unit in OVAMAH is defined to represent the "world" for the system in which the agents participate by default. Additionally, the roles are defined in each unit. The roles represent the functionality that is necessary for obtaining the objective of each unit. The PK component directs the basic services on a multi-agent platform and incorporates mechanisms for transporting messages that facilitate the interaction among the various entities.

From a global perspective, the architecture offers a total integration enabling agents to transparently offer and request services from other agents or entities, at the same time allowing external entities to interact with agents in the architecture by using the services provided. The development of MAS is typically based on a design that focuses on each agent independently, and is geared towards each agent's structure and performance. This research presents a new focus in which the design is directed at the organizational aspects of the agents, establishing two descriptive levels: the organization and the agent [4]. The system we developed used the GORMAS [1] organizational methodology.

3 Case of Study: Ambient Intelligence Based Multi-Agent System

ALZ-MAS [8] is a distributed multi-agent system designed upon Ambient Intelligence and aimed at enhancing the assistance and health care for Alzheimer patients living in geriatric residences. In the remainder of this section, the main characteristics

[1] http://www.fipa.org (*Foundation for Intelligent Physical Agents*).

of ALZ-MAS are described, followed by a description of the new ALZ-MAS system developed by means of the OVAMAH architecture. ALZ-MAS structure had five different deliberative agents based on the BDI model (BDI Agents), each one with specific roles and capabilities. The description of the functionality of these agents can be seen in [8]. The MAS was implementing into a geriatric residence to improve health care of the patients. Within the requirements of the problem are: the patient's personal data and behaviour (monitoring, location, daily tasks, and anomalies); the doctors, which treats patients; the nurse schedules, i.e., the nurse's working day obtaining dynamic plans depending on the tasks needed for each assigned patient; the patients' location and manages locks and alarms; and finally, the medical record database and the doctor-patient and nurse-patient assignment.

In previous versions of ALZ-MAS [8], each agent integrates its own functionalities into their structure. If an agent needs to perform a task which involves another agent, it must communicate with that agent to request it. So, if the agent is disengaged, all its functionalities will be unavailable to the rest of agents.

In the version of ALZ-MAS presented in this paper, called OVACARE (*Adaptive Virtual Organizations for Assistance and Health Care*), these mechanisms have been modelled as services in an open agent organization, so any agent can make use of them. The entire ALZ-MAS structure has been modified according to the OVAMAH model, separating most of the agents' functionalities and roles from those to be modelled as services.

OVACARE is a multi-agent organization-based system. The system was designed according to the findings in [1], which apply a MDD (Model Driven Development) focus on organization-oriented methodologies. It is possible to design an organization that is unified, intuitive, and has a high level of abstraction. Given these features, it becomes easier and simpler to design a MAS for organizations and obtain a model for a virtual organization that can be implemented on different platforms. The fundamental idea is to create different models for different levels of abstraction, and then combine them to achieve a full implementation. We used GORMAS (*Guidelines for Organization-based MultiAgent Systems*) [2] as the design methodology and OAVAMAH as the final platform design.

GORMAS is a guide methodology for the design of open MAS from the perspective of human organizations. It includes an analysis phase, a structural organization design phase, and a dynamic organization design phase. Following the guidelines indicated in the methodological guide from [2], one of the first step in analyzing and designing the problem is to define the following roles that will exist within the architecture:

Communicator: in charge of managing the connections that each user makes.

User: in charge of managing the users' personal data and behaviour (monitoring, location, daily tasks, and anomalies). User Agent maintains continuous communication with the rest of the system agents, especially with the ScheduleUser Agent (through which the scheduled-users can communicate the result of their assigned tasks) and with the SuperUser Agent. The User Agent must ensure that all the actions indicated by the SuperUser are carried out, and sends a copy of its memory base (goals and plans) to the Admin Agent in order to maintain backups.

SuperUser: in charge of inserting new tasks into the Admin Agent to be processed by a Case-Based Reasoning mechanism. It also needs to interact with the User Agents to impose new tasks and receive periodic reports, and with the ScheduleUser Agents to ascertain the evolution of each plan. There is one agent for each doctor connected to the system.

Scheduler: in charge of managing a Case-Based Planning (CBP) mechanism embedded in its structure. It schedules the users' daily activities and obtains dynamic plans depending on the tasks needed for each user. It manages scheduled-users profiles (preferences, habits, holidays, etc.), tasks, available time and resources. Every agent generates personalized plans depending on the scheduled-user profile. There is one ScheduleUser Agents for each nurse connected to the system.

Admin: plays two roles: the security role that monitors the users' location and physical building status (temperature, lights, alarms, etc.) through continuous communication with the Devices Agent; and the manager role that handles the databases and the task assignment. It must provide security for the users and ensure the efficiency of the tasks assignments. There is just one Admin Agent running in the system.

Device Manager: makes it possible for the interactive elements within the environment to interact. It deals with devices that use technologies such as RFID (Radio Frequency Identification), etc. It monitors the users' location (continuously obtaining/updating data from sensors), interacts with sensors and actuators to receive information and control physical services (temperature, lights, door locks, alarms, etc.), and also checks the status of the wireless devices connected to the system (e.g. PDA or Laptops). The information obtained is sent to the Admin Agent for processing. This agent runs on a Workstation. There is just one Devices Agent running in the system

Incident Manager: manages and resolves incidents, offers a user location service, and manages an alarm system.

We have also designed an organizational structure. We will first analyze its dimensions, and then proceed to identify the structure that is best suited to apply to the system [2]. Our case study is modeled as a conglomerate (*OVACareUnit*) made up of four units, each one dedicated to one type of functionality within the setting. The five units are: (i) *PlanningUnit*, contains the roles dealing with users' daily activities and dynamic plans: *Scheduler*; (ii) *UserUnit*, contains the roles associated with the user: *Communicator*, *User*, and *SuperUser*; (iii) *AdminUnit*, contains the roles assigned with global management tasks for the health-care system: *Admin*, *Incident Manager*, and *Analyst*;(iv) *DeviceUnit*, which contains the roles associated with the management of devices: *Device Manager*. This role inherits two specialized subroles: *CommApp*, which is responsible for all communications between applications and the platform, and *CommServ*, which is responsible for all communications between services and the platform.

The diagram in Figure 1a provides a structural view of the organizational model, which is adapted according to a conglomerate pattern. Different services are provided within each unit of the organization. In addition, *Provider* and *Client* roles are refined into these new units to specialize in functionality or in the use of specific services associated. The following section defines the services offered by the units, and uses an example to detail each one and how it has been modeled and described in the architecture. The type of role, the inputs and outputs, and a summary of the functionality for each unit are all explained. Figure 1b shows part of the internal model of the *UserUnit*. The internal structure for *OVACareUnit* and the remaining units was modeled in the same way.

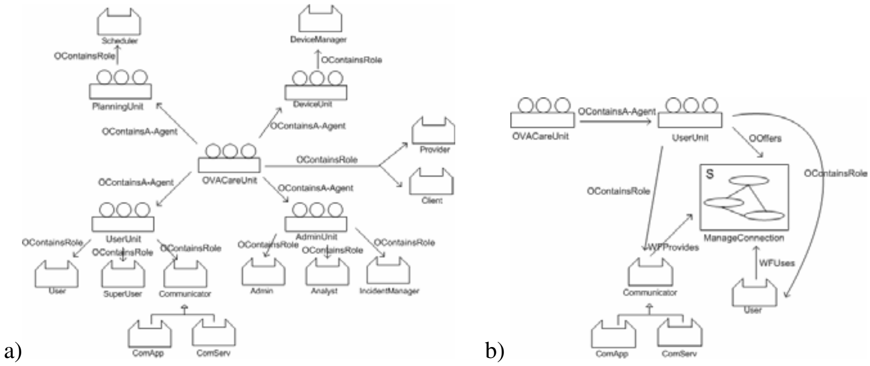


Fig. 1. a) Diagram of organization model: structural view b) Diagram of organization model: functional view of *UserUnit*

One side of the diagram models the functional views of the units, which allows us to identify the services specific to each domain, while the other side precisely details the behavior of the organization services, how they interact with the environment, which interactions are established between the system entities, and how they approach the aspects of an open system. The next step is to define the rules in order to establish the control and management of the services. For example, the basic service provided by *UserUnit* will be *ManageConnection*, which is provided by the agents that take on the role of *Communicator*. The functionalities offered by this service will allow the users to control their connection to the system.

Similarly, within the the services related to *AdminUnit* involve the overall management tasks within a health-care system (e.g., system incidents, data analysis, surveys, user management, notices, etc.). *PlanningUnit* is comprised of services that request guides based on user preferences and certain restrictions (time, specialization of the nurses, etc.). It also includes planning and replanning the route that the nurses will follow based on the suggested plans, and determines the validity and value of the proposed routes. The *DeviceUnit* services deal with the sensors embedded in the physical system (RFID).

The type of services offered is controlled by the system according to the established norms [11]. The internal functionality of the services is responsible for the agents that are offered, but the system is what specifies the agent profiles, as well as the rules to follow for ordering requests or offering results. In this way, when faced with illicit or improper user performance, the system can act to impose sanctions. The OMS will internally save the list of norms that define the role involved, the content of the norms, and the roles in charge of ensuring that the norm is met. We have defined a set of norms in our system for controlling the performance within each unit. This way, for example, an agent within *UserUnit* that acts like *Communicator* is required to register a service as *manageConnection*. If it does not abide by these norms, it will be punished and expelled from the unit. The punishment is logical given that if the agent does not establish a connection within the allocated time, it cannot perform any of the other system tasks. *OBLIGED Communicator REGISTER manageConnection(?requestTime, ?connectionData, ?operation) BEFORE deadline SANCTION (OBLIGED OMS SERVE Expulse (?agentID Comunicator UserUnit))*

Similarly, we have defined a complete set of norms that will control all of the system performances.

Table 1. *ManageConnection* service in *UserUnit*

Service Specification					
Name: ManageConnection **Supplied by:** SF		**Description:** Manages the connection of an user **Required by:** **ClientRole:** User **ProviderRole:** Communicator			
Input Parameters					
Name	**Description**	**Mand.**	**Type**	**Value Range**	**Default**
requestTime connectionData operacion	Time connection Data connection Type of operation on the connection	Yes Yes Yes	date string string		
Output Parameters					
Name	**Description**	**Mand.**	**Type**	**Value Range**	**Default**
connection	Established connection between devices	Yes	connection		
Preconditions and Postcondition: ...					

3.1.1 Example of Service Planning with OVAMAH

The system considers the available time, time to initiate the task, task description, priority of the task, length of the task, and the patient associated with each task, and proposes the optimal route according to the nurse's profile. We will see the series of steps that are taken within the system when a planning route is requested, and how OVAMAH generates the system configuration that will give way to the plan. The first thing is to define the structural components of the organization, that is, the units that will be involved (which are initially empty), the system roles and norms. The indicated service requirements will be registered in the SF. To do so, either the basic OMS services for registering structural components will be used, or the API will directly execute the same functionality. This way, a community type *OVACareUnit* will be created, representing the organization, whose purpose is to control the helathcare system. It has four internal unit planes: *UserUnit*, *PlanningUnit*, *AdminUnit* and *DeviceUnit*, each of which is dedicated to the functionalities we have previously seen. Each unit defines the existing roles, indicating their attributes (visibility, position, etc) and who they inherit them from. The SF will announce basic services that are required for the overall system functionality. The basic services indicate which services are required (according to the defined norms) when creating the units.

From this moment on, the external agents can request the list of existing services and decide whether or not to enter and form part of the organization and with which roles. In our case we have users (nurses, doctors) that use their mobile device to send a request to the system so that it can inform them on the optimal route to take within the system. In order to carry out this function, we have, for example *Co1*, *De1* and *Sc1* acting as agents that will carry out the roles of *Communicator*, *DeviceManager* and *Scheduler* respectively. Agents *U1* and *U2* represent the users that would like to receive a planning route.

Table 2. SF: Basic services

Service Facilitator					
Entity	**Action**	**Service**	**ClientRole**	**ProvRole**	**Profile**
UserUnit	Requires	manageConnection	User	Communicator	UserSP
DeviceUnit	Requires	Locate	Communicator/IncidentManager	DeviceManager	DeviceSP
…	…	…	…	…	…

Initially, all the agents head towards the OVAMAH platform and are associated with the virtual "world" organization. As such, the OMS will play the *member* role in the "world" organization. When SF is asked about existing services in the system, the following response is obtained: `UserUnit Requires manageConnection ClientRole=User;ProvRole=Comunicator;`

Because the service doesn't have an assigned *grounding*, it cannot be requested. But a functionality can be added, thus obtaining the *Communicator* role.

The *Co1* agent wants to offer that functionality, for which it requests receiving the *Communicator* role for the *UserUnit*: `AcquireRole(UserUnit, Communicator)`
If all goes well, the OMS will register *Co1* in the role of *Communicator* in *UserUnit* within the *Entity Play List*. This list shows the roles that the different agents assume within OVAMAH.

The *Co1* agent has carried out all of the regular steps for acquiring a role within OVAMAH. This process is illustrated in Figure 2 where once *Co1* has been registered as a member of the THOMAS platform, it asks SF which defined services have a profile similar to its own "communicator information service". This request is carried out using the SF *SearchService* (message 1), in which *CommunicatorInformationServiceProfile* corresponds to the profile of the *manageConnection* service implemented by *Co1*. The SF returns service identifiers that satisfy these search requirements together with a ranking value for each service (message 2). Ranking value indicates the degree of suitability between a service and a specified service purpose. Then *Co1* executes GetProfile (message 3) in order to obtain detailed information about the *manageConnection* service. Service outputs are "*service goal*" and "*profile*" (message 4). The *manageConnection* profile specifies that service providers have to play a *Communicator* role within *UserUnit*. Thus, *Co1* requests the *AcquireRole* service from the OMS in order to acquire this provider role (message 5). *AcquireRole* service is carried out successfully (message 6), because *UserUnit* is accessible from the virtual organization, thus *Co1* is registered as a *Communicator*. There will be another inquiry regarding which services exist within the units. *AdminUnit*, *PlanningUnit* and *DeviceUnit* will return the services that are necessary for planning. The SF will again return a list (similar to Table 2).

Based on the profiles, we will determine that *Co1* is interested in acquiring the role of *DeviceManager* since in this case it wants to interact with the elements within the environment. *Co1* will use this role to act as intermediary to process the signals that come from the users devices and make them comprehensible to the system. It will allow the order requested by the user from a mobile device to be understood and executed by the specific device that is the object of the request. (`AcquireRole(DeviceUnit, DeviceManager)`).

The agent will now be registered as a member of *DeviceUnit* with the role of *DeviceManager*. This role will require the agent to register the *Locate* service, associating it with the *process* and *grounding* that it considers to most useful. If this is not done within the allocated time, the agent will be expelled. The actual norm is as

Table 3. *Entity Play List*

Entity Play List		
Entity	**Unit**	**Role**
Co1	UserUnit	Communicator
Co1	DeviceUnit	DeviceManager
Sc1	PlanningUnit	Scheduler
U1	AdminUnit	User
U2	AdminUnit	User

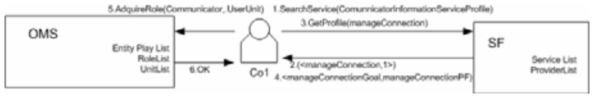


Fig. 2. Agent *Co1* registering

follows: `OBLIGED DeviceManager REGISTER Locate(?route) BEFORE deadline SANCTION (OBLIGED OMS SERVE Expulse (?agentID DeviceManager DeviceUnit))`

The agent will be informed of the norm upon carrying out the "AcquireRole", so that it can take it into consideration if it is a normative agent (otherwise ignore it). To avoid external agents assuming the role of *DeviceManager*, the agent registers a new incompatibility norm in the system. This norm will make it impossible for other agents to take on the same role: `RegisterNorm ("norm1", "FORBIDDEN Member REQUEST AcquireRole Message(CONTENT(role 'DeviceManager))")`

The *De1* and *Sc1* agents will act in a similar fashion, registering at the end for the corresponding units *DeviceManager* and *Scheduler*. They too will be required to register the services as indicated by the defined norms. (*GenerateProfile*, *ConsultProfile*, *UpdateProfile*, *MSState, UpdateMSGState, Replan, ValidateRoute, ValueRoute, TaskListRecovery*) Each one is required for generating the optimal route for the user to follow. The *U1* and *U2* agents will request acquiring the *User* and *SuperUser* roles in order to access the basic services: *FindUserst, GenerateProfile, ConsultProfile, UpdateProfile, MSGState, and UpdateMSGState*.

The agents will also consider whether to acquire other system roles that might be necessary for the required functionality. *U1* can request existing services from the SF, and will receive a list with all the agents that offer their services. The *Entity Play List* would end up as shown in Table 3.

4 Conclusions

An important issue in the development of real open multi-agent systems is to provide developers with methods, tools and appropriate architectures which support all of the requirements of these kinds of systems. Traditional MAS development methodologies are not suitable for developing open MAS because they assume a fixed number of agents that are specified during the system analysis phase. It then becomes necessary to have an infrastructure that can use the concept of agent technology in the development process, and apply decomposition, abstraction and organization methods. We propose a methodology that incorporates decomposition and abstraction via the OVAMAH architecture for a dynamic MAS environment. This architecture has allowed us to directly model the organization of a health-care system according to a previous basic analysis, to dynamically and openly define the agent roles, functionalities and restrictions, and to obtain beforehand the service management capabilities (discovery, directory, etc.). OVAMAH provides us with the level of abstraction necessary for the development of our system, and the set of tools that facilitate its development. OVACARE makes use of OVAMAH distributing resources and enhancing its performance. It is demonstrated that a open approach is adequate to build distributed and highly dynamic Ambient Intelligence based multi-agent systems. In OVAMAH architecture, agents can transparently offer and invoke services from other agents, virtual organizations or entities. Additionally, external entities can interact with agents through the use of the services offered. OVACARE was employed as an illustration of not only the usage of OVAMAH components and services, but also of the dynamics of the applications to be developed with this architecture. In this way, examples of OVAMAH service calls have been shown through several scenarios, along with the evolution of different dynamic virtual organizations.

Acknowledgements. This research has been partially supported through projects JCyL BU006A08, MICINN TIN 2009-13839-C03-03 and MICINN CIT-020000-2008-2.

References

1. Agüero, J., Rebollo, M., Carrascosa, C., Julián, V.: Agent design using model driven development. In: 7th International Conference on Practical Applications of Agents and Multi-Agent Systems (PAAMS 2009), pp. 60–69 (2009)
2. Argente Villaplana, E.: Gormas: Guías para el desarrollo de sistemas multi-agente abiertos basados en organizaciones. Ph.D. thesis, Universidad Politécnica de Valencia (2008)
3. Bernon, C., Cossentino, M., Pavón, J.: Agent Oriented Software Engineering. The Knowledge Engineering Review 20(5), 99–116 (2005)
4. Boissier, O., Hübner, J.F., Sichman, J.S.: Organization Oriented Programming: From Closed to Open Organizations. In: O'Hare, G.M.P., Ricci, A., O'Grady, M.J., Dikenelli, O. (eds.) ESAW 2006. LNCS (LNAI), vol. 4457, pp. 86–105. Springer, Heidelberg (2007)
5. Carbó, J., Molina, J.M., Dávila, J.: Fuzzy Referral based Cooperation in Social Networks of Agents. AI Communications 8(1), 1–13 (2005)
6. Carrascosa, C., Giret, A., Julian, V., Rebollo, M., Argente, E., Botti, V.: Service Oriented MAS: An open architecture. In: Actas del AAMAS 2009 (2009) (in press)
7. Cesta, A., Bahadori, S., Cortellesa, G., Grisetti, G., Giuliani, M., Locchi, L., Leone, G., Nardo, D., Oddi, A., Pecora, F., Rasconi, R., Saggese, A., Scopelliti, M.: The RoboCare Project, Cognitive Systems for the Care of the Elderly. In: Proceedings of International Conference on Aging, Disability and Independence (ICADI), Washington, D.C, USA (2003)
8. Corchado, J.M., Bajo, J., De Paz, Y., Tapia, D.I.: Intelligent Environment for Monitoring Alzheimer Patients, Agent Technology for Health Care. In: Decision Support Systems. Eslevier, Amsterdam (2008b)
9. Decker, K., Li, J.: Coordinated hospital patient scheduling. In: Demazeau, Y. (ed.) Proceedings of ICMAS 1998, Paris, France, vol. 1134, pp. 104–111 (1998)
10. Glez-Bedia, M., Corchado, J.M.: A planning strategy based on variational calculus for deliberative agents. Computing and Information Systems Journal 10(1), 2–14 (2002)
11. GTI-IA. An Abstract Architecture for Virtual Organizations: The THOMAS project, `http://www.fipa.org/docs/THOMASarchitecture.pdf`
12. Dignum, V., Dignum, F.: A landscape of agent systems for the real world. Technical report 44-cs-2006- 061, Institute of Information and Computing Sciences, Utrecht (2006)
13. Lanzola, G., Gatti, L., Falasconi, S., Stefanelli, M.: A framework for building cooperative software agents in medical applications. Artificial Intelligence in Medicine 16(3), 223–249 (1999)
14. Nealon, J., Moreno, A.: Applications of Software Agent Technology in the Health Care domain. Whitestein series in Software Agent Technologies. Birkhauser, Basel (2003)

Talking Agents in Ambient-Assisted Living

José M. Fernández de Alba and Juan Pavón

Universidad Complutense de Madrid, Facultad de Informática,
28040 Madrid, Spain
jmfernandezdalba@gmail.com,
jpavon@fdi.ucm.es

Abstract. A natural way to interact with persons in Ambient Asssited Living (AAL) scenarios is through speech. Speech recognition techniques usually require of training for achieving good results. However, this training is not possible when different people can enter and leave. In these scenarios, the system should get information from the context, such as the state of the environment or previous dialogues, to support the speech recognition process. This is made in this work with Talking Agents. These are autonomous entities, with the ability to interact with each other and with other agents in the environment, and with the responsibility of interacting with persons through speech. They have a modular architecture to facilitate the application of different mechanisms and strategies, depending on the concrete AAL scenario.

1 Introduction

In this this work, are shown reusable components for building distributed multi-agent systems (MAS) in Ambient Assisted Living (AAL) scenarios (e.g., [2] and [7]) where main interaction with users occurs through speech. There are different types of agents in this kind of MAS, and here the focus is on those taking care of speech processing, which are called Talking Agents. A Talking Agent has the ability to manage interactions with users, specially through speech but considering other media as well, and adapting its behavior to the concrete context. This has required the integration of a diversity of software packages for speech recognition and synthesis, as well as coordination of heterogeneous elements in varying contexts.

Talking Agents are reusable software components, with autonomous and social behavior. They have the ability to interact with humans using speech or other communication channels, and work in a distributed computing environment where they can communicate with other agents and find resources. In most scenarios, Talking Agents use heterogeneous resources, which vary depending on the needs of each particular deployment.

Talking Agents running on a system are organized as a society of agents that may cooperate and interact among themselves. Depending on the situation, the user will create an interaction one by one with each Talking Agent, or with several at the same time. The communication among Talking Agents help them

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 328–336, 2010.

to get more knowledge about the context and help each other to better manage the interaction with the user. This idea of working with the context is common in AAL (see, for instance, [9]) and has been started to be applied in related fields [5]. Also, each Talking Agent may implement a different strategy for managing interactions, and the resulting behavior will emerge from the combination of the responses given by the Talking Agents participating in each concrete AAL scenario.

With this purpose, several agent frameworks have been considered to build the architecture, finally adopting ICARO [8], an open source project that promotes the development of distributed applications as organizations of agents and resources. With respect to more popular agent frameworks (e.g. Jade, Jack), ICARO has the advantage of promoting an organization-based view of the multi-agent system, being flexible in the communication mechanisms and protocols among agents (not forcing, for instance, to use FIPA standards).

The rest of the paper is structured as follows. Section 2 discusses the Talking Agents interaction model, which takes many concepts from Dialog Systems. Section 3 presents the multi-agent architecture view. The last section presents some conclusions and future lines of work.

2 Talking-Agents Architecture

Talking Agents are not isolated entities, as they collaborate as a multi-agent system. The organization of the multi-agent system determines the global system architecture, which is described in section 3. Also, each individual agent has its own architecture, with well-defined component interfaces. This section describes the architecture of the Talking Agent, showing the relationships among its components.

Figure 1 depicts a diagram showing an example snapshot of the components being used by a single Talking Agent as a conversational agent: perception, interpretation, planning and execution components, and the Talking Agent core. These concepts relates with the traditional Dialog Systems, as in [3]. Each component implements a well-defined generic interface, depending on the dialog process activity it performs, thus enabling the Talking Agent to manage an interaction of any kind (multimodal), in a generic manner, by customizing the respective components. The Talking Agent core holds both the responsibility of deciding (using a rule engine) the appropriate intention given a certain knowledge, and the responsibility of coordination of the available resources, using a reactive control model (from ICARO platform, as described later).

2.1 Perception

A perception component is a thread that generates events with some information, either when it detects certain situations for which it is prepared or in a periodic manner, depending on the type of stimulators and the way it expects to obtain the information.

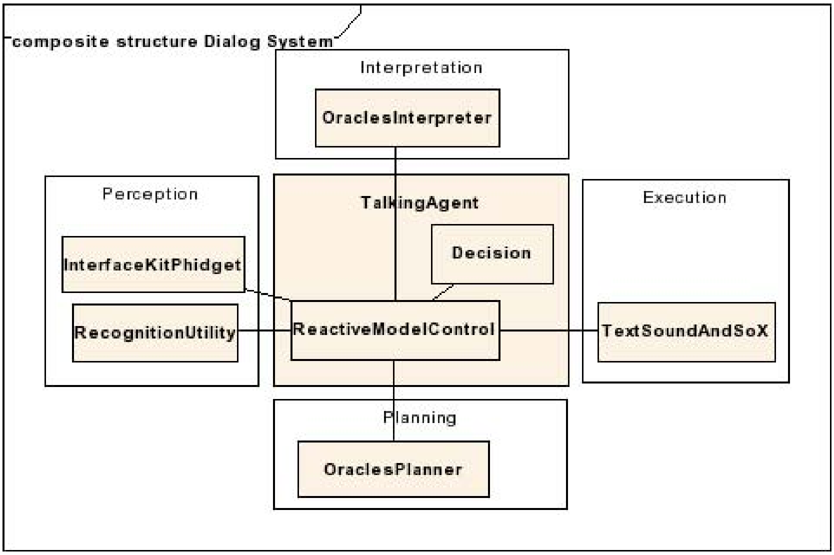


Fig. 1. Talking Agent components snapshot example

Any perception component must implement an interface that has the following methods (only the main operations are shown):

– **AddSubscription(Subscription):** adds a subscription object to the component. This object contains the information necessary for the component to notify a subscriber when a change in certain subject in the context of the perception component has occurred.
– **RemoveSubscription(SubscriptionId):** removes a subscription object on the component.
– **NotifySubscriptors(Subject, Info):** notifies the subscribers when a change has been perceived in certain subject, with an associated information, by sending the AcceptPerceptionEvent signal to the subscribers, as will be described later.

In Figure 2, two Talking Agents subscribe to a perception component. Later, when the components receive and process a speech, they notify the event to the subscribed agents.

2.2 Interpretation

The interpretation elements take the events generated by the perception elements and try to extract the meaning of the information.

The result of the interpretation is a certain amount of knowledge the agent can use in its decision process. The knowledge is represented in a symbolic manner so that it can be processed by a reasoning engine, like a rule-based reasoning engine.

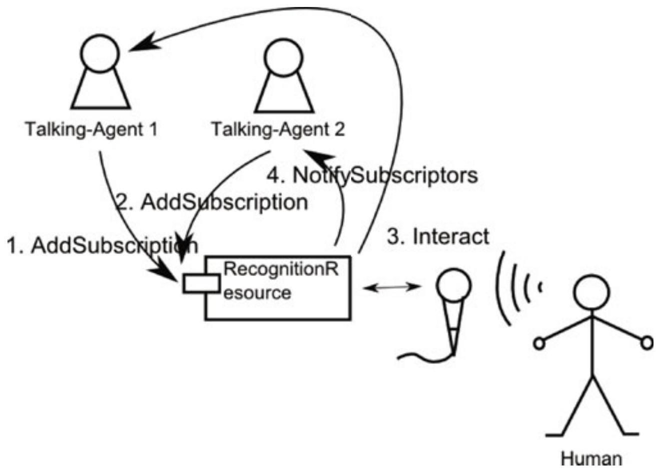


Fig. 2. Perception event subscription

An interpretation element is composed of various heterogeneous elements, in order to be able to interpret the different perception events on the system. For example, considering the language-based kind of interaction, the Natural Language Understanding technology may be ised, so that the speech events can be interpreted by its meaning in certain domain (for instance, obtaining the topic of the query [11]).

Any interpretation component must implement an interface that has the following methods:

- **UpdateMentalState(NewMentalState, AgentId):** updates the mental state associated with an agent in the interpreter component, in order to include the last perceived events and other available knowledge, and triggers a new interpretation based on the new information.
- **SendInterpretation(NewMentalState, AgentId):** sends the UpdateMentalState signal to the agent, if the interpretation process has changed the MentalState object in result, or the NoMentalStateChange signal otherwise.

2.3 Decision

This is the part in which the agent coordinates its managed modules and its functional cycle and decides which immediate objectives to pursue, depending on the knowledge of its own state and the environment (the agent's beliefs). The decision is performed at two different levels: control and intention.

At the *control* level, the agent only decides which phase of the information flow is going to process next. This decision process is modeled as a state machine following the reactive agent behavior model (from the ICARO framework), which controls each of the functions performed by the agent at every moment. This kind of state machine reacts to certain received signals, which are generated by the resources, indicating the completion of each processing. Although it is recommended to use the provided state machine in order to implement the Talking Agent behavior, one may modify it to address certain concrete requirements.

The signals accepted by this state machine are the following:

– **AcceptPerceptionEvent(PerceptionEvent[]):** in the perception phase, includes the new perception event in the events buffer of its MentalState object and calls the UpdateMentalState method of its interpretation component. Then changes to the interpretation phase.
– **AcceptInterpretation(NewMentalState):** in the interpretation phase, updates the mental state of the agent and lets the intention level run, as we will see later. Then changes to the decision phase.
– **DecisionFinalized:** in the decision phase, this signal indicate that the rule engine has ended its processing. The UpdateMentalState method of the planning components is called. Then changes to the planning phase.
– **AcceptPlan(Action):** in the planning phase, includes the new action to the actions buffer of its MentalState object and then calls the Execute method of the action, as we will se later. Then changes to the execution phase.
– **ActionFinalized:** in the execution phase, this signal indicates that the action has been executed. The ObtainResults method of the action is called, updating the mental state. Then changes to perception phase.
– **NoMentalStateChange:** in any phase, changes to the perception phase.

At the *intention* level, the agent decides its next immediate objectives, or intentions. This is implemented using the Drools [1] rule engine, which produces a set of intentions based on a set of input beliefs, corresponding to the specified rules that models the agent behavior. With this formalization, the Belief objects, obtained from interpretations, and distinguished by its code field, are put into the when part, thus triggering the creation of certain Intention objects, which will determine the next plans.

After each execution of the rule engine, a DecisionFinalized signal is sent to the state machine of the agent it there have been any change on the mental state, or a NoMentalStateChange signal otherwise.

2.4 Planning

The planning elements take the intentions generated in previous phase and create a set of concrete actions to be done.

Basically, the intentions are broken down into actions through several iterations, starting from the more abstract, generic actions to the more specific, concrete actions, which will trigger operations on the associated execution resources.

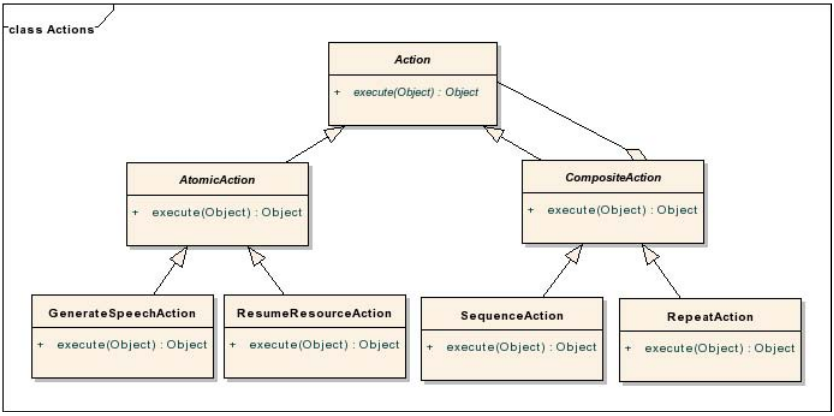


Fig. 3. Actions hierarchy

Figure 3 shows the hierarchy of actions generated by the planning resource.

Any planning component must implement an interface with the following methods:

- **UpdateMentalState(NewMentalState, AgentId):** updates the mental state associated with an agent in the planner component, in order to include the last perceived events and other available knowledge, and triggers a new interpretation based on the new information.
- **SendPlanning(NewMentalState, AgentId):** generates a list of actions associated with the specified agent's mental state and sends the results to the agent.

2.5 Execution

The execution phase is performed by the agent by executing the actions generated in the previous phase. Figure 4 shows an example of execution.

Each Action class holds the necessary information to perform the desired operation on a execution component, which may implement any arbitrary interface, as long as the planning components are able to create the corresponding actions to execute them.

3 Multi-Agent System Architecture

Talking Agents collaborate among themselves and with other agents in the ICARO agent framework [8], which provides communication and management services that facilitate the distribution, configuration and monitoring of the multi-agent system. The global architecture arises from the collaboration of the different agents while they make their own processing as described before.

Figure 5 shows an example of deployment of a Talking System, using different resources and communicating among them.

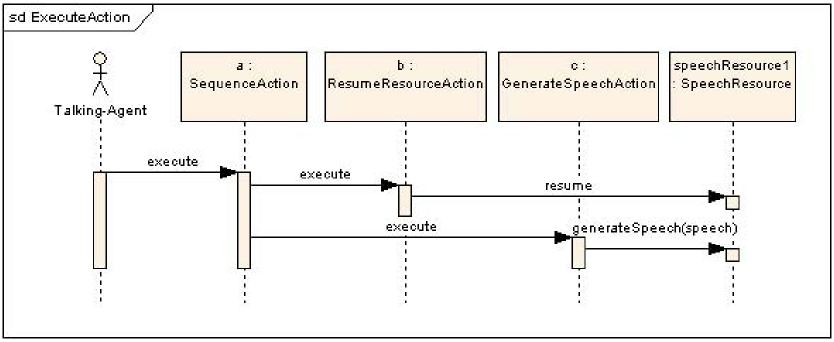


Fig. 4. Example of execution of action

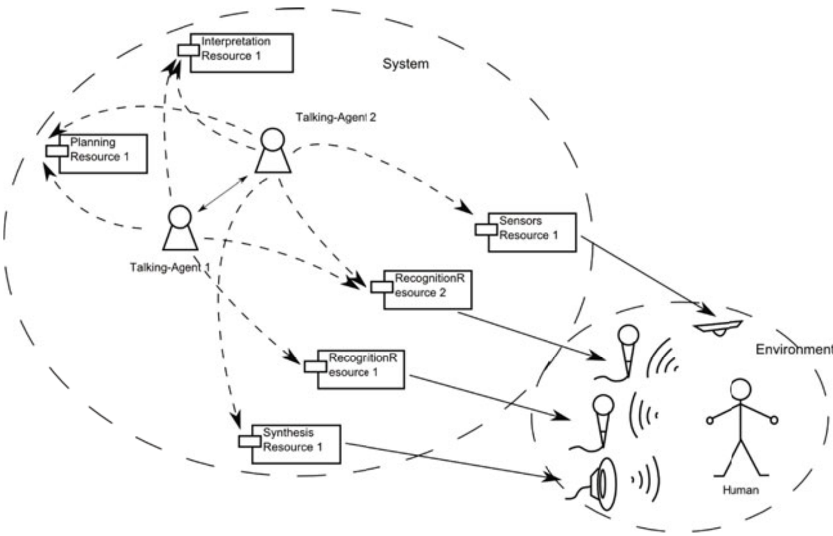


Fig. 5. Talking-Agents system deployment example

3.1 Inter-agent Communication Model

The inter-agent communication capability is the main feature that distinguishes the multi-agent architecture view from the single-agent architecture view. This feature has been integrated into the latter in the following way: any agent is able to execute an action to send an AcceptPerceptionEvent signal to another agent, as if it was a perception resource. This way, the receiver agent is able to process the communication the same manner it process any other perception event from the environment. The concrete protocols used to establish collaborations are implemented when defining a concrete system using this principle.

4 Conclusions and Future Work

According to the definitions found in literature about AAL (like [10]), there exists common "enabling technologies" that should be integrated in order to develop systems of this kind. These technologies are divided in different functions: sensing, reasoning, acting, interacting and communicating. The result should be an electronic environment, sensitive and responsive to the presence of people. The architecture described in this paper manages to integrate all these functions in an abstract processing sequence, controlled by agents, which is able to manage many different modes of interaction. In addition, the use of multi-agent system concept permits the creation of more flexible systems, which make use of social emergent behavior, and component distribution.

One important issue for further work relates with real-time processing [6]. Since this kind of systems usually requires immediate answers, we are working in a set of agent collaboration protocols in order to maximize parallelization and task division. This collaboration should be developed in a robust and flexible manner to allow the system to grow and evolve without compromising the system performance.

Acknowledgements. This work has been done in the context of the project Agent-based Modeling and Simulation of Complex Social Systems (SiCoSSys), supported by Spanish Council for Science and Innovation, with grant TIN2008-06464-C03-01. In addition, we acknowledge support from the Programa de Creacin y Consolidacin de Grupos de Investigacin UCM-Banco Santander with reference GR58/08.

References

1. Bali, M.: Drools JBoss Rules 5.0 Developer's Guide, ed: Packt (2009)
2. Botía, J., et al.: Detecting Domestic Problems of Elderly People: Simple and Unobstrusive Sensors to Generate the Context of the Attended. In: Omatu, S., Rocha, M.P., Bravo, J., Fernández, F., Corchado, E., Bustillo, A., Corchado, J.M. (eds.) IWANN 2009. LNCS, vol. 5518, pp. 819–826. Springer, Heidelberg (2009)
3. Eckert, W., Kuhn, T., Niemann, H., Rieck, S., Scheuer, A., Schujat-takamazzini, E.G.: A spoken dialogue system for german intercity train timetable inquiries. In: Proc. European Conference on Speech Technology, pp. 1871–1874 (1993)
4. Fernández de Alba, J.M., Pavón, J.: Talking Agents Design on the ICARO Framework. In: Corchado, E., Yin, H. (eds.) IDEAL 2009. LNCS, vol. 5788, pp. 494–501. Springer, Heidelberg (2009)
5. Griol, D., Sánchez-Pi, N., Carbó, J., Molina, J.M.: A Context-Aware Architecture to Provide Adaptive Services by means of Spoken Dialogue Interaction. In: International Conference on Artificial Intelligence (ICAI 2009), in WORLDCOM 2009, vol. II, pp. 912–918 (2009)
6. Julian, V., Botti, V.: Developing real-time multi-agent sys-tems. Integrated Computer-Aided Engineering 11(2), 135–149 (2004)
7. López, M., Fernández-Caballero, A., Fernández, M.A., Mira, J., Delgado, A.: Visual surveillance by dynamic visual attention method. Pattern Recognition 39(11), 2194–2211 (2006)

8. Morfeo Community, ICARO Project Website: http://icaro.morfeo-project.org
9. Nieto-Carvajal, I., Botía, J., Gómez-Skarmeta, A.: Information and Hybrid Architecture Model of the OCP Contextual Information Management System. J. UCS 12(3), 357–366 (2006)
10. The European Ambient Assisted Living Innovation Alliance (AALIANCE Project), Ambient-Assisted Living Roadmap: http://www.aaliance.eu/public/documents/aaliance-roadmap/
11. Yang, Y., Liu, X.: A Re-Examination of Text Categorization Methods. In: Proc: ACM SIGIR Conference on Research and Development in Information Retrieval - SIGIR 1999, pp. 42–49. ACM Press, New York (1999)

A System for Recognizing Activities of Daily Living Using Everyday Objects

María Ros, Miguel Delgado, and Amparo Vila

Department of Computer Science and Artificial Intelligence
University of Granada, Granada, 18071
{marosiz,mdelgado,vila}@decsai.ugr.es
http://decsai.ugr.es

Abstract. The population is quickly ageing. It is estimated that 25 % of the European population will be made up of people aged over 65 [9]. This ageing provokes that the government has to provide more resources to manage the elderly requirements. However, its capacity for facing them is not strong enough. In addition, there is other essential problem: the elders' rejection to move to a nursing house or to be continuously monitored. The objective is to design a model that is able to combine both interests: the independence of living in their home and the security of living in an elderly people home. We propose a system which tries to monitor user daily activities in order to provide services, currently in an alarm way, and to support the caregiver works. The system uses everyday objects to know what object the user is touching, by means of sensors, and hand-held devices to collect all the information and inform the user (or their families or even their medical assistance centre).

Keywords: Activities of Daily Living (ADLs), RFID, Data Mining, Regular Grammar, Fuzzy Logic, Temporal Activities.

1 Introduction

Population age is increasing, and as a result, their health status is deteriorated. Because of our age's population is high, the resources and services should be increased at the same pace. However this is not happening. Therefore, we need to find other ways to cover their necessities. A common alternative is to force them to leave their homes and their lifestyle, and to move to a nursing home. In these places, they are continuously subject to nurse's care and the services are assembled. However, this change provokes some troubles for them, who want to remain at home as long as possible. The great majority of them does not admit their constraints and turns any help down, due to their wish of keeping on with their life as they are at present. For that, this forced relocation usually causes mental deterioration, physical decline and even premature death [13].

Ambient Assisted Living tries to avoid these troubles including the assistance in their own homes. Among its objectives, we could emphasizes the aim of providing to older adults with a *safe independence* as long as they are healthy

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 337–346, 2010.

enough to live alone. When we say *safe independence*, we refer to an environment which is capable to provide all necessary services and to monitor what the user is doing in every moment. When something strange happens, such as the user has fallen out or has forgotten the burner on, the system warns the nurses or puts the burner out. In short, the system has to avoid the dangerous situation. These systems do not try to replace caregivers but let them focus their attention on care-giving and not on the daily activities. The ADL monitoring is an important task in health care, because they must record and report the ADLs in order to get Medicare funds. The caregivers spend considerable time measuring and tracking ADL [12]. So, designing and implementing automated aids that can improve caregiver work practice is the following step.

In the existing literature about this subject, we can find several projects whose objective consists of inferring user's current activities, but using different methods to extract the information. In [12], the authors propose a system, called PROACT (Proactive Activity Toolkit) which is able to represent activities as a probabilistic sequence of used objects, to use RFID technology to sense the objects being used across various sensing and use contexts, and to mine probabilistic models from descriptions of activities. On the other hand, in [10], Naeem et al. describe a tiered approach to deal with recognition of activities, paying attention to the missing sensor events. They design a *Hierarchy of Activities of Daily Life (HADL)* where they decompose de ADLS into different levels regarding its size and importance. However, these both systems have to get the data from other resources, such as textual descriptions. Others projects based their strategy on extracting the common pattern for an activity and recognizing when something different happens. In [13], Rantz et al. propose a method based on temporal clustering to detect a gradual change as a result of a deteriorating condition. This proposal is included in the TigerPlace project [13] where the authors propose an educational and research environment based on the Aging in Place concept. This project involves many different academic disciplines as social work, physical therapy, environmental design, health informatics, among others. In [4] eight projects about Health care Systems and Other applications are summarized. Among the European projects, we emphasize the *Cogknow*[11][1] a project to help mild dementia people navigate using contextual information and wireless and mobile device technologies, including location-based services or the *Context Awareness in Residence for elders* [8][2] developed in the University of Amsterdam which let elders retain their independence by living at home longer. Using sensors such as switches or pressure mats, they monitor inhabitants' behaviour unobtrusively, benefiting both families and the elders.

Focusing the study on the system to infer the general *human behaviour*, the current trend is the use of probabilistic models. They have emerged as an effective tool for modelling behaviours, thanks to the advantage of combining prior knowledge of a certain situation with the evidence observed form a variety of sensors [1]. Several works in this area are based on HMMs (and their extension to hierarchical

[1] `http://www.cogknow.eu/`
[2] `www.science.uva.nl/~tlmkaste`

models). One such approach is presented by Wilson et al. in [18] where they use a Viterbi algorithm to execute and analyse the episode recovery experiments. This approach obtains great outcomes in the unsupervised task identification. Problems appear when the actions are executed in any order. For instance, Naeem et al. in [10] base their approach on *Multiple Behavioural Hidden Markov Models.* They create multiple hidden Markov models for every variation of an action. One of the advantages of this approach is that, even if the elderly person has no finished completely the behaviour, it is still possible for the MBHMM to determine which task is currently active. In [10], we can find an extensive comparison of both models to task identification in the Home Environment.

Opposite of these approaches, models based on fuzzy logic[3] or data mining[6] have emerged. For example, the approach presented in [3], where the authors describe a novel life-long learning approach for intelligent agents that are embedded in intelligent environment. They propose a method to extract fuzzy membership functions and rules that represent the user's particularized behaviours in the environment. The Guralnik and Haigh's approach uses sequential patterns learning to learn models of human behaviour and incorporate reasoning about the time of the activities.

In this paper, we propose a system to monitor the Activities of Daily Living (ADLs) and to provide services when dangerous situations happen. Our objective is to maintain the independence of elders in a safely way. Our system is embedded in their environment, in their own home. In this way, we neither bother them nor force to change their normal life. They continue with their life, doing the same things that they used to do, and in the same way. The system learns what the user does in every moment and identifies which the common behaviours are. It uses sensors, and specifically RFID technology, to collect user information, and a hand-held device to inform the right person (the user, the family or even the medical assistance centre) when something is wrong. Our objective is not to inform about every user life's moment but dangerous situations.

This paper is organized in four sections. In section 2 we present the proposed system in detail. In the following section, we explain the selected implementation based on RFID technology and hand-held devices. Finally, the conclusions and future projects are presented.

2 A System to Recognize ADLs in a Smart Environment

In depth, the system consists of two parts: one based on Data Mining techniques and other based on Regular Grammar concept, together with a specific temporal activities processing. The temporal study is realized through the Fuzzy Temporal Window concept [16] [15]. The system uses the information collected by the sensors to study behaviours through a process based on Frequent Itemsets [2]. Once we have the knowledge about the common behaviours, the next step is to provide services. Currently, our services are alarms that are sent to a specific person. This part is based on Regular Grammar concepts, structuring the information as a tree, called as Temporal Behaviour Tree [2]. We should not forget the temporal activities knowledge, using it to get better results.

For the work in this paper, behaviours have been modelled as a sequence of actions. We define an action as an activity (fact) that happens using a specific object. However, the order relationship between the elements in the sequence is as important as the actions, because it determines when a simple sequence could be considered as a right behaviour.

2.1 Extracting Behaviour Patterns

Our system uses Data Mining techniques, specifically *Frequent Itemsets*, to extract behaviour patterns from a database. Our idea is based on the fact that in general, the most common actions are always the most frequent actions that we perform to achieve our objective. See example 1.

Example 1. Let us suppose that we want to study `to leave home` behaviour. When we observe the user, we could detect some actions that they always perform: `to take the keys` and `to close the door`. They might execute neither the actions in the same way or order nor the same actions every time (they could do other actions too). However, these two actions are essential to finish the behaviour: without keys we could not close the door and if we did not close the door, we could be robbed.

Therefore, we identify *frequent itemsets* with the common sequence of actions of a specific behaviour. This method is applied over a database which has two constraints. First of all, the database has to be represented as a transactional database in order to be able to apply the Data Mining method. Secondly it collects what the user performs during the whole day. This is too much information to apply the method directly. So, we delimit a piece of the database where we believe we will find the sequences related to a specific behaviour.

In order to achieve this objective, we introduce the temporal variable in our system. If we observe daily activities, they are usually carried out in a specific time. See example 2:

Example 2. Rose leaves home at 7:30 to go to work.

In general, there are activities that almost always happen at the same time, and even more when the user is an older adult. However, the time is not exact, but it is vague. Our method pays attention to the time user activities in order to manage this vagueness. See example 3.

Example 3. Let us suppose that `Rose leaves home at 7:30`. She is generally going to leave home about 7:30, sometimes before and sometimes later. Then, we should supervise actions that happen in an interval "about 7:30", for instance from 7:20 to 7:40.

To handle this imprecision, we propose a structure based on temporal constraints that establishes the temporal interval where we should search the behaviour pattern. This is called as *Fuzzy Temporal Window* [15][16].

The *Fuzzy Temporal Window* is represented as a fuzzy set, defined generally as a Gauss bell curve shape. It is also possible to use a trapezoidal or triangular shape. To delimit the database, we establish the centre of the fuzzy set (the mean) on the specific moment related to the behaviour, whereas its size is determined by the deviation of the fuzzy set. Using a fuzzy set, we can attach an importance to every action in the temporal interval: the actions closer to the temporal knowledge attach more importance than the others. The outcome of applying this structure over the *transactional database* is a *fuzzy transactional database*.

The *Fuzzy Temporal Window* is an essential structure in our system, not only is it useful to extract the behaviour, but it is essential in the recognizing process, as we see later. For more details about this structure, please see [15][16].

We use the *Fuzzy Temporal Window* in order to delimit the database where we will search the sequence of actions. This process is based on the Frequent Itemsets search. However, not only do we extract itemsets but we need a sequence of actions that is correct according to a specific behaviour meaning. The extraction process has two main points [2]:

1. Extract common actions for a specific behaviour, identifying them with the frequent actions.
2. Extract those permutations of the extracted itemset that keep the order relationship among the actions, regarding the started database.

In [2] we present this method as a crisp model. However, after applying the FTW, every object has associated a fuzzy value according to its position in the FTW. The outcome of this process is a fuzzy transactional database. Therefore, we have to transform the data from fuzzy data to crisp one to apply the method proposed in [2]. We use the α-cut concept to represent the fuzzy database on several crisp database, one for every different value and the *Decomposition Theorem* to join the behaviour patterns in only one.

Using the α-cuts over a fuzzy transactional database, we can obtain a new transactional database. Those items whose membership degree was bigger or equal than the defined α-cut value are studied (the membership degree will be 1), whereas the ones whose membership degree was lower than it are avoided (the membership degree will be 0). This process is repeated with every α-cut value. In this point, we will have different patterns for the same behaviour depending on the used α-cut value. To combine all of the patterns, we apply the *Decomposition Theorem* and get a unique pattern, where the α-cut value is used as an importance degree. For more details about this process, please see [15][16].

2.2 Recognizing Behaviour Patterns

Once, we have extracted patterns that represent a specific behaviour, the next step consists of using this knowledge to recognize the user activities and control when the user is correctly doing the actions. As it happens in the previous section, we have to deal with the temporal information.

The recognizing system has two main points[17]: a support structure, called Temporal Behaviour Tree (TBT), and a designed algorithm, which is able to manage the temporal information, control the user's currently sequence of actions, and all possible behaviours.

This TBT is inspired in automata, the structure used to recognize a Regular Grammar. We identify a *word* of the *Regular Grammars* with a sequence of actions and the alphabet with the possible actions. Thus, the process of recognizing the user activities consists of accepting a Grammar word, matching the actions with the characters that compose the word. However, we cannot use an automaton to recognize our behaviours, so we have to develop other structure. In [2] we present an extensive study to justify our choice.

Our structure is a tree where intermediate nodes symbolize the actions, whereas leaf nodes identify the final behaviour. In addition to the sequence of actions, the TBT stores the FTW used to extract the patterns. That let us process the temporal information and discard the actions out of time.

To build the tree, we use the knowledge extracted in the previous step. However, the patterns have some parameters that will determine the results and its reliability, such as the minimum support and the $\alpha-$cut value. This last parameter will be used as threshold by the algorithm to determine when the user has carried a specific behaviour out.

Together with this structure, we define an algorithm that is able to detect not only what behaviour could be, but also what actions the user has omitted. A user hardly ever performs a behaviour at the same way, but they usually do it in different order, forget actions (that could be key ones or not) or insert other actions between the current behaviour's actions. All these possibilities should be taken into consideration to get a high reliability rate.

The algorithm is based on the tree walking process and the possibility of having several alternative paths. We walk on the tree in function of the input, comparing it with a current action from the studied path. If the algorithm finds a match between their children or grandchildren, the system updates this path and waits for the next input. On the other hand, if the algorithm finds a final node, then the system indicates if the behaviour is done correctly or if the user has forgotten a key action.

In addition, the walk on the tree requires a temporal process, since we have to take into consideration if the user has carried out more than one action from a specific behaviour in the correct time. Let suppose that the algorithm is walking on a branch related to a specific behaviour, then it is natural that the more deep the process arrives the more likely the user is doing this behaviour. When an input value coincides with a node, the algorithm evaluates its current instant of time with the FW related to the studied node. If the obtained evaluation value is bigger than 0, this value will be borne in mind in the next execution. For the next input, if it is a child (or grandchild) of the previous node, its temporal evaluation is the aggregation of the obtained value for the current time and the accumulated evaluations of the previous nodes. To make the aggregation, we could use any Fuzzy logic aggregation operator. In this case, we choose the OWA operator [19]

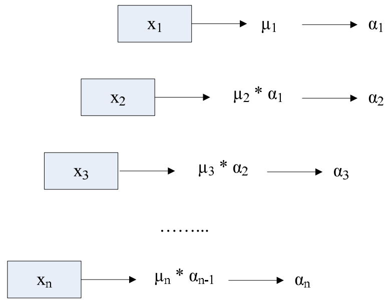


Fig. 1. Temporal information processing over a tree branch

due to the facilities to provide different importance degree to the accumulated value and the current evaluation(See Figure 1). This accumulated value is used to determine if the behaviour is being performed or not. We compare this value with a threshold, determined by the $\alpha-$value used to select behaviour patterns. If this value is bigger than the threshold, then the system confirms that the behaviour is correct, in other case, the system rejects the studied behaviour.

Apart from those details, the algorithm uses a *Sliding Windows* to attach more importance to the recent actions than the distant actions. We should establish a temporal size with the aim of restricting the process between the current moment and the past moment. This last moment will be defined as the current moment minus the fixed windows size. According to the time, the *Sliding Window* is scrolled, discarding some data and including others. The actions that composed a behaviour usually have a temporal relationship among them. For this reason, the fixed size for the used *Sliding Window* will determine the system results.

3 Current Architecture: Experimentation

The implementation of this kind of systems follows a basic principle: its implementation must not affect the user's daily life. In other words, the users may not be aware of the system, since the system should be invisible to them and their habits. This aspect gains special prominence when we deal with elderly people, adversed to any changes in their life. For that, in the implementation of these systems it is necessary to pay attention to used hardware in order to minimize the changes over the environment.

First of all, we have to choose a specific hardware to collect the information that the system uses to extract the *pattern behaviours*. We have used RFID tags, as sensors, to identify every object in the environment. These sensors let us identify every object when the user is touching them. The tags have to be read using a specific reader. In our implementation, we have used an SD reader, which is connected in a PDA. In addition, we require a server to control everything that happens in the environment, to do main reasoning and to be connected. Please see figure 2.

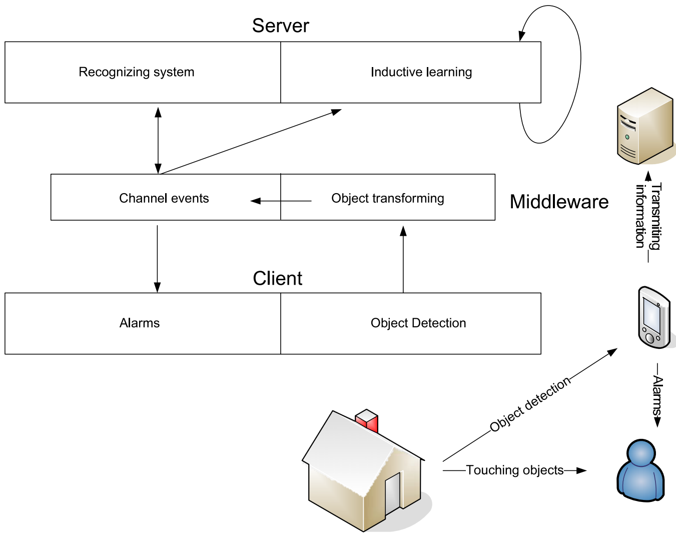


Fig. 2. Simple Diagram of the System Architecture

In this point, we should clarify that we have chosen this technology due to its price and its simplicity to use. We are not actually interested in the sensor network. Our objective is to design a system that is able to receive a sequence of objects and to extract adequate conclusions. However, we are not interested in what sensor has gotten the information, but its related object.

The system is implemented as a Client/Server architecture. The Client is generated as a Client mobile whose aim is to be an intermediary between the Recognizing process (setting up on the server) and the user. It is responsible for reading the passive tags, for sending the read tags to the server, for receiving the results and for sending the alarms, whenever necessary.

As far as the Server's concerned, it implements the Recognizing Process. This has to store two databases: the Behaviour Database, which is the database that collects all the behaviour patterns obtained in the Learning process, and the Object Database, which collects the identification number of objects (RFID identification number) and the position in the environment.

The system has been implemented using JAVA technology [5], to be precise JAVA2 and JAVA Web Service for the server, and JAVAME for the Client mobile. For both the Server and the Client, we have developed a visual interface to see the reasoning plan and alarms, respectively. The alarms are composed by extracted behaviour and services when corresponds, i.e., when user has forgotten something or when the system remind him to do something. Apart from that, we use a middleware, to be precise Zero C Ice [3] [7], to abstract the system from the type of inputs besides to implement the channel communication between client and server.

[3] `http://www.zeroc.com/download/Ice/3.4/Ice-3.4.0.pdf` the lasted access on 03/16/2010.

In order to test our system, we have developed a simulation about some rooms in a house, specifically, a kitchen, an entrance hall and a bedroom. We have collected data for several days and created a behaviour data base with two main behaviours: *to go to the bed* and *to leave home*. To show the obtained results, we present a video that can be visualized in `http://decsai.ugr.es/ marosiz/ Laboratory_Video.wmv`.

4 Conclusions and Future Projects

In this paper, we have presented a system to recognize the Activities of Daily Living. Our objective is to monitor the user activities to help the user, knowing if they are in a dangerous or wrong situation and provide services in this case. In depth, the system is based on Data Mining and Regular Grammar concepts. Nevertheless, the temporal processing is basic in this system. We use this knowledge to extract the behaviour patterns and to recognize the activities. To achieve this objective, we introduce a new structure, called as *Temporal Fuzzy Windows*, which lets us fix fuzzy intervals over the databases.

Besides the theoretical part of the system, we present a possible implementation based on RFID and hand-held devices. The system collects the user activities using their interaction with the objects identified by a RFID tag.

As future project, we want to include semantic to the system, organizing the behaviours as a hierarchical or an ontology in order to abstract the concepts. Furthermore, the laboratory is being extended to study all possible situations and a huge number of behaviours.

References

1. Atallah, L., Yang, G.: Review: The use of pervasive sensing for behaviour profiling - a survey. Pervasive Mob. Comput. 5(5), 447–464 (2009)
2. Delgado, M., Ros, M., Amparo Vila, M.: Correct behavior identification system in a Tagged World. Expert Syst. Appl. 36(6), 9899–9906 (2009), `http://dx.doi.org/10.1016/j.eswa.2009.01.077`
3. Doctor, F., Hagras, H., Callaghan, V.: A fuzzy embedded agent-based approach for realizing ambient intelligence in intelligent inhabited environments. IEEE Transactions on Systems, Man, and Cybernetics, Part A 35(1), 55–65 (2005)
4. Droes, R., Mulvenna, M., Nugent, C., Finlay, D., Donnelly, M., Mikalsen, M., Walderhaug, S., Kasteren, T.V., Krose, B., Puglia, S., Scanu, F., Migliori, M.O., Ucar, E., Atlig, C., Kilicaslan, Y., Ucar, O., Hou, J.: Healthcare Systems and Other Applications. IEEE Pervasive Computing 6(1), 59–63 (2007), `http://dx.doi.org/10.1109/MPRV.2007.12`
5. Eckel, B.: Thinking in Java: the Definitive Introduction to Object-Oriented Programming in the Language of the World Wide Web. Prentice-Hall, Inc., Englewood Cliffs (1998)
6. Guralnik, V., Haigh, K.Z.: Learning Models of Human Behaviour with Sequential Patterns. In: Proceedings of the AAAI 2002 Workshop "Automation as Caregiver", pp. 24–30 (2002)

7. Henning, M.: A New Approach to Object-Oriented Middleware. IEEE Internet Computing 8(1), 66–75 (2004), http://dx.doi.org/10.1109/MIC.2004.1260706
8. Kröse, B., van Kasteren, T., Gibson, C., van den Dool, T.: CARE: Context Awareness in Residences for Elderly. In: Proceedings of ISG 2008: The 6th International Conference of the International Society for Gerontechnology, pp. 101–105 (2008)
9. Mann, W.C.: The Aging Population and Its Needs. IEEE Pervasive Computing 3(2), 12–14 (2004), http://dx.doi.org/10.1109/MPRV.2004.1316812
10. Naeem, U., Bigham, J., Wang, J.: Recognising Activities of Daily Life using Hierarchical Plans. In: Smart Sensing and Context, pp. 175–189 (October 2007)
11. Nugent, C.D., Mulvenna, M.D., Moelaert, F., Bergvall-Kåreborn, B., Meiland, F., Craig, D., Davies, R., Reinersmann, A., Hettinga, M., Andersson, A.-L., Dröes, R.-M., Bengtsson, J.: Home-based Assistive Technologies for People with Mild Dementia. In: Okadome, T., Yamazaki, T., Makhtari, M. (eds.) ICOST. LNCS, vol. 4541, pp. 63–69. Springer, Heidelberg (2007)
12. Philipose, M., Fishkin, K.P., Perkowitz, M., Patterson, D.J., Fox, D., Kautz, H., Hahnel, D.: Inferring Activities from Interactions with Objects. IEEE Pervasive Computing 3(4), 50–57 (2004), http://dx.doi.org/10.1109/MPRV.2004.7
13. Rantz, M., Aud, M., Alexander, G., Oliver, D., Minner, D., Skubic, M., Keller, J., He, Z., Popescu, M., Demiris, G., Miller, S.: TigerPlace: An innovative educational and research environment. In: Proceedings from AAAI Fall Symposium: AI in ElderCare, New Solutions to Old Problems, November 7-9, pp. 84–91 (2008)
14. Riva, G., Loreti, P., Lunghi, M., Vatalaro, F., Davide, F.: Presence 2010: The Emergence of Ambient Intellegence. In: Riva, G., Davide, F., IJsselsteijn, W.A. (eds.) Being There: Concepts, Effects and Measurement of User Presence in Synthetic Environments, Ios Press, Amsterdam (2003)
15. Ros, M., Delgado, M., Vila, M.A.: Fuzzy method to disclose behaviour patterns in a Tagged World. Submitted Expert System with Applications (2009)
16. Ros, M., Delgado, M., Vila, M.A.: A fuzzy approach for the model of sliding window. An application to behaviour patterns mining. In: IFSA-EUSFLAT, July 20-24, pp. 1211–1216 (2009)
17. Ros, M., Delgado, M., Vila, A.: A system to supervise behaviours using temporal and sensor information. Accepted to IEEE World Congress on Computations Intelligence, July 18-23 (2010)
18. Wilson, D.H., Wyaat, D., Philipose, M.: Using Context History for Data Collection in the Home. In: Gellersen, H.-W., Want, R., Schmidt, A. (eds.) PERVASIVE 2005. LNCS, vol. 3468. Springer, Heidelberg (2005)
19. Yager, R.R.: On ordered weighted averaging aggregation operators in multicriteria decisionmaking. IEEE Trans. Syst. Man Cybern. 18(1), 183–190 (1988), doi: http://dx.doi.org/10.1109/21.87068

A Normality Analysis-Based Approach to Monitor Behaviors in AAL Domains

D. Vallejo, J. Albusac, C. Glez-Morcillo, and L. Jimenez

University of Castilla-La Mancha, Paseo de la Universidad 4, Ciudad Real (Spain)
David.Vallejo@uclm.es
http://oreto.esi.uclm.es

Abstract. In this paper we discuss how an existing model for normality analysis of behaviors and a multi-agent architecture that gives support to such a model can be used on the Ambient Assisted Living domain. The use of this kind of models and architectures can contribute to support and help users in particular scenarios. A case study of an indoor environment in a hospital is described paying special attention to elderly people and patients.

Keywords: Normality Analysis; MultiAgent Architectures; Ambient Assisted Living Environments.

1 Introduction

Ambient Assisted Living (AAL) involves the infrastructure required to provide unobtrusive support in everyday life. This infrastructure covers the methods, software systems and specific devices for users, e.g. handicapped and elderly people, to get benefit and improve their quality of life. This paradigm establishes a direct link between the user and the environment so that the existing interactions represent a relevant issue to be addressed.

AAL is strongly related to Ambient Intelligence (AI) [1], which is based on the Ubiquitous Computing paradigm and refers to environments where people is surrounded by intelligent devices that recognize and respond to the users' need in every moment. Within this context, transparency plays the most relevant role so that the users are unaware of the provided services in the environment. In the last few years, AI has covered a wide range of applications and scenarios [4], being some of them shared by the AAL field.

In fact, the potential range of support systems in the AAL domain is huge [8]. Some relevant fields are emergency prediction/detection, assistance in daily tasks such as cooking or cleaning, services for finding things, or even orientation services. As previously introduced, the environment where the user lives plays an important role, usually distinguishing between indoor and outdoor environments. On the one hand, systems for indoor environments work in a well-defined space so that there are more limits when deploying software and hardware elements. On the other hand, systems for outdoor environments have to deal with more unstable conditions, such as the uncertainty when gathering information or the

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 347–356, 2010.

higher complexity of the environment. Depending on this criteria, the artificial system that gives support to AAL will act in one way or another.

In this paper we discuss how to apply our background in Intelligent Surveillance Systems [9] in the context of AAL. A direct application in which we are very interested in consists in detecting and analyzing the elderly people's and patients' behavior in order to detect specific actions and prevent possible accidents. To address this challenge, we are using two different tools: i) a formal model for behavior analysis [2] that allows to define the normality of specific events of interest, e.g. trajectories of people, and ii) a multi-agent [11] architecture that allows to instantiate these analysis modules and manages the hardware devices deployed in the environment [10]. Both the designed model and the implemented architecture can be perfectly adapted to be used on the AAL field in order to increment the intelligence and improve the capabilities of this kind of systems.

The rest of the paper is structured as follows. Section 2 goes into detail on our background in behavior analysis, analyzing how to adapt it to AAL. Section 3 discusses our approach for Ambient Assisted Living Environments. A case study of such an approach is shown in Section 4, which is mainly focused on monitoring the trajectories of elderly people in a hospital. Finally, concluding remarks and some future research lines are summarized in Section 5.

2 Related Work

2.1 Model for Normality Analysis

The monitoring of public and private spaces is increasingly more complex due to the very large number of sensors deployed to gather information from the environment. This tendency together with the impossibility for the security personnel to accurately watch large spaces and to monitor multiple events of interests at the same time demand for new surveillance approaches. Recently, we proposed a formal model [2] that allows us to design the named *surveillance concepts*, which define the normality of an aspect independently of the monitored environment. Thus, it is possible to instantiate a concept in a particular environment either learning the normality of such a concept, through machine learning algorithms [3], or explicitly defining it, through knowledge acquisition tools. In [2], the concept of *normal trajectory* is designed and instantiated in order to monitor a urban traffic environment.

Since these surveillance concepts are designed in a general domain-independent way, by means of fuzzy constraints [12], to define and monitor the normality of aspects or events of interest, this approach can be translated to the Ambient Assisted Living domain[1]. For instance, an immediate application consists in analyzing the movements of elderly people, both in indoor and outdoor environments, to detect whether they move normally and prevent accidents or falls. The current definition of normal trajectory comprises the following constraints:

[1] Although we have mainly used security cameras to gather information from the environment, it is possible to use less intrusive devices such as presence sensors or identity cards.

- *Spatial*, which could be used to check whether elderly people take the *correct* path when walking around, that is, that they follow the correct sequence of zones to get the desired position without invading a restricted area.
- *Temporal*, which could be used to check whether they spend the adequate time to accomplish a task.
- *Role*, which could be used to distinguish individual subjects or groups of people.
- *Destination*, which could be used to determine if elderly people move towards a specific place to accomplish a task (e.g. do they mean to walk towards their rooms?).

This concept of normal trajectory is only an example of the potential possibilities of the proposed method in the field of AAL. Nevertheless, it is important to remark that current software systems under this approach work in a *forensic* way, that is, they can make decisions of infer knowledge after detecting that a normality constraint was violated. Although the developed prototypes only need milliseconds (see [2] [3] for a detailed discussion) to give a response and allow to react to anomalies, our work in the domain of AAL when applying this model should be towards the prevention of undesired situations.

2.2 Multi-agent Architecture for Normality Analysis

To support the deployment of surveillance concepts in specific environments and provide the hardware and software resources to carry out the task of monitoring, a multi-agent architecture for cognitive surveillance [10] was designed and developed. We decided to adopt this approach since the monitoring of current environments involves a complex and open problem where the information is distributed around the environment and there is a need for decentralized control to improve robustness [11].

The use of an agent-based approach in AAL can contribute to improve the design of a solution based on the intelligent management of distributed knowledge by offering a scalable and flexible approach when integrating new analysis modules. Thus, the intelligent agents are responsible for giving support to assisted living. In this way, it is possible to conceive an approach where such agents provide and request services or even compose them to solve more complex tasks.

This is why we have extended the existing architecture in [10] to include new classes of intelligent agents and new communication mechanisms in order to address the requirements of AAL environments. In the next section, the new architecture, aimed to cover the needs of normality analysis in the AAL domain, is described more in detail.

3 Our Approach

3.1 General Overview of the Architecture

The architecture is structured in three layers (see Figure 1): the reactive level, the deliberative level and the user level. All of them communicate with one another through the communication framework, which allows the interaction of the

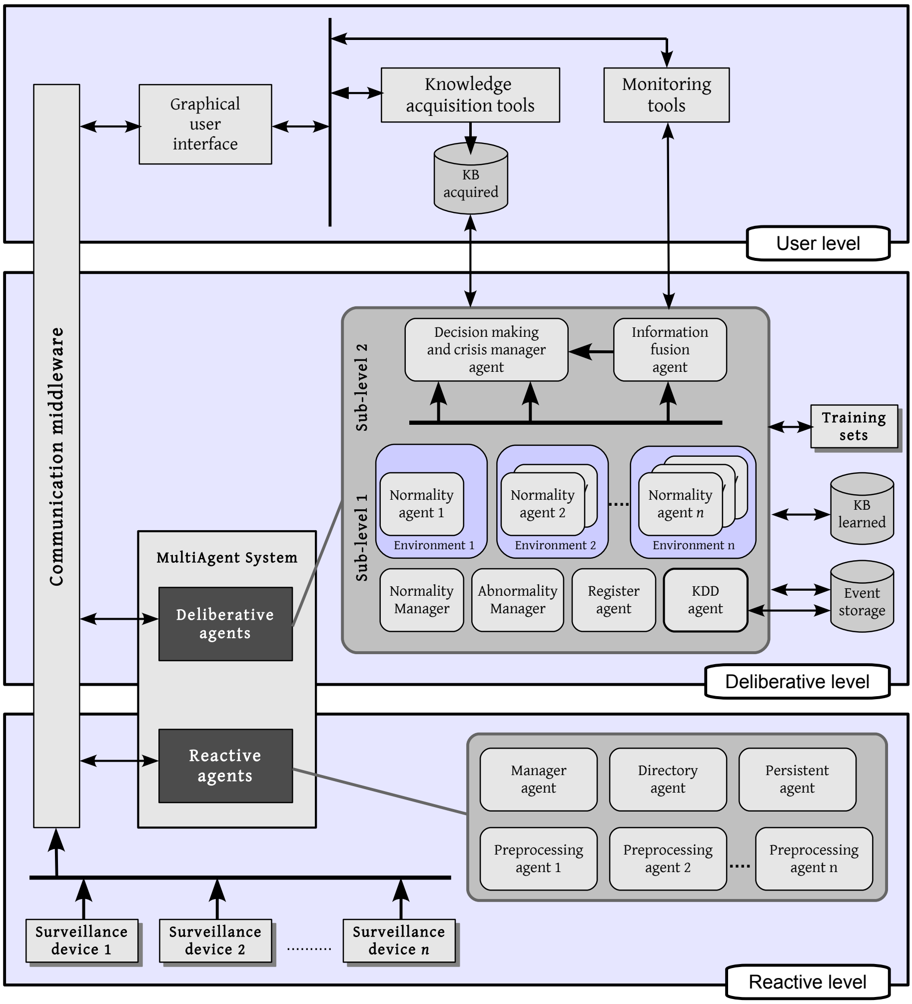


Fig. 1. Proposed Multi-Agent Architecture for Ambient Assisted Living environments

components of the architecture and the integration of new elements. The reactive level is composed of the information retrieval devices from the environment and those management agents that are based on a reactive scheme. The deliberative level is responsible for managing the high-level knowledge in order to provide the system with all the services needed to assist or analyze behaviors within the environment in an intelligent way. This level brings together all the agents that maintain and use a domain knowledge model with the aim of analysing the normality of the environment from different points of view, as well as those that carry out the information fusion obtained from the different perceptions, the decision-making, and the alarm management. Finally, the user level allows the interaction with the user through knowledge acquisition and monitoring tools.

The communication between all the components of the artificial surveillance system is done by means of a component shared by all the architecture levels. This component is the communication middleware (see Figure 1), currently implemented through the ICE middleware [7], which can be defined as a service that allows the communication between the heterogeneous components that compose the artificial system. Within this context, the concept of heterogeneity refers to the diversity of components (hardware devices, software agents, user interfaces, monitoring tools. . .) and the variety of platforms (operating systems, programming languages, communication networks. . .).

3.2 Reactive Level

This first functional level associated to the environmental data to get basic information is supported by the reactive level of the architecture. As depicted at the bottom of Figure 1, the multi-agent system hosts the agents specialised (*preprocessing agents*) in processing the data obtained from the devices deployed in the environment.

The rest of agents of the reactive level are management agents (design according to the FIPA [6] specifications) responsible for watching over the proper working of the multi-agent system from a managerial point of view. Their main tasks involve registering the agents within the agent platform and providing the basic services for the agents to strike up conversations and exchange messages. These two agents are as follows:

– *Manager Agent*: responsible for the administrative management of the MAS, that is, registering agents within the artificial surveillance system and the controlling over the hardware and software resources of the global system.
– *Directory Agent*: responsible for providing the system with yellow-pages service, that is, it allows to search agents or components from a functional description. For instance, if the segmentation agent needs the tracking service, then it could ask the directory agent for the agent that provides such service.

3.3 Deliberative Level

The deliberative level is composed of those agents that maintain and use a knowledge model in order to analyse the normality and the abnormality of the environment from different points of view (normality concepts), as well as those agents that aggregate the information obtained from the sub-environments that establish the whole scenario and make decisions.

In the domain of AAL, the more relevant agents are as follows:

– *Normality Agent*: agent that manages a knowledge model of the environment, where it was deployed, and a reasoning model that allows to analyze the normality of its associated concept by emulating the reasoning model of the human in charge. The architecture of the system allows to instantiate as many normality agents as needed, depending on the requirements of the AAL environment.

- *Normality Manager*: agent responsible for coordinating the global normality when monitoring the whole environment. The main service provided by this manager consists in checking the normality of issues independent of the concepts used to perform the analysis.
- *Abnormality Manager*: agent in charge of managing a knowledge model based on the abnormality to complement the normality model used by normality agents. In fact, this agent knows the more common anomalous situations of concrete environments by identifying when a particular object violates the constraints associated to the normality of the concepts that are being analysed by the artificial system. If an object violates the normality of a concept, this agent interacts with the normality agents to know what kind of constraints were not satisfied and the temporal evolution of the monitored object. The abnormality agent makes use of this information to infer the anomalous situation, previously stored in the agent knowledge base by means of knowledge acquisition tools.
- *Register Agent*: this agent is responsible for the persistent storage of all the knowledge generated by the deliberative agents in the database *event storage*. The internal model of this agent has into account the knowledge obtained by the different normality agents to store all the relevant information for applying machine learning techniques or knowledge discovering and data mining.
- *Knowledge Discovery and Data Mining (KDD) Agent*: this agent employs the information stored in the database *event storage* to deploy knowledge discovery processes [5] over such information. This kind of techniques allows the extraction of useful knowledge from information sources of huge volume. In the case of the artificial system devised in this work, all the information generated by the normality agents (from the same or different local environments), by the information fusion agent, and by the abnormality manager can be employed to extract behaviour patterns useful for providing advanced analysis. One of the main applications if behaviour prediction.
- *Decision Making and Crisis Manager Agent*: this agent gives support to the user when making decisions and managing actions as a consequence of detecting an anomalous situation. The model of this agent allows to suggest decisions from the knowledge generated by the normality agents and the information fusion agent. Since the decision making in view of dangerous situations is a delicate issue, a scheme in which the agent suggests possible solutions has been adopted so that the human in charge has the last word. In fact, this agent can be defined as a decision support agent.

3.4 User Level

The user level represents the top layer of the architecture (see Figure 1) and provides the mechanisms of interaction between the user and the artificial system. Two kind of tools are distinguished within this level: knowledge acquisition tools and monitoring tools. The first ones are complemented with machine learning algorithms for knowledge generation or with previous experience that the system will use to understand what happens from the information coming from sensors. In other words, these tools allow to turn the tacit knowledge of AAL experts into explicit knowledge, which will be used to establish the normality of a concept within the monitored concept or, in other words, to establish the reasoning model of a normality agent. On the other hand, the monitoring tools allow the user to observe what is happening in a concrete environment, as well as debug the normality analysis made by the system.

3.5 Communication Mechanisms

This section summarizes the communication mechanisms provided by the multi-agent architecture in order to give support for the agents to communicate with one another when dealing with ALL environments: a) event-based communication, b) direct communication between agents, and c) communication through a blackboard system.

Event-based communication is used to deploy typed event channels, that is, the communication revolves around the kind of information which is sent and received instead of taking into account the identity of senders or receivers. This class of communication perfectly fits in the interaction between the components of the reactive level (sensors and preprocessing agents) and the agents of the deliberative level (normality agents, among others).

Direct communication allows an agent to communicate with one or more agents whose physical address is explicitly established within the message to be sent. This class of communication can be done with no intermediary (agent-agent) or through an agent that implements the message transport system specified by FIPA (*Message Transport Service*).

Finally, the communication through a **blackboard system** supported by the architecture makes possible that several agents share knowledge when analyzing an environment. Typically, the blackboard will be used to write results by the normality agents that analyse a common concept, e.g. trajectories.

4 Case Study: Analyzing Trajectories in a Hospital

Figure 2 graphically shows one indoor environment of a real hospital where the discussed multi-agent architecture can be deployed to analyze the normality and give support to the hospital staff. Our main intention is to monitor the elderly people's behavior, paying special attention to the concept of *normal trajectory*. To do that, we are redefining such a concept (previously defined in [2] for a urban traffic environment) by considering the sensors deployed in the hospital and the kind of people (paramedics, staff, patients, etc) we find there.

Let us assume there is a camera that covers the whole hall where elderly people usually walk around. In this case study, we are supposing a cenital camera to simplify the explanation. To monitor trajectories, we first need to define the set of regions of interest in which the scenario can be divided into (see Figure 3). Next, we need to identify elderly people and track their position somehow. One possible

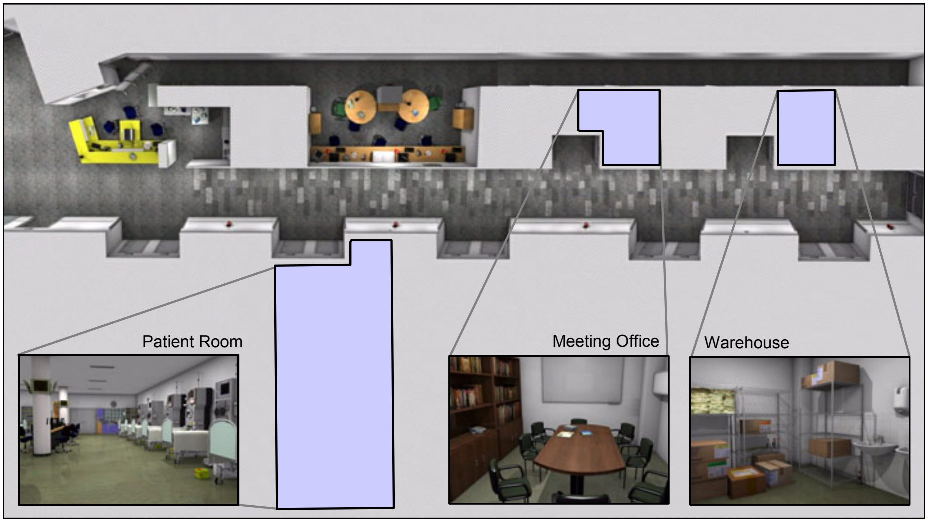


Fig. 2. Virtual view of a real hospital. The division into multiple zones was used to define normal trajectories.

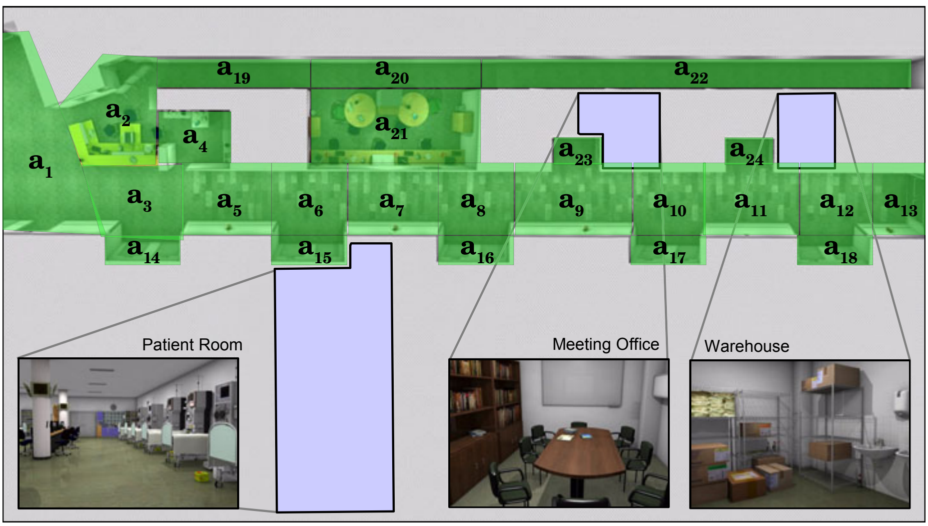


Fig. 3. Division into multiple zones of the hospital hall

way to solve these problems is assign identity cards to people, distinguishing between different roles (patient, nurses, paramedics, etc).

Now, the normal trajectories for elderly people can be defined. Currently, we are using a knowledge acquisition tool with a graphical interface to explicitly define such trajectories. However, the trend should be towards the learning of trajectories in order to reduce the human dependence when accomplishing this task (see [3] for the concept of normal speed). Within this context, we can use such a tool to define the constraints of normal trajectories (role, spatial, temporal, etc). In [2] we discuss in depth how to define and deploy the concept of normal trajectory.

According to Figure 3, a normal trajectory for a patient would involve the zones a_1, a_3, a_5, and a_{15}. Similarly, a normal trajectory for a nurse would involve the zones a_1, a_2, a_{19}, a_{20}, and a_{21}. However, this would not be a normal trajectory for patients since they are not allowed to access to the reserved area for the hospital staff. Trajectories for elderly people also include temporal constraints to check if they spend an abnormal period of time in particular areas (e.g. prowling in the main hall).

From the architectural point of view, the sensors deployed in the reactive level gather the environmental information (e.g. the cenital camera, the reader for identity cards, etc). To analyze behaviors we instantiate the normality agents distinguishing two options (see [10] for implementation details on the architecture): i) one normality agent per concept (e.g. trajectories) or ii) one agent per detected human being, depending on the particular environment. In the hospital case study, we are deploying a single agent. However, if we are more interested in the interactions between people when monitoring a particular event of interest, the second option may be more suitable.

In case of detecting some normality violation (e.g. an elderly man/woman who spends too much time in the bathroom), the normality agent sends an alarm to the decision making agent (see Figure 1), which can notify this situation to the nurse's PC/PDA.

5 Conclusions

This paper raises how a model for normality analysis and a multi-agent architecture that supports it can be translated to deal with topics related to Ambient Assisted Living. The combination of such proposals can improve the scalability at knowledge management and architectural level in AAL environments. Within this context, this paper discusses how to expand our systems to detect abnormal situations so that the staff in charge can correctly assist people. The studied approach for normality analysis can be used on any environment to monitor any event of interest since it is general and domain-independent. On the other hand, the discussed multi-agent architecture covers the most important requirements of this kind of problematic: i) the reactive level provides the communication mechanisms for any kind of sensor, ii) the deliberative level hosts the agents responsible for analyzing different concepts and supporting the decision-making task, and the iii) user level establishes the communication with the knowledge acquisition and monitoring tools.

Our future work is aimed to define more normality concepts for the AAL domain and enrich the definition of the currently designed ones. For instance to support handicapped people when they are moving by using multiple sensors. Another current research line is towards the learning of normal activities to avoid the definition of explicit knowledge.

Acknowledgements

This work has been funded by the Regional Government of Castilla-La Mancha (Research Project PII1C09-0137-6488) and by the Ministry of Science and Innovation (Research Project TIN2009-14538-C02-02).

References

1. Weber, W., Rabaey, J., Aarts, E.H.L.: Ambient intelligence. Springer, Heidelberg (2005)
2. Albusac, J., Vallejo, D., Jimenez-Linares, L., Castro-Schez, J.J., Rodriguez-Benitez, L.: Intelligent Surveillance based on Normality Analysis to Detect Abnormal Behaviors. International Journal of Pattern Recognition and Artificial Intelligence 23(7), 1223–1224 (2009)
3. Albusac, J., Castro-Schez, J.J., Lopez-Lopez, L.M., Vallejo, D., Jimenez, L.: A Supervised Learning Approach to Automate the Acquisition of Knowledge in Surveillance Systems. Signal Processing 89(12), 2400–2414 (2009)
4. Ducatel, K., Bogdanowicz, M., Scapolo, F., Leijten, J., Burgelman, J.C.: Scenarios for ambient intelligence in 2010. In: Office for Official Publications of the European Communities (2010)
5. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: The KDD process for extracting useful knowledge from volumes of data. Communications of the ACM 39(11), 27–34 (1996)
6. Foundation for Intelligent Physical Agents FIPA, FIPA Agent Management Specification (2004), `http://www.fipa.org/specs/fipa00023/index.html`
7. Henning, M.: A new approach to object-oriented middleware. IEEE Internet Computing 8(1), 66–75 (2004)
8. Nehmer, J., Becker, M., Karshmer, A., Lamm, R.: Living assistance systems: an ambient intelligence approach. In: 28th International Conference on Software Engineering, pp. 43–50. ACM, New York (2006)
9. Valera, M., Velastin, S.: Intelligent distributed surveillance systems: a review. IEEE Proceedings Vision, Image and Signal Processing 152(2), 192–204 (2005)
10. Vallejo, D., Albusac, J., Gonzalez-Morcillo, C., Jimenez, L.: A Service-Oriented Multi-Agent Architecture for Cognitive Surveillance. In: Klusch, M., Pěchouček, M., Polleres, A. (eds.) CIA 2008. LNCS (LNAI), vol. 5180, pp. 101–115. Springer, Heidelberg (2008)
11. Weiss, G.: Multiagent systems: a modern approach to distributed artificial intelligence. The MIT Press, Cambridge (2000)
12. Zadeh, L.A.: Fuzzy sets. Information and Control 8(3), 338–353 (1965)

Adaptation of an Evaluation System for e-Health Environments

Nayat Sánchez-Pi and José Manuel Molina

Carlos III University of Madrid,
Avda de la Universidad Carlos III, 22. 28270,
Madrid, Spain
nayat.sanchez@uc3m.es, josemanuel.molina@uc3m.es

Abstract. The increase in ageing of European population implies a high cost in economy and society in any European country and it can be reduced if we pay attention and develop home care systems. Evaluation of these systems is a critical and challenging issue but seldom tackled. It is important before evaluating a system to figure out what is the evaluation goal. In our case, such a goal is to evaluate enhanced user experience and beyond the evaluation goal it is also a central concern about what to evaluate. In this paper we propose a multi-agent home care system where we describe how agents coordinate their decisions to provide e-services to patients when at home after hospitalization. Finally we center our proposal on the adaptation of an evaluation system, previously developed, to support the challenges of an e-Health environment and also the multi-user evaluation. These evaluation methods (online/offline) will provide user's (patients, patient's relatives and healthcare professionals) feedback into the system.

1 Introduction

The feasibility of using information and communications technologies (ICT), such as sensor networks, radio frequency identification (RFID) and Universal Mobile Telecommunications System (UMTS), has led to improve e-services and applications in the ambience of electronic healthcare (e-Health) [1]. The European Furthermore, the increase in ageing of European population and the treatment of chronic and disabled patients implies a high cost in terms of time and effort. Sometimes patients and also healthcare workers consider treatments in health centres unnecessary as they could collapse national health services and increase costs. On the other hand, we face the problem of the patients living in rural areas, where is difficult to access. To face these challenges we need to differentiate medical assistance in health centres from. assistance in a ubiquitous way that it is possible due to the advances in communication technologies. Ubiquitous healthcare (u-health) is an emerging area of technology that uses a large number of environmental and patient sensors to monitor and improve patients' physical and mental condition. U-Health focuses on e-Health applications that can provide health care to people anywhere at any time using information and

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 357–364, 2010.

communication technologies. Besides, innovative approaches in mobile healthcare (m-Health) have also been developed as a footbridge between e-Health and u-Health. Several initiatives, such as Mobihealth [2], XMotion [3] and MyHearth [4] have investigated the feasibility and benefits of mobile healthcare services and applications. There have been also several attempts to developed home care systems, such as Gator Tech House [5], PAUL (Personal Assistant Unit for Living) [6] and AMADE [7]. However, these initiatives do not provide a clear application framework that simplifies the development of e-health applications and at the same time do not provide an evaluation system taking into account the contextual information as well as the user's opinion.

The main contributions of this work are the following:

- i) Effectively adapting the design and architecture of an agent-based developed in previous works [8] [9] to an agent-based home care system in order to provide m-services to patients/ caregivers/ patient's relatives when at home after patients hospitalization.
- ii) Adapting the evaluation system developed in previous work [10] to support the challenges of u-health environment, and also the challenge of providing multi-user evaluation and feedback into the system.

The remainder of this paper is structured as follows: Section 2 presents a scenario of u-commerce environment and its evaluation from the users' point of view developed in previous work. Section 3 describes the MAS adaptation to e-health environment. Section 4 presents a scenario of e-health, and the evaluation system adaptation to e-health environment. Section 5 presents some conclusions and directions for future work.

2 User Evaluation for Context-Aware System Based in MAS

There are several methods [11] and approaches used to evaluate context-aware systems. Evaluation of these systems is a critical and challenging issue but seldom tackled. As there are no established evaluation frameworks in literature, in previous work [10] we have used, first of all, a pre-implementation evaluation method as the 'Wizard of Oz' [12]. And latter we used a method called 'revisiting the hypotheses' [11], which he divides it into four hypotheses to be investigated. Regarding to the evaluation from the users point of view, in [9] we based on two propositions an offline customers evaluation (once the user has finished using the system and that can be accessible via: www.giaa.inf.uc3m.es/u-shopping/myfeedback/) and an online customers evaluation in a u-commerce environment. In any of them, customer makes a quantitative evaluation by defining a utility function and mapping the satisfaction state to numeric value. The value of quality of shopping/vendor services corresponds to the users feedback and can have values like: (1) Correct, (2) Different order or (3) Incorrect for every attribute.

The evaluation for a multi-agent context-aware system for a u-commerce environment is represented by the following scenario where a young customer 'John'

Scenario u-Commerce:
User name: John Mayer
Location: Parque Sur Shopping Center
Date and Time: 2009-12-14, 17:00
Device: PDAQ 00-18-41-32-0B-59
Objective: To buy some new clothes, find new offers

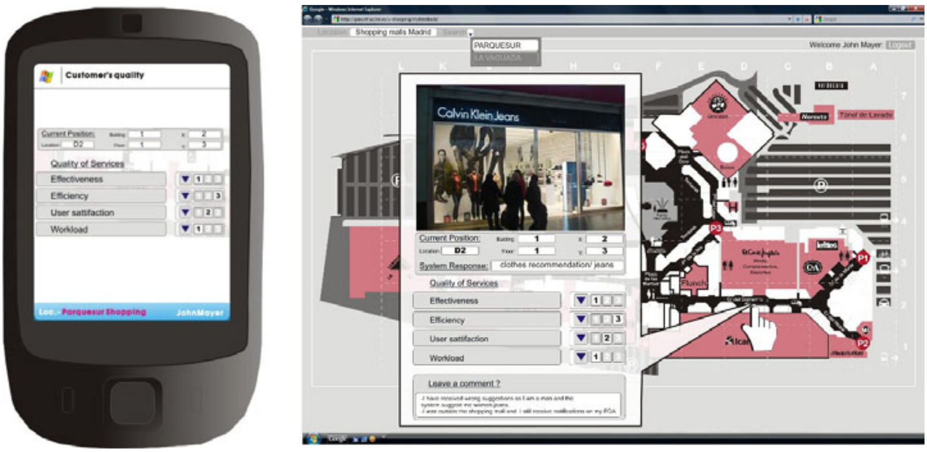


Fig. 1. Online (left) and Offline (right) quality evaluation

goes to 'Parque Sur' shopping mall because he needs to buy some new clothes. See Fig 1.

3 MAS Adaptation to e-Health Environment

Tourism, Healthcare, Education, Transportation, etc., are some of the environments where been developed ubiquitous systems and applications above AmI vision. In the case of modern healthcare, it includes user mobility allowing people at risk or patients with proved health problems to continue their usual life at their homes and work places. Furthermore, health care professionals also need to access and input medical or patient information from anywhere, at any time in their daily ward rounds [13] [14] [15]. Hence, mobile healthcare systems can facilitate efficient and effective patient care information input and access at the point of patient care.

Taking into account all of the above, we have adapted the design and architecture of a multi-agent system developed in previous works to support u-health services provisioning [8] [10]. The redesign includes new features to support u-health services and applications where users will become 'patients or caregivers or patient's relatives' and where a new categories of agent appear: 'evaluation

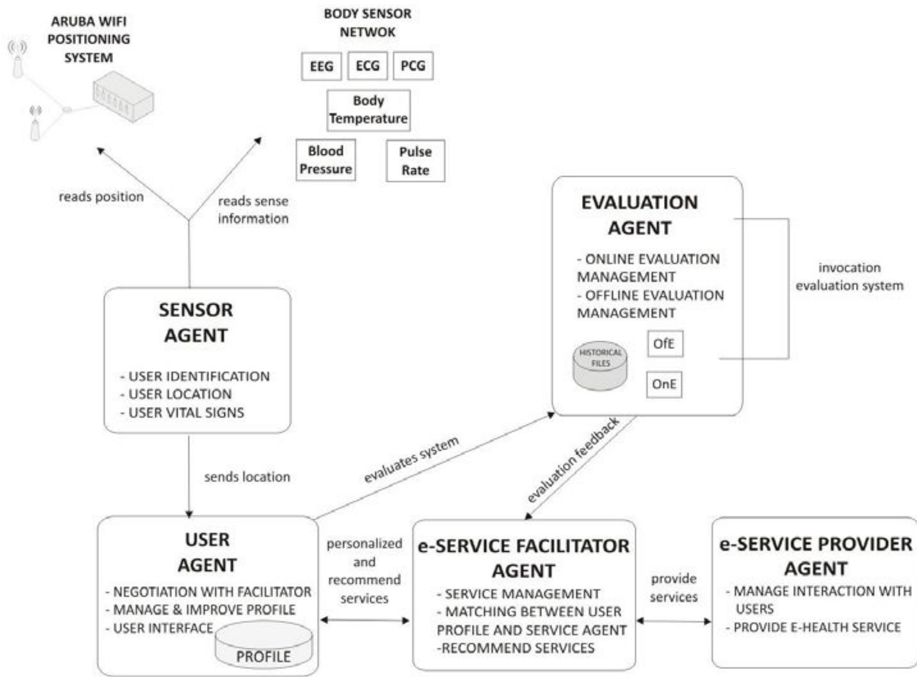


Fig. 2. MAS architecture adaptation

agent'. See fig 2. The interaction between the different agents is described by the following sequence of phases:

- i) The Aruba Positioning system the patient's position while at WiFi home network and the Body Sensor Network provides the information regarding each sensor parameter;
- i) The sensor agent provides sensor information to the user agent;
- iii) Once the user agent knows its location, and its vital signs, sends it to the e-service facilitator agent. It also provides information regarding the type of user is using the system (patient/caregivers/patient's relatives);
- iv) The e-service facilitator agent sends the e-service provider agent the identification of the user agent that provides the kind of e-services required by the user agent;
- v) The user agent asks a specific e-service provider agent to provide it with the required e-service;
- vi) E-service provider agent asks the user agent about context information to be used during the interaction to provide the personalized e-service;
- vii) User agent provides the required context information to the e-service provider agent;
- viii) Interaction between the user agent and the e-service provider agent using the adaptation provided by the previous step;

– ix) Once the interaction and the provisioning of e-service tool place, evaluation agent ask the e-service facilitator agent about the type of user;
– x) Evaluation agent invokes Online/Offline evaluation system depending of the type of user; xi) Historical files are stored for analysis and feedback is provided into the system.

4 User Evaluation for e-Health System Based in MAS

Adaptation of the evaluation system developed in previous work [10], to suit e-health environments takes into account not only the characteristics of the environment, but also the quantitative user evaluation. Quality process has two distinct facets: technical quality and functional quality. Technical quality refers to the accuracy of medical diagnoses and procedures, and is generally comprehensible to the professional community, but not to patients [16]. Patients essentially perceive functional quality as the manner in which the service is delivered; while healthcare professional can be capable of making a technical quality evaluation.

There are several proposals regarding service quality measurement. Some of them are: SERVQUAL instrument proposed by Parasuraman [17]; SERVPERF [18] [19]; Yoo and Donthu [20] and Zhang and Prybutok [21]. Regarding this, for e-health environment we consider two groups of users: patients/caregivers/patient's relatives and health professionals. The first group will be able to make an online evaluation (OnE) of the system, for which we have defined some service quality measurement, and the second one, an offline evaluation (OffE) with other service quality measurement that evaluates, in this case, the technical quality of the system response.

Main contributions regarding adaptation of user evaluation are:

– i) First, as the awareness of the system has been adapted for the e-health environment, the evaluation will be done based not only on the patients' location (as we did in [10], but also on his vital signs: blood pressure (BP); pulse rate (PR); respiration rate (RR) and body temperature (BT). We based on the fact that the system is composed of a set of different sensors connected to a PDA that transmits, in a secure way, all the patient data (location and vital signs) to a central server in the hospital. The authorized doctors can access this medical information from their computers (inside the hospital or even outside) afterwards.
– ii) In the case of OnE, main contributions related to the adaptation of the service quality measurement to e-health are: Quality parameters measures the service quality gap between client expectations and perceptions of 5 quality attributes (on a five-point scale: strongly disagree = 1 to strongly agree = 5). Attributes are: easy of use; proceeding speed and effectiveness; reliability.
– iii) For OffE, we explore the e-service quality dimensions based on a review of the development of e-service quality scales and the SERVQUAL scale [17]. It proposes an 8-dimension scale but we will adapt this scale and

Scenario e-Health: Elderly in a wheelchair + e-health services
Event part: When the wheelchair (it is supposed to be the elderly person) with RFID-tag is detected in the TV room.
Condition part: (and) vital signs are: BP: 130/85; PR: 80; RR=17; BT: 37ºC.
Action part: (then) activate the EMERGENCY MEDICATION ALERT SERVICE and VoIP functionality on the PDA and through a voice message indicates \"Mr. Mark you should take the medication for high pressure. Please take 1 pill right now".

add 2 more dimensions: system design (Appealing and well organized website; Consistent and standardized navigation; Well-organized appearance of user interface; Quickly downloading), reliability (Accurate delivery e-service; Complete order e-service; System being truthful about its diagnosis; The online e-service always correct; Keeping e-service promise; Accurate online e-service records; Website always available), fulfilment (Information on e-services available when need it; System runs smoothly in the transaction process; Accurate promises about delivery e-service when scheduled; Available to modify and/or defer the e-service process at any time without commitment), security (Protect the personal data of customers; Good reputation), responsiveness (Adequate contact information and performance; Prompt responses to customers; Timely responses to customers; Adequate response

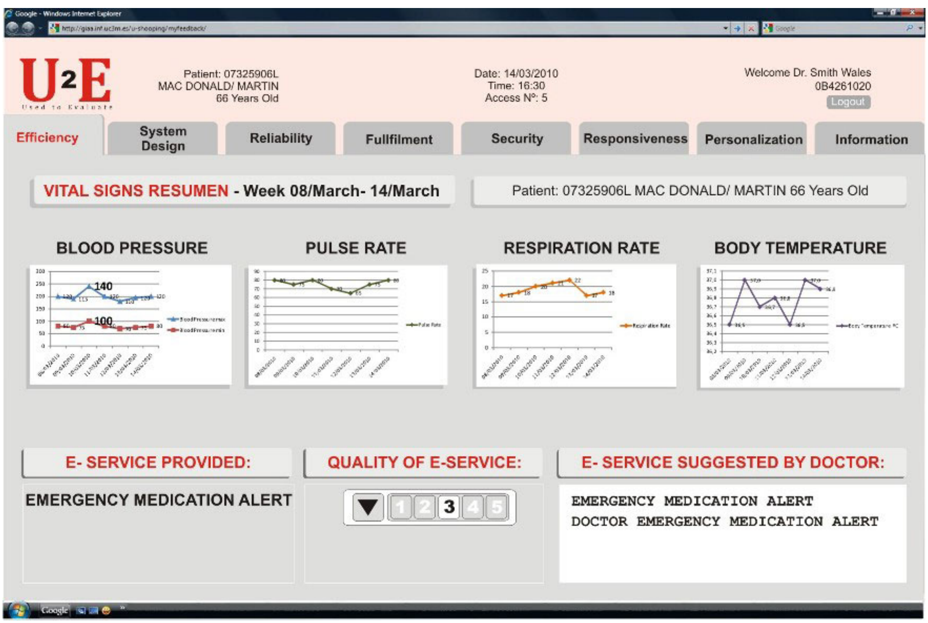


Fig. 3. Online (left) and Offline (right) quality evaluation

time), personalization, information and efficiency. Likerts's five point scale is used (strongly disagree = 1 to strongly agree = 5).

The evaluation for a multi-agent e-health system is represented by the following scenario where an elder patient stays home after hospitalization for a treatment. So, once the e-service is provided to the patient, the OnE evaluation system is invoked by the evaluation agent and patient/caregiver/patient relatives can make the evaluation of the e-services received filling the evaluation form. Doctor or health professional in charge of following the patients' file, can also evaluate the system behaviour as see in Fig 3. In this case Dr. makes an offline evaluation of the behaviour of the system during a week. Dr. suggests the system, in a similar case, to activate the DOCTOR EMERGENCY MEDICATION'S ALERT that will send a message to the doctor, so he can be notified immediately. So our system could then: i) discriminate between contextual information; ii) allow authorized health professional to suggest the correct e-services that should be provided in each case; iii) allow patients/caregiver/patients' relatives give their opinion about the received e-service in each case. In order to give a general measure of the e-service facilitator agent results over the satisfactory cases, we evaluate how close the e-service facilitator agent's ranking is to the user's own ranking as presented in [10].

5 Conclusions and Future Research Agenda

Evaluation of e-health systems from the users' point of view is a critical and challenging issue but seldom tackled. It is important before evaluating a system to figure out what is the evaluation goal and who will evaluate the system. We adapted an evaluation system, previously developed, to support the challenges of e-Health environment, and also the multi-user evaluation. Patients, patient's relatives and caregivers are part of the same group of user evaluators that uses the online user evaluation system through the PDA to evaluate the system output. On the contrary, authorized healthcare professionals can provide feedback into the system taking into account the vital signs of the patient, and the e-service provided by the system. In order to give a general measure of the e-service facilitator agent's results over the satisfactory cases, we evaluate how close the e-service facilitator agent's ranking is to the user's own ranking. For this, we chose the Manhattan distance between the position of the first three products selected by the user and their position in the system ranking [10]. As future trends, we plan to give the statistical results of this analysis and provide the feedback to the system.

Acknowledgements

This work was supported in part by Projects CICYT TIN2008-06742-C02-02/TSI, CICYT TEC2008-06732-C02-02/TEC, CAM CONTEXTS (S2009/TIC-1485) and DPS2008-07029-C02-02.

References

1. Wyatt, J., Sullivan, F.: Ehealth and the future: promise or peril? British Medical Journal 331, 1391 (2005)
2. Van Halteren, A., Bults, R., Wac, K., Konstantas, D., Widya, I., Dokovsky, N., Koprinkov, G., Jones, V., Herzog, R.: Mobile patient monitoring: The mobihealth system. The Journal on Information Technology in Healthcare 2(5), 365–373 (2004)
3. X-motion project webpage (2004)
4. Myheart project webpage
5. Helal, S., Mann, W., El-Zabadani, H., King, J., Kaddoura, Y., Jansen, E.: The gator tech smart house: A programmable pervasive space. Computer, 50–60 (2005)
6. Floeck, M., Litz, L.: Integration of home automation technology into an assisted living concept. In: Assisted Living Systems-Models, Architectures and Engineering Approaches (2007)
7. Fraile, J., Bajo, J., Corchado, J.: Am ade: Developing a multi-agent architecture for home care environments (2008)
8. Sánchez-Pi, N., Carbó, J., Molina, J.: Jade/leap agents in an aml domain. In: Corchado, E., Abraham, A., Pedrycz, W. (eds.) HAIS 2008. LNCS (LNAI), vol. 5271, pp. 62–69. Springer, Heidelberg (2008)
9. Sánchez-Pi, N., Molina, J.: A multi-agent platform for the provisioning of u-commerce services. In: 28th North American Fuzzy Information Processing Society Annual Conference (2009)
10. Sánchez-Pi, N., Molina, J.: A multi-agent approach for the provisioning of e-services in u-commerce environment. Internet Research 20 (2010), ISSN: 1066-2243
11. Schmidt, A.: Ubiquitous Computing- Computing in Context. PhD thesis (November 2002)
12. Dahlback, N., Jonsson, A., Ahrenberg, L.: Wizard of oz studies-why and how. Knowledge-Based Systems 6(4), 258–266 (1993)
13. Haux, R.: Health information systems-past, present, future. International Journal of Medical Informatics 75(3-4), 268–281 (2006)
14. Johnson, C., Turley, J.: The significance of cognitive modeling in building healthcare interfaces. International Journal of Medical Informatics 75(2), 163–172 (2006)
15. Reuss, E., Menozzi, M., Buchi, M., Koller, J., Krueger, H.: Information access at the point of care: what can we learn for designing a mobile cpr system? International Journal of Medical Informatics 73(4), 363–369 (2004)
16. Bopp, K.: How patients evaluate the quality of ambulatory medical encounters: a marketing perspective. Journal of Health Care Marketing 10(1), 6–15 (1990)
17. Parasuraman, A., Zeithaml, V., Berry, L.: Servqual: a multiple-item scale for measuring consumer perceptions of service quality. Retailing: Critical Concepts 64(1), 140 (2002)
18. Gronroos, C.: Service management and marketing: managing the moments of truth in service competition. Jossey-Bass, San Francisco (1990)
19. Cronin Jr., J., Taylor, S.: Measuring service quality: a reexamination and extension. The Journal of Marketing 56(3), 55–68 (1992)
20. Yoo, B., Donthu, N.: Developing a scale to measure the perceived quality of an internet shopping site (sitequal). Quarterly Journal of Electronic Commerce 2, 31–46 (2001)
21. Zhang, X., Prybutok, V.: A consumer perspective of e-service quality. IEEE Transactions on Engineering Management 52(4), 461–477 (2005)

OutCare: Supporting Dementia Patients in Outdoor Scenarios

Jie Wan[1], Caroline Byrne[1], Gregory M.P. O'Hare[2], and Michael J. O'Grady[2]

[1] Science School, Institute of Technology Carlow, Kilkenny Road, Carlow, Ireland
[2] CLARITY: Centre for Sensor Web Technologies, University College Dublin, Ireland
`Jessie.wan@yahoo.co.uk, caroline.byrne@itcarlow.ie, {Gregory.OHare,Michael.J.OGrady}@ucd.ie`

Abstract. Ambient Intelligence (AmI) is a profound vision of computing power that is invisibly embedded into the fabric of everyday life. It is accessible through intelligent interfaces which are so natural that they can be used unconsciously. The increasing maturity, performance and miniaturization of processors, networking technologies and wireless sensor networks enable a move towards pervasive computing, ubiquitous connectivity and adaptive software. The Intelligent Agent paradigm has proven itself to be a promising branch of Artificial Intelligence (AI), complementing the pervasive trend of network engineering. One of the possible usages of AmI is Ambient Assisted Living (AAL) which attempts to utilize appropriate technologies to support citizens in living in their preferred environment independently, for a longer period of time than may otherwise be the case. This paper proposes an AAL solution for Alzheimer's patients based on the intelligent agent platform, exploring the practical delivery of intelligent environments that are sensitive and responsive to the patients, supporting them in performing daily activities and tasks in a natural, unobtrusive way.

Keywords: Ubiquitous Computing, Ambient Assisted Living, Intelligent agents, pervasive health.

1 Introduction

Demographics are changing at an alarming rate in both Europe and worldwide. Life expectancy is increasing; the percentage of older Irish male citizen aged 65 or over will rise from 9.7% in 2002 to 14.1 percent in 2021 [1]. The percentage of Irish female aged 65 or over will rise from 12.5% in 2002 to 16.4% in 2021. The absolute number of older citizens was 436,401 in 2002; the corresponding figure will increase to 730,000 by 2021 with the life expectancy rising to 84. The equivalent increase of people aged 75 or over will be approximately 52% by 2021. It is also projected that the number of older people living alone will increase substantially between 2002 and 2021, in line with the growth of the overall number of old people. By 2021 there will be over 30% of those aged 65, approximately 211,000 and over living alone.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 365–374, 2010.

Between 2002 and 2036, the approximate number of people with dementia in Ireland is expected to increase by 303% while the total population will increase by less than 40%; however dementia worldwide will witness a four folds increase by 2050 reaching 117 million.

The cost of healthcare is continuously increasing; for example, the fiscal cost of a hospital-bed has become punitive, running at a weekly cost of €778 on average [2]. There are nearly 50,000 people are currently involved in caring someone with one of the six symptoms of dementia [3]. In 2005 Ireland spent 8.2% of GDP on healthcare; in 2009 it rose to 16%, which is estimated to double again in 2035.

As the figures alluded to testify, the aging demography profile will give rise to significant problems, both financial and social. Thus a key issue for consideration concerns how technology might be prudently harnessed to such that these costs may be reduced.

1.1 Ambient Intelligence

Ambient intelligence [4] (AmI) is a multidisciplinary paradigm built upon ubiquitous computing. It fosters a novel anthropomorphic human-centric computer interaction design. AmI enables the basic criteria to build intelligent environments where devices can be so miniature as to disappear into the background, being seamlessly embedded into the environment. Fundamental to AmI is the recognition of objects and their situational context, and the delivery of services that are personalized, adaptive and anticipatory [5].

Intelligent agents [6] have become a mature technology for integrating ambient intelligence into smart applications. Intelligent agents are able to perceive their environment, apply deductive and practical reasoning mechanisms to respond in a timely fashion to changes in order to satisfy objectives. Secondly, they are able to exhibit goal-directed behaviours by taking the initiative, and capable of interacting with other agents. Their distributed nature makes agents are a particularly apt model for realising AmI.

1.2 Artificial Intelligence (AI) in Healthcare

Ambient Assisted Living [7] has been conceived as a paradigm for enabling people live on their own for longer periods than would otherwise be the case. In this way, costs would be reduced and quality of life maintained. Thought there are a multitude of technologies around which AAL systems may be constructed, AmI offers one particularly promising avenue. In this paper, a prototype mobile services for those afflicted with dementia is described.

2 Related Research

The most significant step in AAL area is enabling the development of an environment based on the concept of ubiquitous sensing. It attempts to develop a distributed and networked computing infrastructure to support users while remaining transparent to them. Combinations of processors and sensors are embedded into a distributed and elaborate network to provide user identification, face/speech recognition, tactile tracking, movement monitoring, pose tracking, facial expression and gesture recognition, audio processing and physical environment detecting functionalities, amongst others. Immense research efforts have been made in this direction and some commercial products are available.

CareLab [8] resembles a one-bedroom apartment which is equipped with a rich sensor network in conjunction with a health and wellness application, to support seniors' independent living. Higher-order behavioural patterns and the state of home infrastructures can be extracted by processing and combining the diverse sensor information.

BelAmI [9] aims at developing innovative technologies in the area of Ambient Intelligence. One of the contributions of this project is building a monitoring and remaindering solution that can address some characteristic requirements such as adaptively, dependability, interoperability, safety/security and resource efficiency. Objects' behaviours and Habit status will be observed and emergency assistance will be provided by enabling people to carry out daily tasks.

PAS [10] aims at maintaining the capability of independent living through a time-based daily activity reminder, non-intrusive monitoring of physical functions and mobility profiles, and more importantly maintaining a real-time communication with remote care providers and clinicians.

Since AAL has become a popular topic, many living labs for the study of ubiquitous technology in home settings have been established. Plenty of research efforts arise from either technological or medical fields, as it is extremely hard to develop any multidisciplinary access to this research realm. AAL requires not only technological or medical expertise but also attempts at including a social and cultural perspective to reach a high level of quality of life. One key observation concerns the predominance of indoor environments in many documented research efforts. In contrast, the work described here focuses on outdoor environments, focusing on the needs of both the patient and their carers.

3 Outdoor Healthcare System (OutCare) Overview

OutCare was designed as an AAL system specifically tailored for citizens with Dementia, for example, Alzheimer's Disease, and its design is based in County Carlow, Ireland. In line with the main symptoms of Alzheimer's disease, applying the intelligent agent platform, OutCare implements an intelligent outdoor assistance to provide remote and unobtrusive services to patient and efficient caring methods to the carers as well. The main functions implemented include:

1. Delivery and presentation of simple memory triggers for patients assisting in the repair of the activity sequence when performing simple routine tasks.
2. Adoption of an intelligent-agent based approach to facilitate distributed system intelligence and a modular system design.

3. Enabling a monitoring capability whereby patients' daily routines can be both learnt and deviations from such noted within a patient profile. Significant deviations from the normal daily signature could be sent as alerts to relevant parties (carers, relatives, medical-centres) via SMS, email, and voice.
4. Development of efficient and convenient caring mechanisms, as well as the collection of data that underpins longitudinal studies of the onset of such conditions and fosters a greater understanding of the condition itself.

3.1 OutCare Architecture

The general architecture of the OutCare system is shown in Fig. 1. There are three components involved in the system: the patient, the carer and a central web station.

– **Patients:** OutCare system is designed to support people with a cognitive impairment, particularly Alzheimer's. For this group of people, who are capable of maintaining basic independent living to some extent, may nonetheless suffer from memory loss, disability of performing complex daily activities and tasks, such as getting lost when walking.

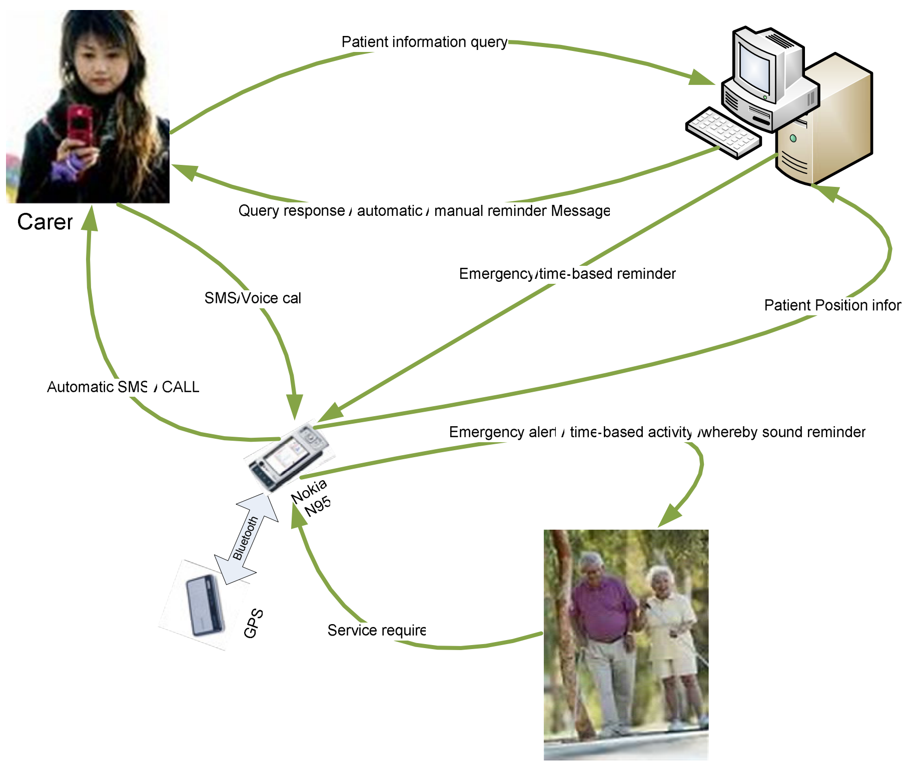


Fig. 1. OutCare System architecture

- **Carers:** people who are involved in taking care of the patients. The ever-increasing care demands result in the increasing requirement of a professional caring service.
- **Server station:** the central server which is responsible for patient profiling, data recording, processing, and visualization. The server station enables an efficient and trustworthy administrate tool, providing a convenient caring method.

3.2 OutCare System Intelligence

To facilitate system intelligence, the patient terminal within the OutCare system is implemented as a mobile multi-agent system [11], as the patient terminal is the start point which takes responsibility for collecting patient context information and boosts the system functions. It could be implemented on any type of mobile device as long as it supports 3G, Wi-Fi, Bluetooth and Java based applications. The Nokia N95 was identified as an archetypical mobile device although it incorporates a mobile device. In addition, it has considerable battery life and memory capacity. Usually the battery of Nokia N95 can last for a couple of days; the GPS unit allows 11 hours continuous use, which is adequate for daily tracking and allows charging occur overnight.

The intelligent agent platform is developed using Agent Factory Micro Edition (AFME) [12]. AFME is broadly based upon a pre-existing framework that support a structured approach to the development and deployment of agent-oriented application named Agent Factory [13]. AFME has been developed and widely used; it is actually supports the deployment of intelligent agents in AmI environment, especially for computationally restricted artefacts and environments. A collaboration of agents are deployed on the patient terminal, and the activities diagram is shown in Fig. 2.

- **GUI agent:** makes decisions about when and what information required to be displayed to the patient, and controls the graphical interface.
- **Device agent:** discovers the available Bluetooth devices, and manages the connection to the GPS device.
- **GPS agent:** collects the GPS position information of the relative patient.
- **Zone agent:** acquires GPS data from the GPS agent, analysis and calculates the patient context information, for example, what zone does the patient currently in, is she/he in a regular position, are they safe in the un-regular routine, etc.
- **Alert agent:** interprets the results from the zone agent, reasoning about what kind of alert is required, such as playing a sound or vibrating the mobile phone to get the patient's attention, and display relevant reminder message by the GUI agent to the patient.
- **Web agent:** updates patient profile information to the web server, such as patient’s current position if the GPS device is active, alert records etc.
- **SMS agent:** communicates with the zone agent and alert agent; decides when and what message needs to be notified to the relevant carer or parties.
- **Map agent:** helps patient to find the correct direction on the map.

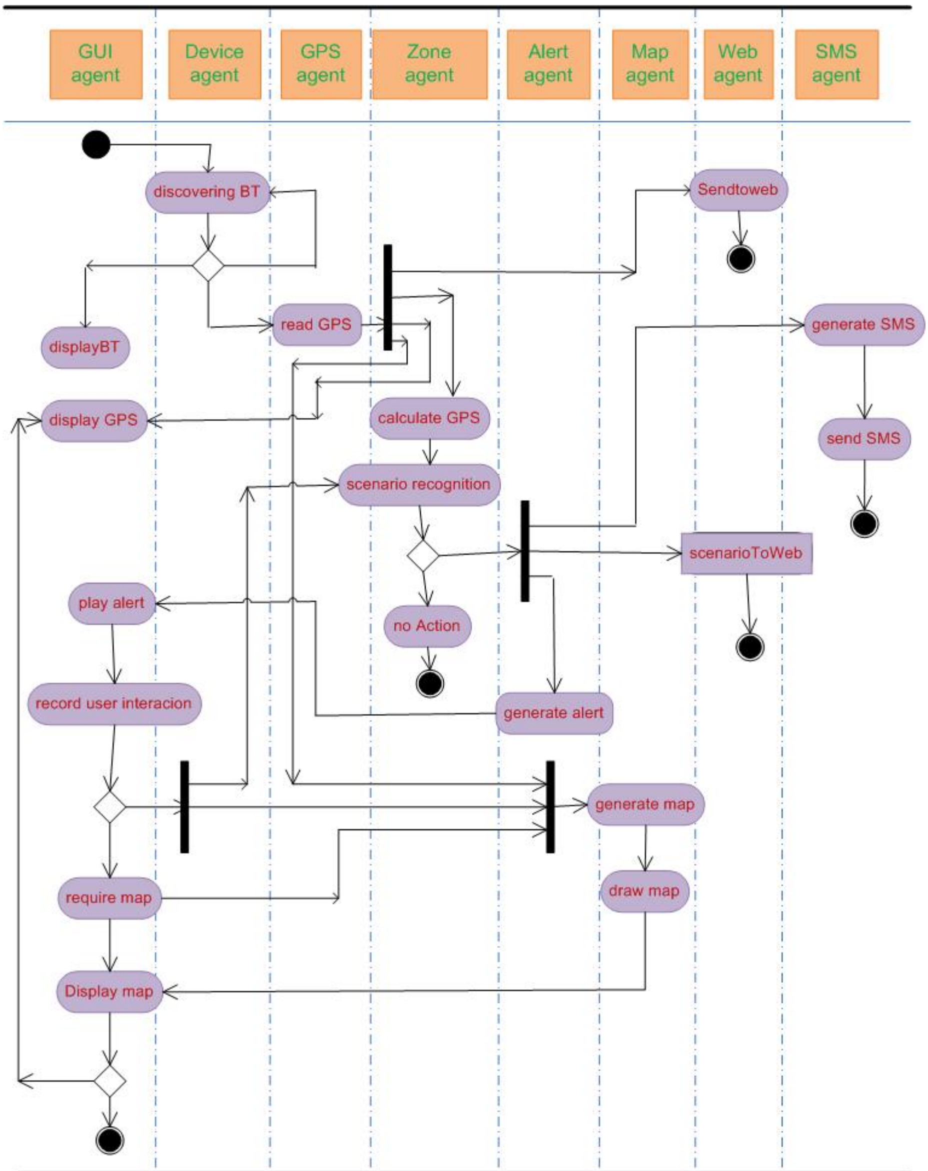


Fig. 2. Intelligent agent activity diagram

4 OutCare Design and Implementation

4.1 Patient Module

The patient module is designed to be embedded with the patient, two components are required - a mobile phone and a GPS receiver. This module harnesses

Fig. 3. Main interface (a), Alert example (b), and Location selection (c)

Fig. 4. Main screen (a), Alert example (b), and Routine query (c)

intelligentagents running on top of the J2ME platform. Due to the fact that patients are often afraid of using technology, and their cognitive impairment may varies according to the stage of Alzheimer's disease, OutCare adopts a user friendly and unobtrusive interface with no interactions required unless any irregular activity is detected. Fig.3 (a) is the service screen; Fig. 3 (b) is an alert example, and Fig. 3 (c) destination selection interface.

4.2 Carer Module

The aim of the carer module is to provide carers with an efficient and convenient caring method. The carer module is loaded onto the mobile device and its main functions include:

5. Receive emergency alerts from both the patient module and the web server (see Fig.4 (b)).

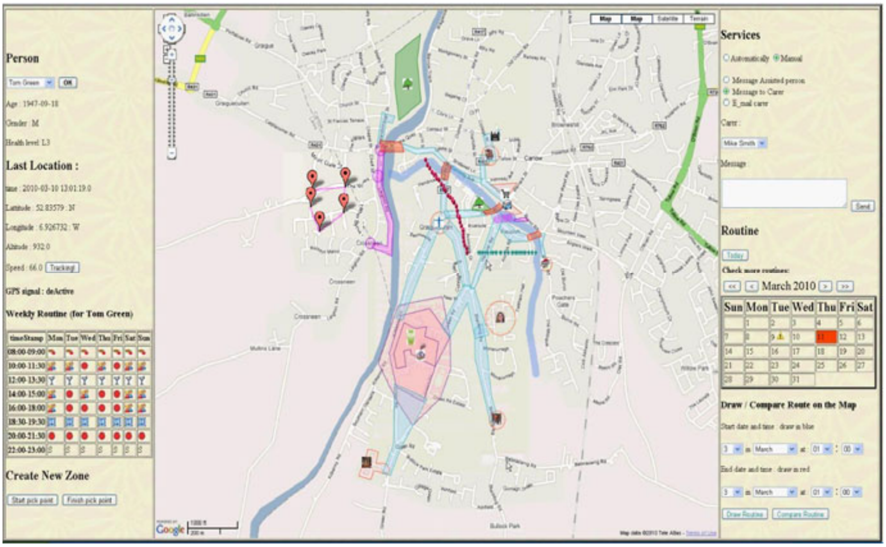


Fig. 5. Web server browser

6. Patient tracking and profile inquiry. Fig. 4 (a) shows the screen where OutCare can request patient's current location, patient tracking, query device information, view patient profile such as the alert history, collection records and so on. Fig. 4 (c) represents an example of a server response when the carer requests the patient routines for a given day.

4.3 Web Server Module

The web server is responsible for patient profiling, data recording, processing and analysing, information visualization and providing a convenient interface. Fig. 5 is the consolidated interface of the server browser. To implement patient monitoring functionality, we deploy patient profile information onto a map (Google maps, in this case), centred on Carlow.

Web server module plays an important role in the backstage of the OutCare system, with which the patient's daily movements can be recorded; real-time movement tracking, movement replay for a given person and time can be deployed on the map; analysis for the comparison of movement across several days according to individual patient's profile using overlays on the map; generate patient's regular routine; allow specification of alert activation conditions and recipient and contact modalities by SMS/Email.

5 Conclusion and Future Work

Given careful investigation and consideration of how to maintain the full capability of independent living for the elderly and particularly Alzheimer's sufferers, how their lives can be influenced by technologies, how to satisfy their needs not only physical but also social and personal, this paper presents an intelligent outdoor routine monitoring and assisting care system. Based on the AmI construct, the OutCare solution can maximise the utilisation of resources and deliver efficient monitoring and assisting services.

One of the novelties of OutCare comes from the location zone awareness and deliberate reasoning mechanism which maintains a patient's capability of independent living, respect of for their independence and dignity, as well as effective resource utilization. As OutCare has been implemented, user trials and evaluations is expected to commence in the near future. One important issue for future investigation is balancing patients' privacy with an efficient monitoring and assistance service. Another interesting issue concerns the integration of more sophisticated outdoor environmental models.

Acknowledgements. Jie Wan and Caroline Byrne gratefully acknowledge the support of the Institute of Technology, Carlow. This work is partially supported by Science Foundation Ireland under grant 07/CE/I1147.

References

1. Connell, P., Pringle, D.: Population Aging Ireland – Projections 2002 – 2021. National Council on Aging and Older People (2004)
2. O'Grady, M.J., Muldoon, C., Dragone, M., Tynan, R., O'Hare, G.M.P.: Towards Evolutionary Ambient Assisted Living Systems. Journal of Ambient Intelligence and Humanized Computing 1(1), 15–29 (2010)
3. The Alzheimer's Society of Ireland, `http://www.alzheimer.ie/eng/Media-Centre/Facts-About-Dementia`
4. Remagnino, P., Foresti, G.L.: Ambient Intelligence: A New Multidisciplinary Paradigm. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 35(1), 1–6 (2005)
5. Kleinberger, T., Becker, M., Ras, E., Holzinger, A., Müller, P.: Ambient Intelligence in Assisted Living: Enable Elderly People to Handle Future Interfaces. In: Stephanidis, C. (ed.) UAHCI 2007, Part II. LNCS, vol. 4555, pp. 103–112. Springer, Heidelberg (2007)
6. Russell, S.J., Norvig, P.: Artificial Intelligence: A Modern Approach, 2nd edn. Prentice Hall, New Jersey (2003)
7. AALIANCE, `http://www.aaliance.eu/public/news/aaliance-initiative-presents-roadmap-for-research-on-aal`
8. Ruyter, B.D., Pelgrim, E.: Ambient Assisted Living Research in CareLab. Interactions 14(4), 30–33 (2007)
9. Nick, M., Becker, M.: A Hybrid Approach to intelligent Living Assistance. In: Seventh International Conference on Hybrid Intelligent System, pp. 283–289. IEEE Computer Society, Washington (2007)
10. Zeng, Z., Yu, S., Shin, W., Hou, J.C.: PAS- A Wireless-Enabled, Cell-phone-incorporated Personal Assistant System for Independent and Assisted Living. In: Proceedings of the 28th International Conference on Distributed Computing Systems, pp. 233–242. IEEE Computer Society, Washington (2008)

11. Ferber, J.: Multi-Agent Systems: An Introduction to Artificial Intelligence. Addison-Wesley, Harlow (1999)
12. Muldoon, C., O'Hare, G.M.P., Collier, R., O'Grady, M.J.: Agent Factory Micro Edition: A Framework for Ambient Applications. In: Alexandrov, V.N., van Albada, G.D., Sloot, P.M.A., Dongarra, J. (eds.) ICCS 2006. LNCS, vol. 3993, pp. 727–734. Springer, Heidelberg (2006)
13. Muldoon, C., O'Hare, G.M.P., Collier, R.W., O'Grady, M.J.: Towards pervasive intelligence: reflections on the evolution of the agent factory framework. In: Multi-Agent Programming: Languages, Platforms and Applications, pp. 187–210. Springer, Berlin (2009)

Frame Semantics in Text-to-Scene Generation

Bob Coyne[1], Owen Rambow[1], Julia Hirschberg[1], and Richard Sproat[2]

[1] Columbia University, New York NY, USA
{coyne,julia,rambow}@cs.columbia.edu,
http://cs.columbia.edu/~coyne
[2] Oregon Health & Science University, Beaverton, Oregon, USA
rws@xoba.org

Abstract. 3D graphics scenes are difficult to create, requiring users to learn and utilize a series of complex menus, dialog boxes, and often tedious direct manipulation techniques. By giving up some amount of control afforded by such interfaces we have found that users can use natural language to quickly and easily create a wide variety of 3D scenes. Natural language offers an interface that is intuitive and immediately accessible by anyone, without requiring any special skill or training. The WordsEye system (http://www.wordseye.com) has been used by several thousand users on the web to create over 10,000 scenes. The system relies on a large database of 3D models and poses to depict entities and actions. We describe how the current version of the system incorporates the type of lexical and real-world knowledge needed to depict scenes from language.

1 Introduction

The work we describe seeks to bridge the gap between language, graphics, and knowledge by modeling the automatic conversion of text into a new type of semantic representation – a virtual 3D scene. 3D scenes provide an intuitive representation of meaning by making explicit the contextual elements implicit in our mental models. The system we are developing centers on a new type of lexical knowledge representation, which we call a Scenario-Based Lexical Knowledge Resource (SBLR). The SBLR will ultimately include information on the semantic categories of words; the semantic relations between predicates (verbs, nouns, adjectives, and prepositions) and their arguments; the types of arguments different predicates typically take; additional contextual knowledge about the visual scenes various events and activities occur in; and the relationship between this linguistic information and the 3D objects in our objects library. The resulting text-to-scene system, utilizing the SBLR, is applicable to several domains:

– **Education:** Seeing words spring to life makes using language fun and memorable. This suggests many uses in education including ESL, EFL, special needs learning, vocabulary building, and creative storytelling. We have done preliminary testing in some of these areas.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 375–384, 2010.

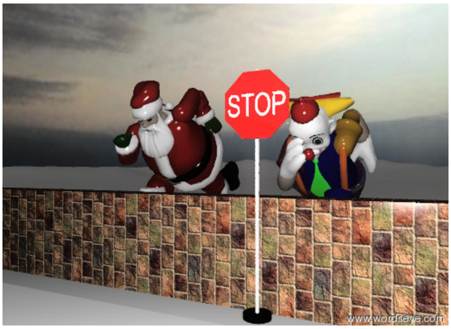
- **Graphics Authoring and Online Communication:** Textually generated scenes can be created in a fraction of the time it takes to make a scene by hand. This increase in speed and lower entry barrier enables a new form of social interaction and promotes "visual banter." We have seen much of this in our online website and gallery.
- **3D Games:** 3D games are painstakingly designed by 3D artists – the malleability of the graphics in games is usually limited to external level design tools and rigid interfaces for character selection and modification. Some recent games, such as Spore (`http://www.spore.com`), allow interesting variability of graphical elements; and spoken language commands are supported by games such as *Tom Clancy's EndWar* [14]. We foresee this trend continuing and games being developed where large parts of the game environment itself are interactively constructed and modified by the players using natural language.

In this paper we describe some of our recent work designing, building, and utilizing the SBLR in order to produce a system with much broader and robust coverage. Examples from the online system are shown in Figure 1. The remainder of the paper is organized as follows. In Section 2 we describe related work with natural language interfaces to 3D graphics. In Section 3 we describe experiences of the system with real online users as well as preliminary testing in schools. In Section 4 we provide an overview of the system. In Section 5 we discuss the SBLR, including the lexicon, semantic relations between lexical items, and frames which group complex relations together. We introduce the notion of a *vignette* and describe the graphical primitives that all semantic representations must resolve to.We conclude and describe future work in Section 6.

2 Related Work

Natural language input has been investigated in a number of very early 3D graphics systems [1][12][4][9] and the Put system [5], which was limited to spatial arrangements of existing objects in a pre-constructed environment. Also, input was restricted to an artificial subset of English consisting of expressions of the form $Put(X, P, Y)$, where X and Y are objects and P is a spatial preposition. Several more recent systems target animation rather than scene construction. Work at the University of Pennsylvania's Center of Human Modeling and Simulation [2] used language to control animated characters in a closed virtual environment. CarSim [7] is domain-specific system where short animations were created from natural language descriptions of accident reports. CONFUCIUS [11] is a multimodal animation system that takes as input a single sentence containing an action verb. The system blends animation channels to animate virtual human characters. Another recent system, from the University of Melbourne [16], uses a machine learning-based approach to create animated storyboards on a pre-made virtual stage. In these systems the referenced objects, attributes, and actions are typically relatively small in number or targeted to specific pre-existing domains.

Redundancy by Bob Coyne



Santa Claus is on the white mountain range. He is racing. The clown is 2 feet to the right of santa claus. The clown is racing. A brick wall is 2 feet in front of the clown. The wall is 20 inches tall. It is 20 feet wide. A small stop sign is in front of the wall. It is cloudy.

Infinite Time by Richard Sproat

The clock is one foot in front of the silver wall. The ground has a grass texture. The texture is one foot wide. A silver wall is two feet in front of the clock. A light is fifty feet above the clock.

Fig. 1. Some Examples

As such, these systems have a natural affinity to the SHRDLU system [15] which used natural language to interact with a “robot” living in a closed virtual world.

Our current system is a rewrite and enhancement of the original version of WordsEye [6], which was the first system to use a large library of 3D objects in order to depict scenes in a more general and free-form manner using natural language. The current system contains 2,000 3D objects and 10,000 images and a lexicon of approximately 15,000 nouns. It supports language-based control of spatial relations, textured and colors, collections, and poses; and it handles simple anaphor resolution, allowing for a variety of ways of referring to objects. The earlier WordsEye system handled 200 verbs in a *ad hoc* manner with no systematic semantic modeling of alternations and argument combinations. In the current system, we are instead adding frame semantics to support verbs, event nouns, and stative relations more robustly. The system also does very little inferring of background locations, default poses of characters, and other contextual features – everything must be stated fairly explicitly. As a result, users must describe scenes in somewhat stilted language. But even with these limitations, the process of creating scenes with the system is quick and enjoyable.

3 User Experiences

Earlier versions of the current WordsEye system have been tested online (`http://www.wordseye.com`) over a several year period. A few thousand real-world users have used the system to create and post a large number textually-generated pictures to our online gallery, and in some cases to their personal webpages and social media sites such as Facebook. The ease and speed of creating and modifying scenes has led to pictures being used as a form of social interaction. We also found, in the course of our testing, that users would try to use the system to do more than it is capable of. In particular they would use language that involved a) complex background settings (e.g., *living room*, *garden*) consisting of many objects b) characters performing actions c) modifications to parts of objects (*the girl's hair is curly*). Our work on the SBLR, to allow the system to better support these areas, has grown out of these experiences.

We also performed some preliminary testing of the system in schools in Spring 2007. After seeing a demo of WordsEye at the Innovate 2007 Exposition (hosted by the State of Virginia Department of Education), K-12 public school teachers from the Albemarle county school system in Virginia asked if they could use it in their classes, believing it to be a useful tool for ESL (English as a second language) remediation, special education, vocabulary enhancement, writing at all levels, technology integration, and art. Feedback from these teachers and their students was quite positive. In one school, with a 10% ESL population, a teacher used it with 5th and 6th graders to reinforce being specific in details of descriptive writing and noted that students are "very eager to use the program and came up with some great pictures." Another teacher tested it with 6th through 8th grade students who were "in a special language class because of their limited reading and writing ability," most reading and writing on a 2nd or 3rd grade level. In addition to its educational value, students found the software fun to use, an important element in motivating learning. As one teacher reported, "One kid who never likes anything we do had a great time yesterday...was laughing out loud."

4 System Overview

WordsEye consists of multiple components, including a user interface, language processing and 3D graphics. In this paper we focus on the lexical and world knowledge used to convert dependency structures into semantic nodes and roles and the subsequent conversion to graphic frames. The overall system works in the following sequence:

- The user types in text to a webpage.
- The input text is parsed into a dependency tree.
- Anaphora and other coreferences are resolved.
- Lexical items and dependency links are resolved to semantic nodes and roles.
- Semantic relations are converted to graphical relations.
- Default graphically-oriented constraints are inserted (such as putting objects on the ground unless otherwise specified).
- The scene is composed from these constraints and rendered in OpenGL (`http://www.opengl.org`) and optionally ray-traced in the Radiance [10] renderer (`http://radsite.lbl.gov/radiance`).
- The user can provide a title and caption to the finished scene and save it in an online gallery where other users can add comments and create their own pictures in response.

5 SBLR – Lexical, Semantic, and Contextual Information

The SBLR contains lexical, semantic, and contextual information. As such it is related to WordNet and FrameNet, and is, in fact, partially derived from them. We first examine some of the features and limitations of WordNet and FrameNet.

5.1 Wordnet

The WordNet ontology provides a taxonomy of words grouped into separate SYNSETS by word sense and related primarily by HYPERNYM (IS-A) relations. This provides useful information such as the fact that a *chair* is *furniture*. The taxonomy fails, however, to provide a rich set of semantic relations between those entries. For example, the WordNet synset for *princess* is a hyponym of *aristocrat*, but there is no encoding of the basic fact that a princess is also a female. Likewise, WordNet often conflates lexical usage with functional semantic categories. For example, *spoon* is classified as a *container*, which is true in some sense; however, it does not match normal colloquial usage, since a spoon is unlikely to be considered a container by typical speakers of English. Also, while WordNet does encode some part-whole and substance-whole relations, it is missing many very common ones, such as the fact that lamps have lightbulbs and that snowballs are made of snow. In addition, there is no encoding of functional properties of objects such as the fact that a mop is used in cleaning floors. A broader set of semantic relations is crucial to the resolution of lexical references in scene construction.

5.2 Framenet

FrameNet is a digital lexical resource for English that groups related words together into semantic frames [3]. It currently contains 10,000 lexical entries including nouns, verbs, and adjectives. Each lexical unit is associated with one of nearly 800 hierarchically-related semantic frames, where each frame represents the joint meaning of the lexical units in that frame. Each lexical unit is also associated with a set of annotated sentences which map the sentences' constituent parts to their frame-based roles. FrameNet, in total, contains over 135,000 annotated sentences across all lexical units. A FrameNet frame consists of a set of frame-based roles, called *frame elements* (FEs) representing the key roles characterizing the meaning of lexical units in that frame. For example, the COMMERCE_SELL frame includes frame elements for SELLER, GOODS, and BUYER and has lexical units for the verbs *retail*, *sell*, *vend* as well as nouns such as *vendor* and *sale*.

The exact expression of FEs for a given annotated sentence constitutes what FrameNet refers to as a *valence pattern*. Valence patterns are represented as FE and grammatical function (GF) pairs. Grammatical functions are *subject*, *obj*, *second object*, and various other dependent phrases (e.g., *Dep/to*, *Dep/on*, *Dep/with*) which designate the particular prepositional phrase type.

FrameNet provides no semantic information to distinguish the meaning of words in the same frame. For example, the SELF_MOTION frame contains a large number of verbs related only by the fact that the SELF_MOVER moves under its own power in a directed fashion without a vehicle. As a result, this frame contains strongly related verbs such as *walk* and *stroll* but also verbs with very different manner of motion such as *swim*, and *swing*. Likewise there is no representation in FrameNet of synonymy, antinomy, or other lexical semantic relations.

5.3 SBLR

The SBLR currently contains about 15,000 nouns representing the 3D objects in our library and related words, including a set of semantic relations seeded from FrameNet and augmented as needed. The lexicon was semi-automatically extracted from WordNet, filtering out obscure words and word senses. The resulting lexicon is defined with multiple inheritance to allow the same lexical item to be classified in several ways. In addition, as needed, we encode semantic relations between lexical items via the corresponding SBLR semantic relation. So, to say that a snowball is made of snow, we use the MADE-OF frame.

In addition to their role in verbal adjuncts, prepositions are especially important in text-to-scene generation, since they directly denote spatial relations in a variety of subtle ways [8]. In the SBLR we extend the coverage of frames and lexical units to better model the range of preposition senses. So, for example, *John is on the sidewalk* and *John is on the phone* will constitute different senses of *on* and hence be in different frames. Likewise, *The flower is in the vase*, *The spider is in the web*, and *The goldfish is in the aquarium* represent different, though related, spatial interpretations of *in*.

In the SBLR we also extend FrameNet's notion of valence pattern to directly include semantic and contextual constraints (including SELECTIONAL RESTRICTIONS), drawing upon the semantic types of words and their semantic and contextual relations to other words as defined in the rest of the SBLR. This allows the appropriate frames and frame elements to be assigned to parsed input text. Consider, for example, a few of the semantic interpretations of the preposition *of* and how they are handled by the system in Table 1 and Figure 2.

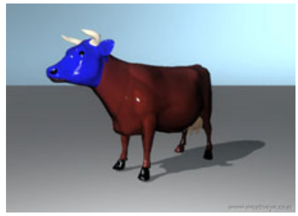
In order to decompose semantic representations into scenes, we need to supply default instruments, settings, poses, and so on. These choices are captured in the notion of a **vignette**. A vignette is a mapping from a frame-semantic representation to the graphical relations that invoke that scene. For example, the verb *wash* has very different meanings depending on the arguments of the verb. Washing a car takes place outside, often in a driveway with the AGENT

Table 1. SBLR: Semantic Mappings for *of*

Text (A of B)	**Valence Patterns for *of***	**Resulting Frame Relation**
bowl of cherries	A=container, B=plurality-or-mass	CONTAINER-OF(bowl, cherries)
slab of concrete	A=entity, B=substance	MADE-OF(slab, concrete)
picture of girl	A=representing-entity, B=entity	REPRESENTS(picture, girl)
arm of the chair	A=part-of (B), B=entity	PART-OF (chair, arm)
height of the tree	A=size-property, B=physical-entity	DIMENSION-OF(height, tree)
stack of plates	A=arrangement, B=plurality	GROUPING-OF (stack, plates)

MADE-OF: *horse of stone*

GROUPING-OF: *stack of cats*



PART-OF: *Head of cow*

DIMENSION: *height of horse*

REPRESENTS: *picture of girl*

CONTAINER-OF: *bowl of cats*

Fig. 2. Depictions of *of*

holding a hose and standing near to the car. Washing dishes usually takes place in a kitchen with the AGENT standing in front of the sink, holding the dishes. Washing the floor likewise invokes yet another prototypical scene. We now consider, in more detail, a simple example: *the truck chased the soldier down the road*.

Valence pattern triggered for *The truck chased the soldier down the road*:

- Frame=COTHEME, FEs=AGENT, COTHEME, PATH, SOURCE, GOAL, ... (This frame contains words that necessarily indicate the motion of two distinct objects.)
- Verb="chase"
- Subject=AGENT, direct-object=COTHEME, dep/down=PATH

Vignette triggered by the above input generates graphical relations putting the soldier (COTHEME) and truck (THEME) on the road (PATH) with the soldier in a running pose in front of the truck:

- Orient-towards: object=THEME, reference=PATH
- Orient-towards: object=COTHEME, reference=PATH
- Position-behind: object=THEME, reference=COTHEME
- Position-on-top: object=THEME, reference=PATH
- Position-on-top: object=COTHEME, reference=PATH
- In-pose: object=COTHEME, pose="running pose"

Figure 3 shows the rendered scene generated using these valence patterns and vignettes.

a) The truck chased the soldier down the road...

b) The soldier ran across the sidewalk...

Fig. 3. Scenes derived from SBLR valence patterns and vignettes. The valence patterns for *chase* and *run* are used to assign semantic frame roles. The semantic relation is then mapped to the appropriate vignettes based on which roles are filled with what. The Vignettes resolve to spatial relations such a) as the COTHEME (*soldier*) being in a running pose and located on the PATH and in front of the THEME (truck) in the case of *chase down* or b) the SELF_MOVER perpendicularly oriented toward the PATH and in a running pose in the case of *run across*. These are combined with explicitly described backgrounds and other objects. In the future, the background settings for different actions will also be defaulted by vignettes.

5.4 Graphical Objects and Relations

Graphical objects and relations are the bedrock to which all semantics must be resolved in order to be depicted. Graphical objects are inserted into the ontology with ISA links referencing existing semantic nodes. This allows them to inherit property values for other objects of their type (such as default size). In addition, almost all 3D objects implicitly contain subtype information (e.g., *dining room chair*, or *antique chinese vase*). And sometimes, 3D objects are compound objects (a lighthouse on a hill) or part of another object (a rose blossom). In all these cases, these properties are represented by semantic relations drawn from our stock of semantic frames.

Graphical objects have various additional geometric and functional properties. These include spatial tags (used in resolving target locations for spatial relations), default size and orientation, and parts. Some of these spatial tags and attributes are described in [6]. The graphical relations which are used in creating a 3D scene include object size, color, position, orientation, texture, aspect ration, facial expressions and poses (for human characters) among others.

The information collected in the SBLR is coming from our own existing resources; external semantic resources such as WordNet, FrameNet, and PropBank, which we are mining for additional information; and information which we will extract from Wikipedia and other corpora . The overall system architecture and the role played by the SBLR is shown in Figure 4.

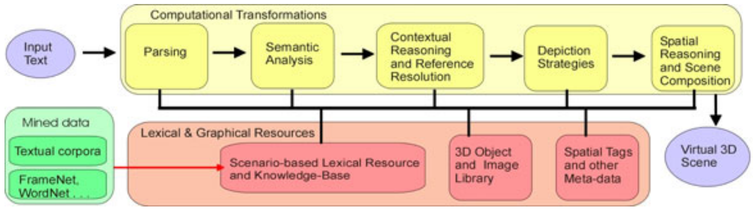


Fig. 4. System Architecture

6 Conclusions and Future Work

The system as described is a work-in-progress. Much remains to be done at the language, graphical, and application levels. Knowledge acquisition, representation, and utilization is our core ongoing task. We are acquiring contextual information such as part relations, default backgrounds for actions, and lexical constraints and verb arguments using Amazon's Mechanical Turk and automatic methods [13]. We have plans in place to evaluate our software in partnership with a non-profit after-school program in New York CIty.

We believe that textually generated scenes offer an exciting new way to interact with a computer and to create visual imagery. It affords access to a wider variety of people than who might otherwise be able to make pictures. And it allows picture-making to be done in completely new settings due to the speed and low overhead involved.

Acknowledgments

This work was supported in part by the NSF IIS- 0904361. Any opinions, findings and conclusions or recommendations expressed in this material are the authors' and do not necessarily reflect those of the sponsors. We thank Steve Feiner for providing references to some early text-to-scene work and Fadi Biadsy and anonymous reviewers for their useful comments. And we thank the many online WordsEye users for their help in testing and validating this new method of creating 3D scenes and interacting socially with words and pictures.

References

1. Adorni, G., Di Manzo, M., Giunchiglia, F.: Natural language driven image generation. In: COLING, pp. 495–500 (1984)
2. Badler, N., Bindiganavale, R., Bourne, J., Palmer, M., Shi, J., Schule, W.: A parameterized action representation for virtual human agents. In: Workshop on Embodied Conversational Characters, Lake Tahoe (1998)
3. Baker, C., Fillmore, C., Lowe, J.: The Berkeley FrameNet Project. In: COLING-ACL (1998)

4. Boberg, R.: Generating Line Drawings from Abstract Scene Descriptions. Master's thesis, Dept. of Elec. Eng., MIT, Cambridge, MA (1972)
5. Clay, S.R., Wilhelms, J.: Put: Language-based interactive manipulation of objects. In: IEEE Computer Graphics and Applications, pp. 31–39 (1996)
6. Coyne, B., Sproat, R.: WordsEye: An automatic text-to-scene conversion system. In: SIGGRAPH, Computer Graphics Proceedings, pp. 487–496 (2001)
7. Dupuy, S., Egges, A., Legendre, V., Nugues, P.: Generating a 3d simulation of a car accident from a written description in natural language: The carsim system. In: Proceedings of ACL Workshop on Temporal and Spatial Information Processing, pp. 1–8 (2001)
8. Herskovits, A.: Language and Spatial Cognition: an Interdisciplinary Study of the Prepositions in English. Cambridge University Press, Cambridge (1986)
9. Kahn, K.: Creation of Computer Animation from Story Descriptions. Ph.D. thesis, MIT, AI Lab, Cambridge, MA (1979)
10. Larson, G., Shakespeare, R.: Rendering with Radiance. The Morgan Kaufmann Series in Computer Graphics (1998)
11. Ma, M.: Automatic Conversion of Natural Language to 3D Animation. Ph.D. thesis, University of Ulster (2006)
12. Simmons, R.: The clowns microworld. In: Proceedings of TINLAP, pp. 17–19 (1998)
13. Sproat, R.: Inferring the environment in a text-to-scene conversion system. In: First International Conference on Knowledge Capture, Victoria, BC (2001)
14. WikiPedia: Tom Clancy's EndWar, `http://en.wikipedia.org/wiki/Tom_Clancy's_EndWar`
15. Winograd, T.: Understanding Natural Language. Ph.D. thesis, Massachusetts Institute of Technology (1972)
16. Ye, P., Baldwin, T.: Towards automatic animated storyboarding. In: Proceedings of the 23rd National Conference on Artificial Intelligence, vol. 1, pp. 578–583 (2008)

SenticSpace: Visualizing Opinions and Sentiments in a Multi-dimensional Vector Space

Erik Cambria[1], Amir Hussain[1], Catherine Havasi[2], and Chris Eckl[3]

[1] Dept. of Computing Science and Maths, University of Stirling, Scotland, UK
[2] MIT Media Lab, MIT, Massachusetts, USA
[3] Sitekit Labs, Sitekit Solutions Ltd, Scotland, UK
{eca,ahu}@cs.stir.ac.uk, havasi@media.mit.edu, chris.eckl@sitekit.net
http://cs.stir.ac.uk/~eca/sentics

Abstract. In a world in which millions of people express their feelings and opinions about any issue in blogs, wikis, fora, chats and social networks, the distillation of knowledge from this huge amount of unstructured information is a challenging task. In this work we build a knowledge base which merges common sense and affective knowledge and visualize it in a multi-dimensional vector space, which we call SenticSpace. In particular we blend ConceptNet and WordNet-Affect and use dimensionality reduction on the resulting knowledge base to build a 24-dimensional vector space in which different vectors represent different ways of making binary distinctions among concepts and sentiments.

Keywords: Sentic Computing, AI, Semantic Networks, NLP, Knowledge Base Management, Opinion Mining and Sentiment Analysis.

1 Introduction

Opinion mining and sentiment analysis is a new discipline which has recently raised more and more interest, especially in fields such as marketing, personal affective profiling and financial market prediction. Given a generic resource (textual, visual, audio or multimodal) containing a set of opinions O about a set of topics T with different polarity $p \in$ [-1,1], we define opinion mining as the process which aims to extract, for each $t \in T$, the subset of opinions $o \subseteq O$ concerning t and determine p.

Existing approaches to automatic identification and extraction of opinions and sentiments from text, in particular, can be grouped into four main categories: keyword spotting [1][2][3], in which text is classified into categories based on the presence of fairly unambiguous affect words, lexical affinity [4][5], which assigns arbitrary words a probabilistic affinity for a particular opinion or emotion, statistical methods [6][7], which consist in calculating the valence of keywords, punctuation and word co-occurrence frequencies on the base of a large training corpus, and hand-crafted models [8].

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 385–393, 2010.

The problem with these approaches is that they mainly rely on parts of text in which attitudes and affective states are explicitly expressed such as positive terms (e.g. good, excellent, superior), negative terms (e.g. bad, poor, wrong) or verbs, adjectives and adverbs of emotion (e.g. to love/to hate, angry/pleased, happily/sadly). Opinions and sentiments, in fact, are more often expressed implicitly through concepts with an affective valence such as 'play a game', 'be laid off' or 'go on a first date'.

Sentic Computing [9] overcomes this problem by using a semantic reasoning approach coupled with a novel emotion categorization born from the idea that our mind consists of four independent emotional spheres, whose different levels of activation make up the total emotional state of the mind.

2 Sentic Computing

The capability of perceiving and expressing emotions is a fundamental component in human experience, cognition, perception, learning and communication.

Today conventional computers lack this kind of skill. They just do what they are programmed to do without caring at all if the user is experiencing fascination or frustration. This is why nowadays we have plenty of programs that exceed the capabilities of world experts but are not able to do what even a puppy can do – understand us from an emotional point of view.

Sentic Computing, whose term derives from the Latin 'sentire' (root of words like sentiment and sensation), is a new opinion mining and sentiment analysis paradigm which exploits AI and Semantic Web techniques to better recognize, interpret and process opinions and sentiments in web-posts i.e. short texts over the Web such as blog and forum entries, RSS feeds, tweets and instant messages.

Differently from statistical affective classification, which generally requires large inputs and thus cannot appraise texts with satisfactory granularity, Sentic Computing allows to analyze web-posts not only on the page or paragraph-level but even on sentence-level. The affective categorization, in fact, is not based on statistical learning models but rather on an affective semantic network (section 2.1) and a novel emotion categorization model (section 2.2).

2.1 AffectNet

When people communicate with each other, they rely on similar background knowledge e.g. the way objects relate to each other in the world, people's goals in their daily lives, the emotional content of events or situations. This 'taken for granted' information is what we call common sense [10] – obvious things people normally know and usually leave unstated.

The Open Mind Common Sense project has been collecting this kind of knowledge from volunteers on the Internet since 2000 to provide intuition to AI systems and applications. ConceptNet [11] represents the information in the Open Mind corpus as a directed graph in which the nodes are concepts and the labeled edges are assertions of common sense that interconnect them.

WordNet-Affect [12] is a linguistic resource for the lexical representation of affective knowledge, developed starting from WordNet [13]. The ontology is built by assigning to a number of WordNet synsets one or more affective labels (a-labels) and then by extending the core with the relations defined in WordNet.

In particular, the affective concepts representing emotional states are identified by synsets marked with the a-label 'emotion', but there are also other a-labels for concepts representing moods, situations eliciting emotions or emotional responses.

The blend of ConceptNet and WordNet-Affect is what we call AffectNet – an affective semantic network in which common sense concepts are linked to a hierarchy of affective domain labels (Fig. 1).

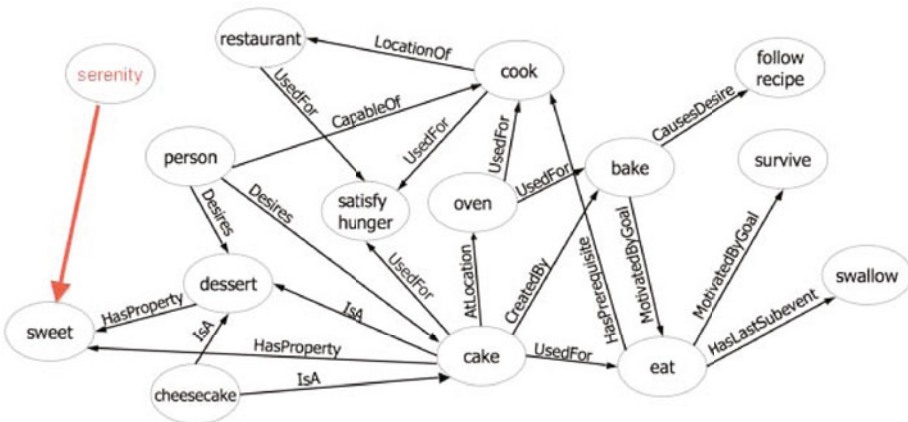


Fig. 1. A sketch of AffectNet

Blending [14] is a technique that performs inference over multiple sources of data simultaneously, taking advantage of the overlap between them. It basically combines two sparse matrices linearly into a single matrix in which the information between the two initial sources is shared.

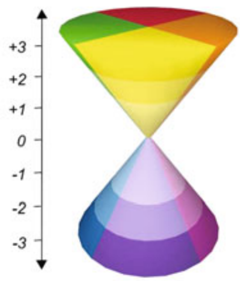
The first step to create a blend is to transform the input data so that it can all be represented in the same matrix. To do this we align the lemma forms of ConceptNet concepts with the lemma forms of the words in WordNet-Affect and map the most common relations in the affective knowledge base into ConceptNet's set of relations, e.g. Hypernym into IsA and Holonym into PartOf.

This alignment operation yields a new semantic network in which common sense and affective knowledge are in fact combined, not just concomitant, i.e. a network in which everyday life concepts like 'have breakfast', 'meet people' or 'watch tv' are linked to affective domain labels like 'joy', 'anger' or 'surprise'.

2.2 The Hourglass of Emotions

This model is a variant of Plutchik's emotion categorization [15] and constitutes an attempt to emulate Marvin Minsky's conception of emotions. Minsky sees the mind as made of thousands of different resources and believes that our emotional states result from turning some set of these resources on and turning another set of them off [16].

Each such selection changes how we think by changing our brain's activities: the state of anger, for example, appears to select a set of resources that help us react with more speed and strength while also suppressing some other resources that usually make us act prudently. The Hourglass of Emotions (Fig. 2) is specifically designed to recognize, understand and express emotions in the context of human-computer interaction (HCI).



	Pleasantness	Attention	Sensitivity	Aptitude
+3	ecstasy	vigilance	rage	admiration
+2	joy	anticipation	anger	trust
+1	serenity	interest	annoyance	acceptance
0	—	—	—	—
-1	pensiveness	distraction	apprehension	boredom
-2	sadness	surprise	fear	disgust
-3	grief	amazement	terror	loathing

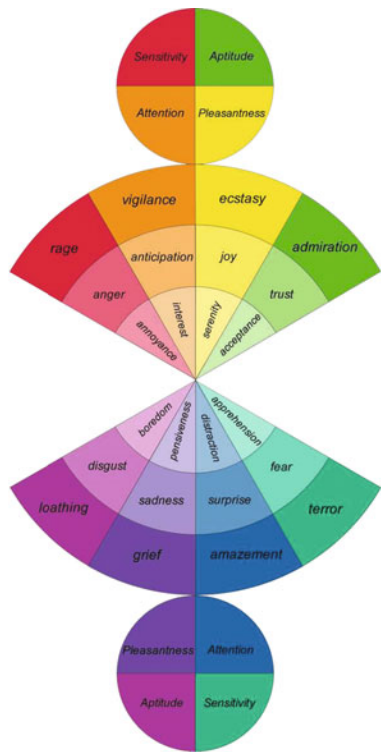


Fig. 2. The Hourglass of Emotions

In the model, in fact, affective states are not classified, as often happens in the field of emotion analysis, into basic emotional categories, but rather into four concomitant but independent dimensions – Pleasantness, Attention, Sensitivity and Aptitude – in order to understand how much respectively:

1. the user is happy with the service provided
2. the user is interested in the information supplied
3. the user is comfortable with the interface
4. the user is disposed to use the application

Each of the 4 affective dimensions is characterized by 6 levels of activation, called 'sentic levels', which determine the intensity of the expressed/perceived emotion, for a total of 24 labels specifying the 'elementary emotions'. The concomitance of the different affective dimensions makes possible the generation of 'compound emotions' such as love, which is given by the sum of positive values of Pleasantness and Aptitude, or aggressiveness, given by the concomitance of Attention and Sensitivity.

3 Visualizing Opinions and Sentiments

The process for the inference of opinions and sentiments from text numbers three main components: a NLP module, which performs a first skim of the document, a Semantic Parser, whose aim is to extract concepts from the stemmed text, and SenticSpace, a vector space in which different vectors represent different ways of making binary distinctions among concepts and sentiments.

The NLP module interprets all the affective valence indicators usually contained in text such as special punctuation, complete upper-case words, onomatopoeic repetitions, exclamation words, negations, degree adverbs and emoticons. The Semantic Parser then deconstructs text into concepts and provides, for each of them, the relative frequency, valence and status i.e. the concept's occurrence in the text, its positive or negative connotation, and the degree of intensity with which the concept is expressed.

SenticSpace is a multi-dimensional vector space obtained by applying principal component analysis (PCA) [17] on AffectNet. It is a language visualisation system which transforms natural language from a linguistic form into a 24D real-time navigable space.

After performing truncated singular value decomposition (TSVD) on AffectNet, we obtain a new matrix $\tilde{A} = U_k * \Sigma_k * V_k^T$, which forms a low-rank approximation of A, the original data. This approximation is based on minimizing the Frobenius norm of the difference between A and $\tilde{A}$ under the constraint $rank(\tilde{A}) = k$. For the Eckart–Young theorem [18] it represents the best approximation of A in the mean-square sense, in fact:

$$\min_{\tilde{A}|rank(\tilde{A})=k} \|A - \tilde{A}\| = \min_{\tilde{A}|rank(\tilde{A})=k} \|\Sigma - U^* \tilde{A} V\| = \min_{\tilde{A}|rank(\tilde{A})=k} \|\Sigma - S\| \qquad (1)$$

assuming that $\tilde{A}$ has the form $\tilde{A} = USV^*$, where S is diagonal. From the rank constraint, i.e. S has k non-zero diagonal entries, the minimum of the above statement is obtained as follows:

$$\min_{s_i} \sqrt{\sum_{i=1}^{n} (\sigma_i - s_i)^2} = \min_{s_i} \sqrt{\sum_{i=1}^{k} (\sigma_i - s_i)^2 + \sum_{i=k+1}^{n} \sigma_i^2} = \sqrt{\sum_{i=k+1}^{n} \sigma_i^2} \qquad (2)$$

Therefore, $\tilde{A}$ of rank k is the best approximation of A in the Frobenius norm sense when $\sigma_i = s_i$ $(i = 1, ..., k)$ and the corresponding singular vectors are same as those of A.

By exploiting the information sharing property of TSVD, concepts with the same affective valence are likely to have similar features i.e. concepts concerning the same sentiment tend to fall near each other in the vector space.

We use a k-means approach to clusterize the resulting space according to the sentic levels of the Hourglass model (Fig. 2). To finally obtain SenticSpace we then reorganize the retrieved affective concepts (i.e. the concepts belonging to at least one cluster) as rows of a new matrix in which the columns are the 24 sentic levels and the entries are the relative distances, i.e. the dot products, between concepts and sentic levels.

In case a concept belongs to more than one cluster at the same time it is assigned to the closest centroid, unless the clusters belong to different affective dimensions (according to the Hourglass model, in fact, affective concepts can convey emotions in more than one affective dimension at the same time).

This operation yields a multi-dimensional vector space in which concepts are represented by vectors of 24 coordinates: these coordinates describe concepts in terms of the Hourglass sentic levels and form the axes of SenticSpace i.e. the basis e_0,...,e_{23} of the vector space. We use Processing, an open source programming language and environment to program animation and interactions, to visualize the space (Fig. 3).

Depending on the cluster they belong to, concepts in SenticSpace are given a different color i.e. the color of the corresponding sentic level of the Hourglass model (Fig. 2). This way it is easy to visualize and compare different concepts according the emotions they convey and hence use this information to semantically analyze opinions and sentiments expressed in natural language documents.

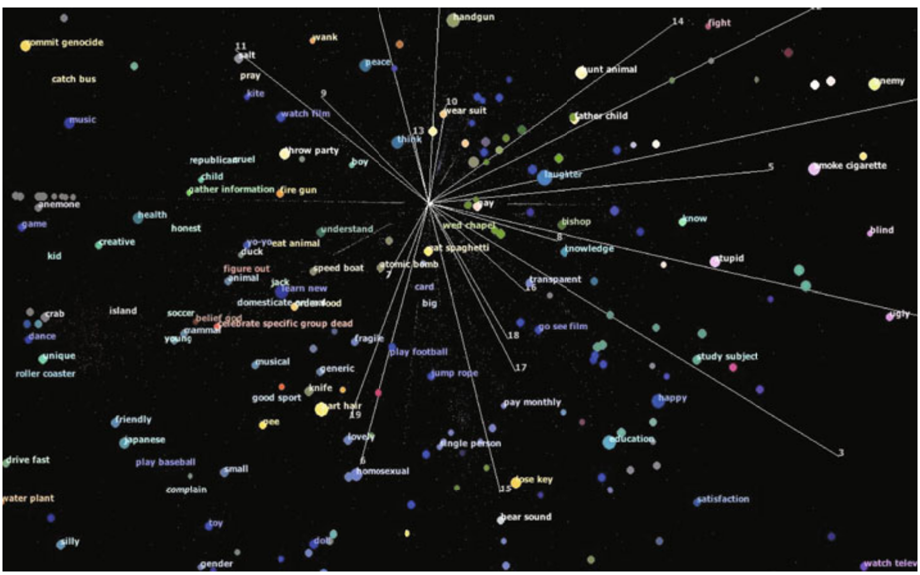


Fig. 3. A sketch of SenticSpace

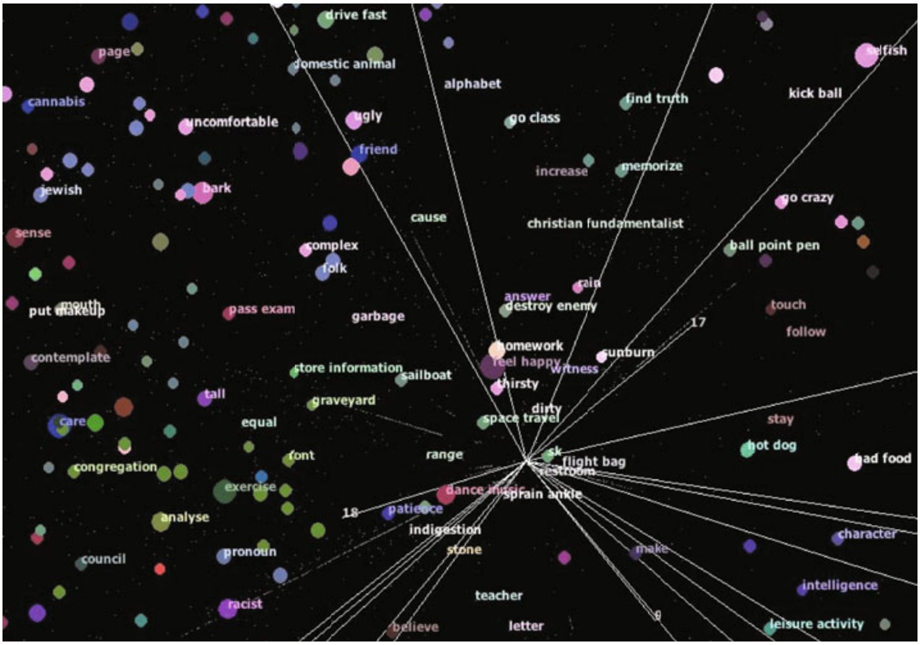


Fig. 4. Navigating SenticSpace

4 Evaluation

An evaluation of the process was performed by considering a corpus of blog-posts from LiveJournal (LJ), a virtual community where web-users can keep a blog, journal or diary. One of the interesting features of this web-site is that LJ bloggers, who number over 23 millions, are allowed to label their posts with a mood tag, by choosing from more than 130 predefined moods or by creating custom mood themes.

Since the indication of the affective status is optional, the mood-tagged posts are likely to reflect the true mood of the authors, and hence form a good test-set for Sentic Computing. Unfortunately there is no possibility to get mood-tagged blog-posts via data feeds so we had to design our own crawler.

After extracting the useful data and meta-data from each post, the crawler selected just the mood-tagged ones and stored them in an ad-hoc database. For each post in the database we extracted the content and gave it as input to the process. The output was then compared with the mood retrieved from the database to calculate statistical classifications such as precision and recall.

In particular we crawled and processed 5,000 web-posts: on average the posts' length was 242 words, the NLP Module detected 7 affective valence indicators per post and the Semantic Parser extracted 32 affective concepts for each post. According to the values of the concepts' relative distance in SenticSpace, we categorized each post and compared the results with the corresponding LJ tags:

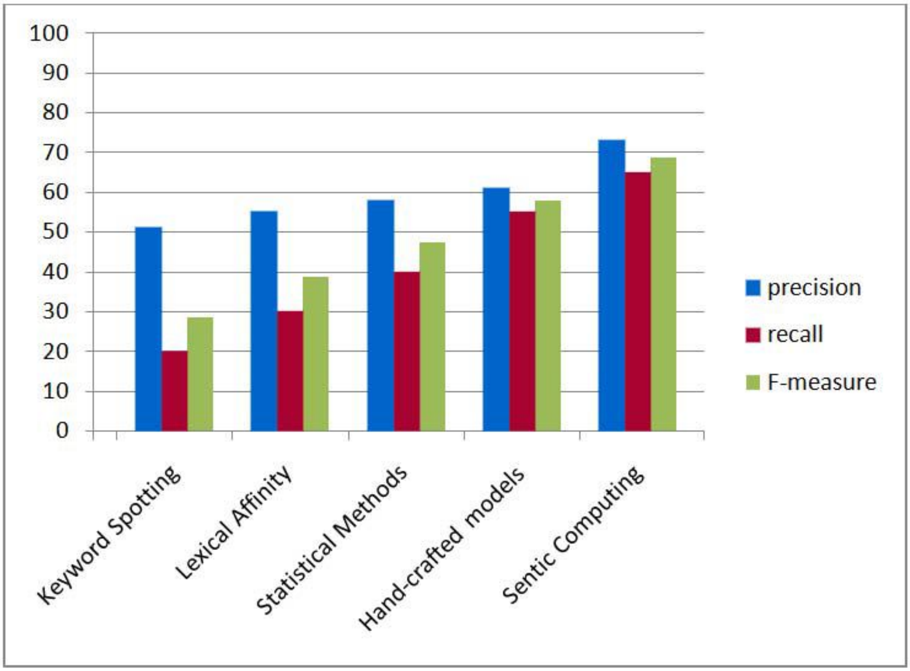


Fig. 5. Comparison of affective categorization methods

the system showed a very high precision (73%) and significantly good recall and F-measure rates (65% and 68% respectively), outperforming the baseline methods. The affective categorization was particularly good, in terms of precision, for positive moods, probably due to the fact that the Open Mind corpus contains more concepts expressing positive emotions than concepts related to sad affective states. Moreover, since the most significant eigenmood represents concepts with positive and negative affective valence, the process turned out to be particularly strong for opinion polarity detection.

5 Conclusion and Future Work

In this work we used PCA to build a language visualisation system which transforms natural language from a linguistic form into a 24D real-time navigable space. SenticSpace allows to discover emotionally related concepts and nicely visualize the affective information of common sense concepts.

In the future we plan to apply different multi-dimensionality reduction techniques, such as independent component analysis (ICA), to better build the vector space, and develop a web interface for interactively visualizing sentiments.

References

1. Elliott, C.: The Affective Reasoner: A Process Model of Emotions in a Multi-Agent System. The Institute for the Learning Sciences, Technical Report No. 32 (1992)
2. Ortony, A., Clore, G., Collins, A.: The Cognitive Structure of Emotions. Cambridge University Press, New York (1988)
3. Wiebe, J., Wilson, T., Cardie, C.: Annotating Expressions of Opinions and Emotions in Language. Language Resources and Evaluation 2(3), 165–210 (2005)
4. Somasundaran, S., Wiebe, J., Ruppenhofer, J.: Discourse Level Opinion Interpretation. In: COLING, Manchester (2008)
5. Wilson, T., Wiebe, J., Hoffmann, P.: Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. In: HLT-EMNLP, Vancouver (2005)
6. Hu, M., Liu, B.: Mining Opinion Features in Customer Reviews. In: AAAI, San Jose (2004)
7. Goertzel, B., Silverman, K., Hartley, C., Bugaj, S., Ross, M.: The Baby Webmind Project. In: AISB, Birmingham (2000)
8. Dyer, M.: Emotions and Their Computations: Three Computer Models. Cognition and Emotion 1(3), 323–347 (1987)
9. Cambria, E., Hussain, A., Havasi, C., Eckl, C.: Sentic Computing: Exploitation of Common Sense for the Development of Emotion-Sensitive Systems. In: Esposito, A., Campbell, N., Vogel, C., Hussain, A., Nijholt, A. (eds.) Development of Multimodal Interfaces: Active Listening and Synchrony. LNCS, vol. 5967, pp. 148–156. Springer, Heidelberg (2010)
10. Cambria, E., Hussain, A., Havasi, C., Eckl, C.: Common Sense Computing: From the Society of Mind to Digital Intuition and Beyond. In: Fierrez, J., Ortega-Garcia, J., Esposito, A., Drygajlo, A., Faundez-Zanuy, M. (eds.) Biometric ID Management and Multimodal Communication. LNCS, vol. 5707, pp. 252–259. Springer, Heidelberg (2009)
11. Havasi, C., Speer, R., Alonso, J.: ConceptNet 3: a Flexible, Multilingual Semantic Network for Common Sense Knowledge. In: RANLP, Borovets (2007)
12. Strapparava, C., Valitutti, A.: WordNet-Affect: an Affective Extension of WordNet. In: LREC, Lisbon (2004)
13. Fellbaum, C.: WordNet: An Electronic Lexical Database. MIT Press, Cambridge (1998)
14. Havasi, C., Speer, R., Pustejovsky, J., Lieberman, H.: Digital Intuition: Applying Common Sense Using Dimensionality Reduction. IEEE Intelligent Systems 24(4), 24–35 (2009)
15. Plutchik, R.: The Nature of Emotions. American Scientist 89(4), 344–350 (2001)
16. Minsky, M.: The Emotion Machine. Simon and Schuster, New York (2006)
17. Wall, M., Rechtsteiner, A., Rocha, L.: Singular Value Decomposition and Principal Component Analysis. In: Berrar, D., et al. (eds.) A Practical Approach to Microarray Data Analysis, pp. 91–109. Kluwer, Norwell (2003)
18. Eckart, C., Young, G.: The Approximation of One Matrix by Another of Lower Rank. Psychometrika 1(3), 211–218 (1936)

Supporting Collaborative Transcription of Recorded Speech with a 3D Game Interface

Saturnino Luz[1], Masood Masoodian[2], and Bill Rogers[2]

[1] School of Computer Science and Statistics
Trinity College Dublin
Dublin, Ireland
luzs@cs.tcd.ie

[2] Department of Computer Science
The University of Waikato
Hamilton, New Zealand
{masood,coms0108}@cs.waikato.ac.nz

Abstract. The amount of speech data available on-line and in institutional repositories, including recordings of lectures, “podcasts”, news broadcasts etc, has increased greatly in the past few years. Effective access to such data demands transcription. While current automatic speech recognition technology can help with this task, results of automatic transcription alone are often unsatisfactory. Recently, approaches which combine automatic speech recognition and collaborative transcription have been proposed in which geographically distributed users edit and correct automatically generated transcripts. These approaches, however, are based on traditional text-editor interfaces which provide little satisfaction to the users who perform these time-consuming tasks, most often on a voluntarily basis. We present a 3D “transcription game” interface which aims at improving the user experience of the transcription task and, ultimately, creating an extra incentive for users to engage in a process of collaborative transcription in the first place.

Keywords: Computer-assisted speech transcription, collaborative transcription, 3D interfaces, automatic speech recognition, single-player games.

1 Introduction

The unprecedented growth in availability of content on the Internet in the past decade has spurred a great deal of research on technologies for rendering this content more easily accessible to users. From an initial focus on fully automated content indexing, technology has steadily moved towards incorporating input from the user community.

Textual content exemplifies this situation quite clearly. Although vast volumes of data can be efficiently searched by the existing large-scale search “engines”, language (access to multilingual content) and semantics (access to structured content) remain as barriers to widespread access. Despite the fact that progress

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 394–401, 2010.

continues to be made in machine translation and text analytics research, tools that support user collaboration in text structuring (through collaborative filtering [13], or extensions to popular "wiki" platforms [16], for instance) and machine translation (via distributed content platforms [3], or though "crowdsourcing" projects etc) have received increasing attention.

The volumes of speech content publicly available on the network, both through audio and video media, have also increased greatly in recent years[1]. In addition to the requirements common to textual data indexing, speech content also requires conversion to text for practical and effective access. For the purposes of information retrieval tasks similar to text document search, techniques based on automatic speech recognition (ASR) have been devised which employ word lattices [18,2] as input to the indexing process. However, other tasks, such as browsing of lecture "podcasts" and instructional video material, require users to be able to read the spoken content as they would a written text. For such tasks, the levels of accuracy attained by current ASR are inadequate and manual correction is often necessary. As in the case of translation and semantic structuring, schemes which harness the contributions of the user community have been devised.

Ogata & Goto [12], for instance present a Web-based system which displays transcriptions of "podcasts" which the user can correct at will. These transcriptions are initially generated by a large vocabulary speech recognizer so that in correction mode the user can access for each word a set of alternatives produced by the recogniser. Any of these words can be selected if they correspond to the speech actually heard by the user, or the user can enter a new word if the speech does not correspond to any word in the list. Corrected sentences can then be fed back to ASR training, potentially improving future performance, in addition to improving the quality of the existing transcriptions. The success of this kind of systems obviously relies heavily on user involvement. Although Ogata & Goto, claim that users of their system found the task of correcting transcriptions "enjoyable", it is hard to imagine many people would feel compelled to do it on a regular basis.

In this paper we describe TRAEDRIS, a "transcription editing game" loosely inspired by the popular Tetris game which aims to provide users with a stronger motivation to correct transcripts, namely, entertainment. TRAEDRIS displays sentence transcription candidates through animated 3D representations of word lattices generated by speech recognition. The user can interact with these sentence representations by selecting the correct paths as the words move towards the background.

1.1 Problem Definition

ASR accuracy varies widely, depending on a number of factors. These include: the level of noise in the recording, variations in accent and voice quality, speaker

[1] This growth in volume is apparently matched by a growth in use. According to a Nielsen's *Netratings* report (`http://en-us.nielsen.com/main/insights/reports`), video content delivered by the main video websites grew by 41% from 2008 to 2009.

change, the accuracy of pre-processing (e.g. sentence boundary detection or other types of segmentation [14]), and the adequacy of the training data to the speech data to be transcribed, in terms of audio quality, vocabulary and language models.

Although relatively versatile user interaction techniques, such as alternative lists [10], "respeaking" [1] and other variants of multimodal interaction [15], have been developed for dictation systems and ASR-assisted transcription [11], users tend to prefer retyping from scratch when the accuracy of the initial ASR-generated transcription is low [7].

Accuracy in speech recognition is usually measured in terms of word error rates (WER), that is, the total number of deletions, insertions and substitutions in the transcribed sentence with respect to the correct sentence. In other words, WER is given by the Levenshtein distance $wer(W, R)$ between the ASR *hypothesis* W and the *reference* sentence R. Another, less forgiving way of measuring accuracy is at the sentence level. A sentence-level error (SER) can be defined as the ratio between the number of sentences containing at least one error and the total number of sentences transcribed. Even the best recognisers available today have typically high SER while exhibiting relatively low WER. This is specially true when the ASR system is used with good quality audio produced by a small number of speakers as is the case of much broadcast, "podcast", and some recorded meeting data available on the Internet. A state of the art system employed to transcribe the Hub5 conversational telephone speech dataset, for example, achieves WER of 26.5% while its SER is about 66.2% [4]. Similar results have been reported for other systems that attempt to minimise WER directly on the Switchboard corpus [9] and the North American Business corpus (NAB'94). On the latter, the system presented in [17] attains WER as low as 11.1% even though its SER is 74%. As regards user input, this disparity between WER and SER basically means that such transcripts are well suited to manual correction in that, most of the time, the user will only need to correct a few words at a time in order to repair an erroneous sentence.

In addition to high accuracy, applications often require precise time alignment between the speech signal and the transcription at a phrase [14] or word level. Given the above described scenario and requirements, the design problem which motivated the development of TRAEDRIS can be stated as follows: to design an interactive 3D graphics game for generation of accurate speech transcripts by rapid, collaborative correction of medium to low WER speech recognition results by users on the Internet.

2 The 3D Transcription Game

TRAEDRIS is a single-player game in which ASR transcribed sentences appear on the screen as their corresponding audio is played, and move slowly towards the background. The player's aim is to keep their screen clear by correcting each sentence before it disappears towards the background, or before it reaches the last unoccupied position on the z axis. When a sentence reaches its limit for placement on the z axis (i.e. either the position in front of a previous sentence or a pre-defined limit point, if the z horizon is empty) a confidence score is computed. If the confidence score is greater than a threshold value, the sentence simply fades out into the horizon. Otherwise, the sentence stops, "piling up" on the screen. As in Tetris, the game can be set to different initial speeds, and the speed at which the sentences move increases as the game progresses. The game ends when the z axis has accumulated enough sentences that a newly appearing sentence cannot move towards the back. The player's total score can be computed in different ways, but for simplicity let us define it as the number of words in the sentences that faded out (i.e. those whose scores exceeded the confidence threshold).

2.1 The Transcription System

The overall architecture of the TRAEDRIS system is shown in Fig. 1. We assume that the speech input is already segmented into sentences and that these sentences are appropriately stored in a speech database. This audio signal is initially fed to the ASR system which decodes it producing, for each sentence W, a word lattice and a confusion network which store the recognition hypotheses for that sentence along with its posterior probabilities $P(W|A)$. This is done offline, in batch, so that all recognition hypotheses are available to the scoring module and the game's graphical front-end when the game starts.

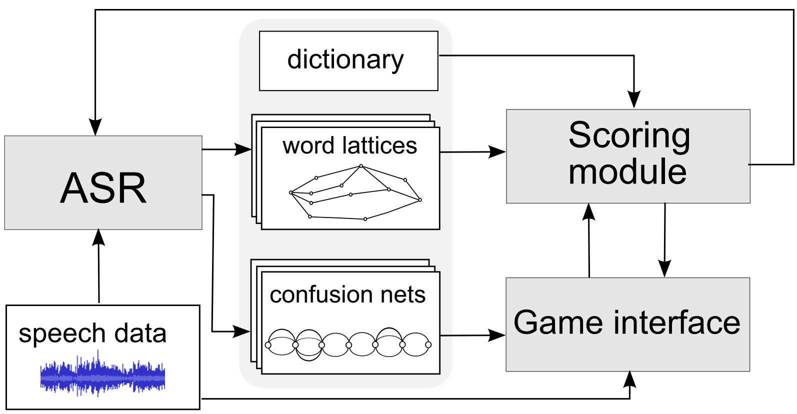


Fig. 1. TRAEDRIS system architecture

Word lattices are employed in computing the scores and determining the confidence thresholds for each user-corrected sentence, as explained in Section 2.2. Optionally, an external, domain-specific dictionary can be used to improve the scores when out-of-vocabulary corrections are entered. The graphical representation of the sentence candidates shown on the user interface, on the other hand, is based on the hypotheses encoded in the confusion network. Confusion nets can be regarded as compact representations of word lattices [9] and are used by TRAEDRIS, as well as other computer-assisted transcription systems [11,12,7], in order to simplify the correction task, from the users perspective.

The game interface also supports audio playback. Once the confusion network for a given sentence is selected for display, the corresponding speech is retrieved from the database and played back to the user.

2.2 The Scoring Method

Ideally, corrected sentences should be scored against a "gold standard", perhaps a faithful transcription produced by a professional transcriber. However, such a scoring mechanism would defeat the point of supporting collaborative transcription in the first place. It is precisely because such reference transcriptions are not readily available for most speech content that one would like to support transcription by the user community. Since there is no gold standard, possible sources of information that can be exploited in generating a score for a user-corrected transcription are: (a) measures derived directly from ASR system's posterior probabilities, and (b) measures that incorporate the changes made by previous players.

The latter might involve, for instance, clustering the various versions of a sentence (based on, say, their Levenshtein distances to each other), aligning all sentences against the best cluster and then scoring the newly edited sentence against the alignments so that edits that contradict the majority get penalised. This scheme would require the system to present the same sentences to many players so as to gather a number of hypotheses large enough to produce a meaningful consensus. However, this would once again defeat the purpose of the game, which is to encourage manual correction of as much transcription as possible without repeating the same transcripts many times over to different players. We therefore chose to focus on a score derived from the posterior probabilities available through the word lattice.

Sentence posteriors in the maximum likelihood approach to ASR [6] are approximated as the product of a language model prior $P(W)$ and acoustic likelihoods $P(W|A)$ learnt from training data:

$$P(W|A) \approx P(W)P(A|W) \tag{1}$$

In this framework, the transcription hypothesis chosen by the system is the one which maximises the posterior (MAP hypothesis). It has been shown, however, that although MAP minimises SER it is not guaranteed to minimise WER [5,9]. Since we would like our score to somehow reflect the number of words the user corrected, we need to base it on an alternative criterion of error minimisation. Goel et Al. [5] propose a decision rule based on minimising the expected word error according to the posterior distribution. We derive our scoring strategy from that rule.

First, we compute the best (i.e. lowest) expected word error for the n best hypotheses in the word lattice [9]:

$$E_{min} = \min_{i=1}^{n} \sum_{j=1}^{n} wer(W_i, W_j) P(W_i|A) \tag{2}$$

We then set the threshold τ to be a fraction of $-E_{min}$ and score the user-corrected sentence W_u by weighting it against the ASR hypotheses as shown in equation (3), so that the sentence is allowed to disappear from the screen if S_u exceeds τ. The loss function $l(\cdot,\cdot)$ for this calculation can be simply $wer(\cdot,\cdot)$ or a modified edit distance which does not penalise substitutions or insertions of words missing from the ASR vocabulary but contained in the domain specific dictionary (see Fig. 1).

$$S_u = -\sum_{j=1}^{n} l(W_u, W_j) P(W_i|A) \tag{3}$$

The scoring mechanism can also be made to vary as new hypotheses are entered by the users (and the transcription presumably improves) by adding such sentences to the n-best list with maximum posterior probabilities.

2.3 Visualisation and Interaction

The TRAEDRIS user interface is based on an earlier prototype developed for visualisation of speech recognition results through a simulated 3D environment

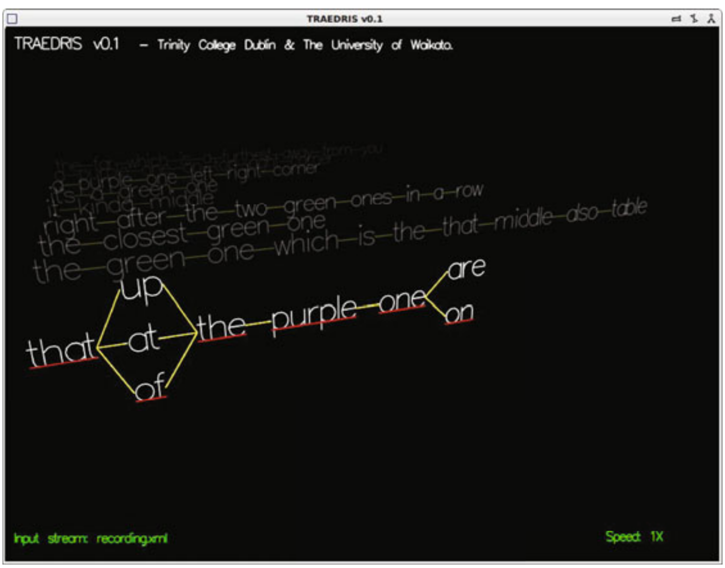


Fig. 2. TRAEDRIS user interface

[8]. The system represents the recognition hypotheses as paths in a graph and displays this representation, initially on the foreground. The recognition alternatives with the greatest posterior probabilities highlighted and connected through red-coloured edges. Hypotheses of lower likelihood are dimmed. As time passes, the word graphs move towards the background until it is no longer possible for the user to interact with them. The interface is shown in Fig. 2.

For speed, the user is allowed to move the cursor through the graph with the keyboard's arrow keys, highlighting the path containing the corrected sentence. Special *split* and *merge* operations can be activated through key combinations if a word must be split into two or more words, or if a group of word slots must be merged into a single word. The user is also allowed to stop the sentence in mid-air a limited number of times (depending on the difficulty level set), in order to enter a correction not shown in the current confusion network by typing it in.

Once a sentence reaches the horizon or the top of the pile, its score is computed in the manner indicated above, and a new confusion network is selected for display. TRAEDRIS displays the player's current accumulative score based on the scores of individual sentences that have been processed. At the end of the game the player's final score and overall ranking are presented.

3 Conclusions and Further Work

In this paper we have presented a 3D single-player game as an alternative to the conventional text editor interfaces common to most manual speech transcription correction systems. We made a design decision to keep the TRAEDRIS game rather simple in its user interaction due to the fact the actual task of transcription correction is mentally demanding, requiring the user to listen to the corresponding speech audio segments while also reading and correcting the ASR generated transcripts. Despite its simplicity, however, TRAEDRIS presents a potentially more engaging activity than standard text-editing. Past experience has also shown that even simple games, such as Tetris, can be fun and stimulating if the player can use simple and quick actions (e.g. the keyboard arrow keys) to play the game. These improvements in user experience are meant to provide a motivation for greater involvement from the community of users of speech archives in collaboratively improving these archives.

We have also introduced a novel application and extension of the consensus technique [9] for scoring user-edited sentences. Planned future work includes a detailed evaluation of TRAEDRIS in order to clarify the type of user experience it provides, and whether it facilitates generation of more accurate transcripts.

References

1. Ainsworth, W.A., Pratt, S.R.: Feedback strategies for error correction in speech recognition systems. International Journal of Man-Machine Studies 36(6), 833–842 (1992)
2. Chelba, C., Silva, J., Acero, A.: Soft indexing of speech content for search in spoken documents. Computer Speech & Language 21(3), 458–478 (2007)

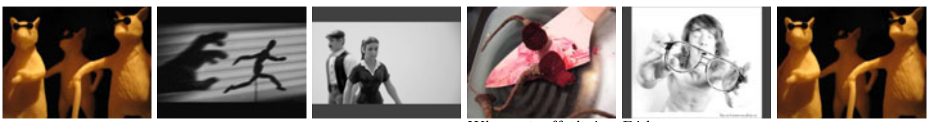
3. Désilets, A., Gonzalez, L., Paquet, S., Stojanovic, M.: Translation the Wiki way. In: WikiSym 2006: Proceedings of the 2006 International Symposium on Wikis, pp. 19–32. ACM, New York (2006)
4. Evermann, G., Woodland, P.C.: Posterior probability decoding, confidence estimation and system combination. In: Proceedings of the Speech Transcription Workshop. College Park, MD (October 2000)
5. Goel, V., Byrne, W., Khudanpur, S.: LVCSR rescoring with modified loss functions: a decision theoretic perspective. In: Procs. of the IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP 1998), vol. 1, pp. 425–428 (1998)
6. Jelinek, F.: Statistical Methods for Speech Recognition. MIT Press, Cambridge (1998)
7. Luz, S., Masoodian, M., Rogers, B., Deering, C.: Interface design strategies for computer-assisted speech transcription. In: Proceedings of the Australasian Conference on Human-Computer Interaction (OZCHI 2008), pp. 203–210. ACM, New York (2008)
8. Luz, S., Masoodian, M., Rogers, B., Zhang, B.: A system for dynamic 3D visualisation of speech recognition paths. In: Bottoni, P., Levialdi, S. (eds.) Proceedings of Advanced Visual Interfaces (AVI 2008), pp. 482–483. ACM Press, New York (2008)
9. Mangu, L., Brill, E., Stolcke, A.: Finding consensus in speech recognition: word error minimization and other applications of confusion networks. Computer Speech & Language 14(4), 373–400 (2000)
10. Munteanu, C., Baecker, R., Penn, G.: Collaborative editing for improved usefulness and usability of transcript-enhanced webcasts. In: Proceedings of the 26th SIGCHI Conference on Human Factors in Computing Systems (CHI 2008), pp. 373–382. ACM, New York (2008)
11. Nanjo, H., Kawahara, T.: Towards an efficient archive of spontaneous speech: Design of computer-assisted speech transcription system. The Journal of the Acoustical Society of America 120, 3042 (2006)
12. Ogata, J., Goto, M.: PodCastle: a spoken document retrieval system for podcasts and its performance improvement by anonymous user contributions. In: SSCS 2009: Proceedings of the ACM Multimedia Workshop on Searching Spontaneous Conversational Speech, pp. 37–38. ACM, New York (2009)
13. Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J.: GroupLens: an open architecture for collaborative filtering of netnews. In: Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work, CSCW 1994, pp. 175–186. ACM, New York (1994)
14. Roy, B., Roy, D.: Fast transcription of unstructured audio recordings. In: Proceedings of the 10th Annual Conference of the International Speech Communication Association (INTERSPEECH), Bristol, UK, p. 4 (2009)
15. Suhm, B., Myers, B., Waibel, A.: Multimodal error correction for speech user interfaces. ACM Transactions on Computer-Human Interaction 8(1), 60–98 (2001)
16. Völkel, M., Krötzsch, M., Vrandecic, D., Haller, H., Studer, R.: Semantic Wikipedia. In: Proceedings of the 15th International Conference on World Wide Web, WWW 2006, pp. 585–594. ACM, New York (2006)
17. Wessel, F., Schluter, R., Ney, H.: Explicit word error minimization using word hypothesis posterior probabilities. In: Procs. of the IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing (ICASSP 2001), vol. 1, pp. 33–36 (2001)
18. Zhou, Z.Y., Yu, P., Chelba, C., Seide, F.: Towards spoken-document retrieval for the internet: lattice indexing for large-scale web-search architectures. In: Proceedings of the Conference of the North American Chapter of the ACL, pp. 415–422 (2006)

Text-to-Video: Story Illustration from Online Photo Collections

Katharina Schwarz[1], Pavel Rojtberg[2], Joachim Caspar[2], Iryna Gurevych[2], Michael Goesele[2], and Hendrik P.A. Lensch[1]

[1] Ulm University
[2] TU Darmstadt

Abstract. We present a first system to semi-automatically create a visual representation for a given, short text. We first parse the input text, decompose it into suitable units, and construct meaningful search terms. Using these search terms we retrieve a set of candidate images from online photo collections. We then select the final images in a user-assisted process and automatically create a storyboard or photomatic animation. We demonstrate promising initial results on several types of texts.



Three blind mice. Three blind mice. | See how they run. See how they run. | They all ran after the farmer's wife, | Who cut off their tails with a carving knife, | Did you ever see such a sight in your life, | As three blind mice?

1 Introduction

Telling a story by natural language is a process everybody gets trained for all his life. Telling a story by images however is significantly harder, because the generation of images let alone video is a time consuming process and the outcome largely depends on skill. On the other hand, pictures have a similar expressive range as natural language, with regard to describing objects, actions, or evoking specific emotions. Our goal is to provide a framework that simplifies the process of telling a story with pictures but eliminates the need to create the images yourself. More precisely, we aim at generating visualizations driven by natural language, augmenting written text semi-automatically by a sequence of images obtained from online photo collections as shown in the example above.

The main task is to obtain a semantically close translation from natural language to image sequences. As currently neither the semantics of written text nor of images can be automatically extracted with sufficient success rates by freely available tools, we provide a user-in-the-loop solution: Rather than semantically analyzing the text we parse individual sentences to extract the functional description for the individual words. In order to determine semantically matching images we rely on the tagging of images in photo-community sites such as Flickr. From the extracted parse trees, we formulate optimized search queries to obtain

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 402–409, 2010.

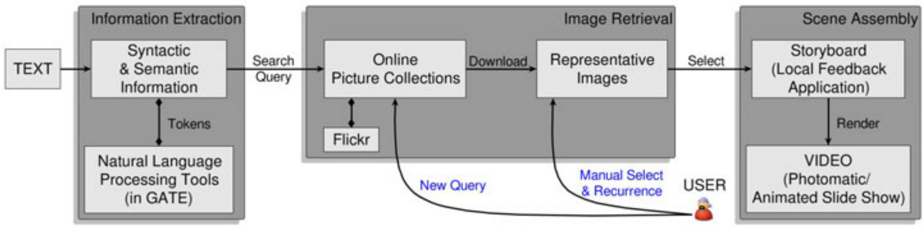


Fig. 1. System overview: information extraction, image retrieval and scene assembly

as specialized images as possible for each part of a sentence. Typically, for each query a set of images is returned. The ultimate choice of which image to include in the final output is left to the user. We provide automatic means to select images with similar color distribution along the sentences. The images can now be presented in various formats, e.g., as storyboard or as slide animation. The quality of the image sequence largely depends on the available tagged imagery and the semantic complexity of the input text. Issues that are not correctly resolved by the parser will typically yield unsatisfying results but can often be corrected with little user intervention.

Figure 1 shows a high level overview of our system based on three parts. The information extraction part consists of automatically parsing and segmenting a given input text into parts of text, or *POTs* (Sec. 3). For each POT we construct an optimized search query (Sec. 4). The image retrieval part (Sec. 2) consists of automatically querying the online collection Flickr and retrieving a set of candidate images for each POT. Finally, the scene assembly part (Sec. 5) consists of automatically or semi-automatically picking the most representative images (*shots*) per POT. We present results on a nursery rhyme, fairy tale, screen play, short story, and a news article (Sec. 6), and close the paper with a short discussion (Sec. 7).

2 Image Retrieval

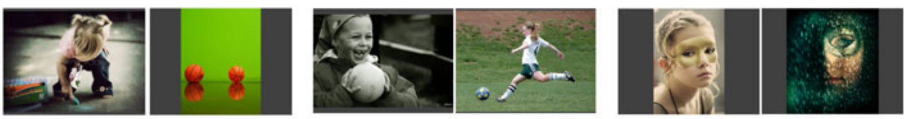
Online photo collections provide a tremendous amount of imagery (e.g., Flickr currently stores more than 4.5 billion images). They are widely used in the graphics community, e.g., as source for clip arts [8], to perform scene completion [6], to create photorealistic images from sketches [1] or computer-generated imagery [7], or to visualize and reconstruct scenes [12].

Images on Flickr are attributed by titles, tags, and texts by users in an informal way yielding a so called *Folksonomy* [5] (as opposed to the more formal ontology). Flickr allows for a full text search in each of these three categories. Each query will potentially return a set of images. One can measure the precision of the answer with respect to a query by manually counting the number of matching images:

$$precision = \frac{|\{relevant\ \ images\}|\cap|\{retrieved\ \ images\}|}{|\{retrieved\ \ images\}|}$$

Table 1. Compound query results for fairy tales: precision in percent and number of total successful queries in parenthesis. A query was counted as successful, if at least one matching image was retrieved.

Search Tokens	Queries	Full Text Search	Title Search	Tag Search
Combined Nouns	25	21% (25)	21% (20)	36% (24)
Nouns & Adjectives/Adverbs	25	17% (25)	25% (19)	36% (21)
Nouns & Verbs	25	8% (25)	13% (17)	16% (11)
Nouns & Averbs	20	5% (12)	10% (10)	12% (5)
Nouns & Adverb Stems	20	8% (9)	3% (7)	14% (5)
Nouns & Verbs	20	8% (20)	13% (13)	15% (6)
Nouns & Verb Stems	20	10% (19)	14% (12)	25% (11)



girl $\vee$ ball   girl $\wedge$ ball   girl $\wedge$ ball $\wedge$ beautiful

Fig. 2. Successive specialization of a shot for "the beautiful girl at the ball"

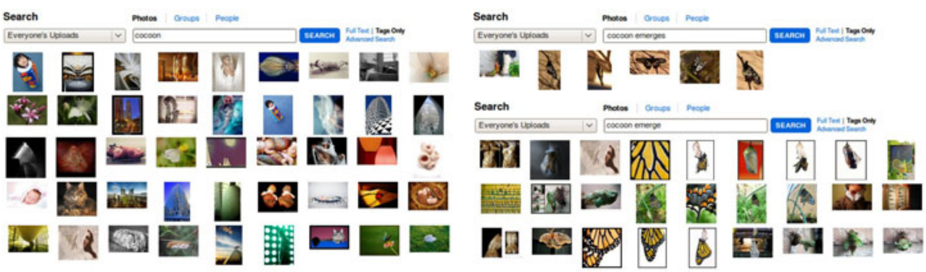


Fig. 3. Improving the correlation of shot and sentence action by combining nouns with verbs or verb stems

Basic Tokens and Stemming. In order to obtain an image which matches the semantics of a POT, we need to assemble a compound query. By far the most frequent category of words in Flickr tags are nouns (about 90%) while verbs, adjectives and adverbs are found less often. Most often, querying for a single noun ignores the information of the remainder of the POT. By combining the nouns and attributes of one POT by conjunction, more and more specific images can be retrieved, such that they finally match the desired semantics (Fig. 2). Thus, combining nouns or adding adjectives, adverbs, or verbs can help retrieving a more specialized collection of images even though the precision for the query per se might drop (Table 1). In particular, the precision for queries including verbs is low. But most often the retrieved images show the action represented by a sentence much better. Fig. 3 demonstrates this with the retrieved images for the queries *cocoon* (noun), *cocoon emerges* (noun & verb), and *cocoon emerge* (noun & stemmed verb). As shown in the second half of Table 1, the precision for adverb or verb queries can also be improved by querying for the stem rather than the inflected form.

The highest precision is typically obtained searching in tags, but verbs and adverbs are rarely found in tags while they are more frequent in full text or title search. We will use these insights in Sec. 4 to assemble an optimal query for each POT.

3 Information Extraction

Information extraction (IE) takes unseen documents as input and produces selectively structured output from the explicitly stated or implied data which is found in the text [2]. In order to form versatile queries for image retrieval, we extract syntactic and semantic information from the story using the General Architecture for Text Engineering (GATE) [3]. GATE is a modular infrastructure for developing and deploying NLP (Natural Language Processing) software components from various developers.

Preprocessing with ANNIE. GATE is distributed with an IE system called ANNIE (A Nearly-New Information Extraction System). Figure 4 shows the components we use from ANNIE. Tokeniser and Sentence Splitter are required to annotate each word or symbol with a part-of-speech tag by the POS Tagger. The tokens determined by this pipeline form the basis for generating our queries.

Additional Tokens from Stemming. Furthermore, we discovered the usage of verb stems as a significant improvement concerning the query answers. GATE provides the Snowball Stemmer as plugin, a tool based on the Porter stemmer for English [11]. It annotates each token with its stem. We apply this to verbs and adverbs only.

Text Segmentation. Starting with sentences as the most general entities, they are split by punctuation marks or coordinating conjunctions in order to receive smaller segments, the POTs, which could be clauses, the entities of an enumeration or similar. Within these POT objects we query ANNIE tokens, especially nouns, verbs, adjectives and adverbs, as they contain the most significant information. For the same reason, we remove auxiliary verbs. The next step is to find an appropriate image (a shot) for each POT.

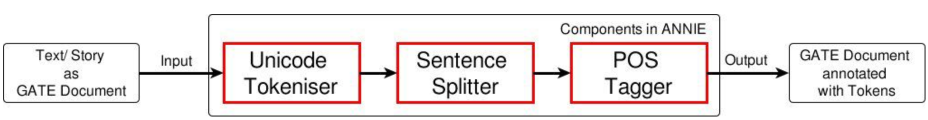


Fig. 4. GATE preprocessing with ANNIE components

4 Forming the Query

Shots for the individual POTs are obtained by submitting proper queries to the online image collections. Using the list of tokens extracted for each POT (Sec. 3), we formulate an appropriate query that results in a high precision and sufficiently many images to choose from. A conjunction of all n_t token T_i in a POT will produce the most specialized query. However, if too many constraints are added it might happen that only an unsatisfactory number of images is reported. We therefore assemble the query iteratively with the goal to find a sufficiently large set of images which are as specific as possible.

Following the evidence of Sec. 2, we create a priority list Q of compound queries for a shot. The first query is the conjunction of all n_t tokens, followed by a disjunction of conjunctions formed by all possible subsets containing $n_t - 1$ tokens, etc., until we end with a disjunction of all n_t individual tokens. For $n_t = 3$, we would assemble the following list of queries:

$$Q = \{(T_1 \wedge T_2 \wedge T_3), ((T_1 \wedge T_2) \vee (T_1 \wedge T_3) \vee (T_2 \wedge T_3)), (T_1 \vee T_2 \vee T_3)\}.$$

Due to their importance in the syntactic analysis and their frequency in the image tags, we treat nouns in a special manner. Queries combining multiple nouns yield the highest precision and therefore, we first treat the conjunction of all nouns as a single token in the algorithm outlined above, and in a second step create a list for each noun separately and append them.

Based on the priority list Q, the system issues a sequence of queries, accumulating the downloaded images. The process is stopped as soon as the number of downloaded images for one shot exceeds a user defined threshold, e.g., 30. For each entry of the list, we first perform a query on Flickr tags and then perform a title search. This way, we were able to download 30 images for almost all of our example shots and due to the structure of the priority queue ensured that the most specialized images are always at the top of the image set.

5 Scene Assembly

We now retrieved a set of candidate images per shot, sorted by relevance, from which we can automatically select the highest ranked image for each shot. Alternatively, this selection can be performed by the user. User selection will, for example, be necessary if the semantics, style, or composition of the highest ranked image does not match the user preference. To simplify the selection process, we provide the user with two sets of tools, one dealing efficiently with *recurring queries*, and another handling *color consistency* between neighboring shots.

Recurrence. It is often desirable to use the same image for similar shots (see first and last image of the teaser figure). Our system therefore reuses by default the selection results for shots with the same query. If the user chooses a different image for one shot, all other shots in this category are updated accordingly. Beyond a literal match, the user can manually group multiple queries together which will then be represented by the same image.

Fig. 5. Color consistency enforced on the selection for one query. The queries are indicated in blue.

Color Consistency. Online images typically vary largely in style and color. Similarity across queries can be increased by sorting the image sets for neighboring shots by color similarity. Mehtre et al. [9] indicate that it is sufficient to perform a coarse comparison to exclude severe color miss-matches in image retrieval. We therefore compute for each image its mean RGB color vector. Given the current representative image I_A for shot A, we select for a neighboring query B the representative image I_B that minimizes the Euclidean distance of the mean color vectors. For the next query C, the comparison is carried out with respect to I_B, and so on. The user is free to indicate whether or not the color matching constraint should be applied and, if so, into which direction the color should be propagated. The improved results are shown in Fig. 5. Alternatively, color consistency can be achieved by processing the selected images (e.g., converting them to sepia or black-and-white representations, applying color style transfer techniques [10]).

6 Results

At this stage, each POT has been assigned a shot. The resulting image set can simply be represented as a storyboard (see teaser image and Fig. 6) or presented as in the accompanying video as an animated slide show, where text and image transitions are achieved by a constant motion (see the paper web page).

We have applied our system on various text types: the nursery rhyme "Three Blind Mice", the fairy tale "Cinderella", a part of the screenplay to "Braveheart", a news article about the Wii controller, and the short novel "Animal Farm". In general, we observed high context-sensitive precision in our tests, considering the actual meaning of the sentence rather than just the queried tokens. This can be seen in Table 2, which shows our results for a range of text types. For each

Table 2. Results on different text types

text type	prec. #1	prec. #10	# words	# queries	#user interactions	query time
nursery rhyme	40%	46%	53	8	5	00:02:33
fairy tale	80%	60%	538	73	45	00:33:18
screen play	80%	55%	230	41	28	00:20:59
news	60%	49%	147	19	12	00:12:02
short novel	40%	38%	967	150	113	01:27:08

Fig. 6. Automatically generated visual story (top row) and manually improved version (bottom row) of a news article about the Wii controller

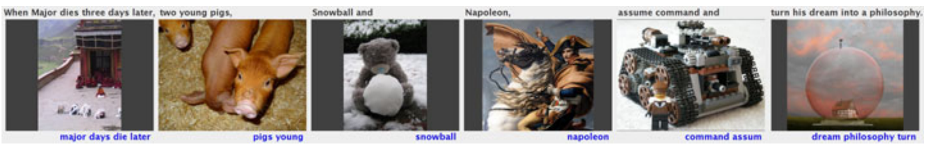


Fig. 7. Semantically mismatching images due to metaphorical character names in the "Animal Farm" novel

query, 30 images were downloaded and the context-sensitive precision for the first and the first 10 images was evaluated. Because of the very good sentence related precision of the first reported images, the required user interaction to construct a semantically close storyboard is moderate.

We think that the variation in the text types correlates with the typical spectrum of submitted Flickr images. The content of the selected nursery rhyme and short novel is slightly more abstract than the content of the other categories.

In Fig. 6, we compare completely automatically generated results against the manually optimized selection. While quite a number of shots received a decent representative automatically, a few clicks were necessary to obtain the final selection where semantic errors and deviations in style have been removed.

In general, we were surprised by how often the retrieved images for our generated queries match the intention of the original text. However, Fig. 7 clearly shows the limits of our approach, namely dealing with word sense ambiguities. The images retrieved for the two pig characters *Snowball* and *Napoleon* from the novel "Animal Farm" do not depict pigs, but the literal or most frequent meaning of the words. A solution for this problem could be the usage of sophisticated lexical semantic analysis, such as word sense disambiguation and named entity recognition.

So far, our parsing is limited to group only tokens that are adjacent in a sentence. In the future, we would like to use dependency graphs [4] to assemble better queries for the non-connected parts of the sentence.

We expect further improvements if anaphora could be automatically resolved or if spatial relations in the scene could be considered in the query.

7 Conclusion

Our system can be seen as one of the first steps towards creating a movie based on a textual input. After parsing the text, the system automatically generates queries and retrieves images from the community site Flickr. Most often, a set of representative images is found automatically. After a few user interactions, a reasonable storyboard is produced.

By enhancing the semantic analysis of the text, e.g., by using WordNet or by considering syntactic as well as semantic relations between objects, the quality of the (semi-)automatically generated storyboards might be further increased. Additionally, an automatic classification of the retrieved images into sets of similar images might facilitate the manual selection process. One natural extension to the presented system is to animate the retrieved images to better visualize the action in a story – as a next step towards creating full-fledged movies.

Acknowledgments. We would like to thank the Flickr users for the images used in our research. This work was funded in part by the DFG Emmy Noether fellowships Le 1341/1-1, GO 1752/3-1, the DFG Research Training Group 1223, and the Lichtenberg Program No. I/82806 of the Volkswagen Foundation.

References

1. Chen, T., Cheng, M.-M., Tan, P., Shamir, A., Hu, S.-M.: Sketch2Photo: Internet image montage. ACM Trans. Graph. 28(5) (2009)
2. Cowie, J., Lehnert, W.: Information extraction. ACM Commun. 39(1), 80–91 (1996)
3. Cunningham, H., Maynard, D., Bontcheva, K., Tablan, V.: GATE: A framework and graphical development environment for robust NLP tools and applications. In: Proc. 40th Anniv. Meeting of the Assoc. for Comp. Ling. (2002)
4. de Marneffe, M.-C., MacCartney, B., Manning, C.D.: Generating typed dependency parses from phrase structure parses. In: LREC 2006 (2006)
5. Guy, M., Tonkin, E.: Folksonomies, tidying up tags? D-Lib Magazine 1 (2006)
6. Hays, J., Efros, A.A.: Scene completion using millions of photographs. ACM Trans. Graph. (2007)
7. Johnson, M.K., Dale, K., Avidan, S., Pfister, H., Freeman, W.T., Matusik, W.: CG2Real: Improving the realism of computer generated images using a large collection of photographs. Technical Report MIT-CSAIL-TR-2009-034 (2009)
8. Lalonde, J.-F., Hoiem, D., Efros, A.A., Rother, C., Winn, J., Criminisi, A.: Photo clip art. ACM Trans. Graph (2007)
9. Mehtre, B.M., Kankanhalli, M.S., Narasimhalu, A.D., Man, G.C.: Color matching for image retrieval. Pattern Recogn. Lett. 16(3), 325–331 (1995)
10. Neumann, L., Neumann, A.: Color style transfer techniques using hue, lightness and saturation histogram matching. In: Comp. Aesthetics in Graphics, Visualization and Imaging, pp. 111–122 (2005)
11. Porter, M.F.: An algorithm for suffix stripping. Program 14(3), 130–137 (1980)
12. Snavely, N., Simon, I., Goesele, M., Szeliski, R., Seitz, S.M.: Scene reconstruction and visualization from community photo collections. Proc. of the IEEE, Special Issue on Internet Vision (2010)

Improving Communication Using 3D Animation

Laurent Ruhlmann, Benoit Ozell, Michel Gagnon, Steve Bourgoin, and Eric Charton

Ecole Polytechnique de Montréal, 2900 boulevard Edouard-Montpetit,
Montréal, QC H3T 1J4, Canada
{laurent.ruhlmann,benoit.ozell,michel.gagnon,
eric.charton}@polymtl.ca, sbourgoin@unimasoft.com

Abstract. This paper presents a high level view of a project which aims at improving the communication between people who do not share the same language by using a 3D animation upper layer to lift the inherent limitations of written text. We will produce animation using the Collada file format, a standard based on the XML format. This project (GITAN) has started in January 2010, and we expect to have the first results by the end of 2010. A limited set of sentences will validate the global process and help us refine the tools developed in the pipeline.

Keywords: Animation, Collada, Natural Language Processing, Ontology.

1 Introduction

Since Aristotle, human language has been considered a partial way of representing the human thoughts. "The soul never thinks without a mental image"[1]. With the development of computer graphics, a new family of tool has emerged to improve communication. In this paper we present the *Gitan Project.* This project aims at creating a new language based on animation sequences in order to swap between text and graphics. More specifically, the project will develop a grammar linking text and animation, thus allowing the conversion between them. Such a grammar will allow us to lift the inherent ambiguity of natural languages. Using the powerful multimedia capacity of animation, it will be able to move from an abstract representation to a more concrete graphical view. There are many challenges that this project will face; one of the most important one is the ability to maintain the number of the existing animation segments at a reasonable level. Another important challenge is the fact that we want to combine algorithmically various animations into a plausible solution. When combining a walk cycle and a translation, in order to represent a walking person, we must ensure that the duration of the animation and the speed of the walk represents a good approximation.

2 Related Work

Animated scenes have already been used as a good medium to help the communication between adults and children [2], to illustrate domain specific activities like car accident simulations [3], or even for storytelling [4]. Another important project was

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 410–419, 2010.

aimed at the object reuse and the adaptation of 3D data for automatic scene creation [5]. The major difference between those projects and this one comes from the application domain and the linguistics approach we are taking. We want to cover a broader set of words, not limiting ourselves to a specific domain like car accidents. We want also limit the number of original animations to a minimum. In the first phase, we will not handle emotions or high level of languages like poetry.

One of the hypotheses of this project is that we will be able to reconcile the text and the animated image which is distinct types of representations [6]. We believe that animations will facilitate the understanding of a text, in different languages. Another hypothesis is that the principle of compositionality [7] is also applicable to animations. Being able to compose various animations in order to produce a plausible sequence is key to the project; we have already validated some simple cases. A third hypothesis is that the computer graphics field is mature and rich enough to be used as a good communication medium. We are not planning to develop videogame-like animations, but the realism achieved by some current games gives us some confidence about the validity of this hypothesis, even if the goal of the project is communication and not realism; In fact too much realism could limit effective communications between non-native speakers, since each culture has its own visual standard for representing emotions or complex actions. A final major hypothesis is that through proper usage of semantic tools (like ontologies), applied to the computer graphics domain, we will be able to minimize the amount of objects and animation to be modeled, in order to produce meaningful animations.

3 The Animation Standards

The term animation has been used throughout the ages to define 'rapid display of a sequence of images of 2-D or 3-D artwork, in order to create an illusion of movement' [8]. It is interesting to note that in 3200 BC, a bowl created with a sequence of 5 images depicting a jumping goat was found in Iran. Since it is easier to use a computer, we will limit our definition to 3D artwork generated and displayed by a computer. In this domain, we will privilege key-frame based animation, since this is the most common used technique, and realism (which is for instance achieved by physically based animation) is not our primary goal. In order to be open and readable by many Digital Content Creation (DCC) software, we will use the Collada (COLLaborative Design Activity) format, which is emerging as a new exchange standard [9]. DCC software (3DS Max, Maya, Softimage, 3dvia,) are already supporting it. Many viewers supporting it are available. It is an XML schema supporting animations, physics and many other features. It allows extensions which will be used to define specific information needed by the project. There are a lot of scripting languages (Maya's MEL, Softimage ICE, Virtools's SDK, 3ds Max's Maxscript, Unreal's Development Kit ...), but they fail to provide a high level language which will allow the complete description of an animation for our project. They all have very good scripting capabilities, and are focused on areas which intersect only partly with our goals. This is why we decided to develop our own formalism and language.

4 Intermediate Format between Text and Animation

Going from text to animation is a complex process which cannot be achieved in one step. We developed an incremental approach, using an intermediate representation based on the XML schema concept. It is an XML schema [10], since this formalism provides a clear and easy way to specify the various elements of a grammar. The file based on this format is split in 2 main parts, the *Data* and the *Process* parts. The *Data* part contains a detailed description of all the elements needed in the animation, and the *Process* part describes how the animation is going to be played. It can be viewed as a textual animation storyboard [11**].** A detailed description of the parts is given in section 6.1 and specific examples in section 6.8.For a complete description of the project's architecture, see figure 1.

5 The Need of Ontologies

There are many definitions of the term ontology depending on the domain to which it applies. For our purpose we will use the definition used in the information science: "An ontology is a specification of a conceptualization." [12]. We will use the ontologies tools (like OWL, [13]), to abstract meaning from the geometries used in the animations. This way we will be able to minimize the amount of 3D objects needed for building the animations. The most interesting project to create and define semantic based models for digital objects is the AIM@SHAPE project [14]. Due to its wide spectrum, it has raised a lot of interest among the semantic web community [15]. This is an avenue where the project may want to invest in the near future, once the animation ontology is mature enough. We are planning to use an ontology for static objects in the scene. This way the word 'Boeing', as well as 'Cessna' and airplane, will be mapped to the same 3D representation of a plane.

6 Architecture of Proposed System

This section describes in details how we are going to produce 3D animations which will carry all the textual information. We give an overview of the architecture and then detail the 2 modules which will produce the final animation. A description of the various data needed for the process is also given.

6.1 Architecture Components

Our system is based on 3 main components. The **Language Engine** produces a semantic form of the text. The **Text To Scene Engine** converts then this annotated text in an animatable format: the animatable scene. This is the input to the **Graphics Engine** which will produce a Collada file, consumable by any compliant viewer. An important part of the project is **the GITAN Repository**, containing partial Collada files which will be part of the final Collada scene.

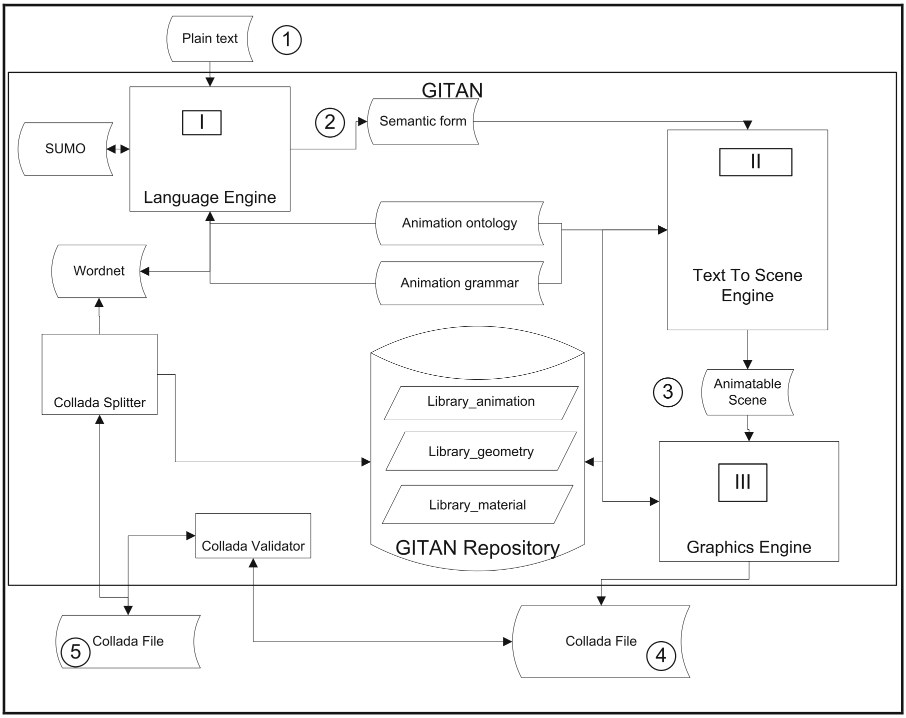


Fig. 1. Gitan's architecture

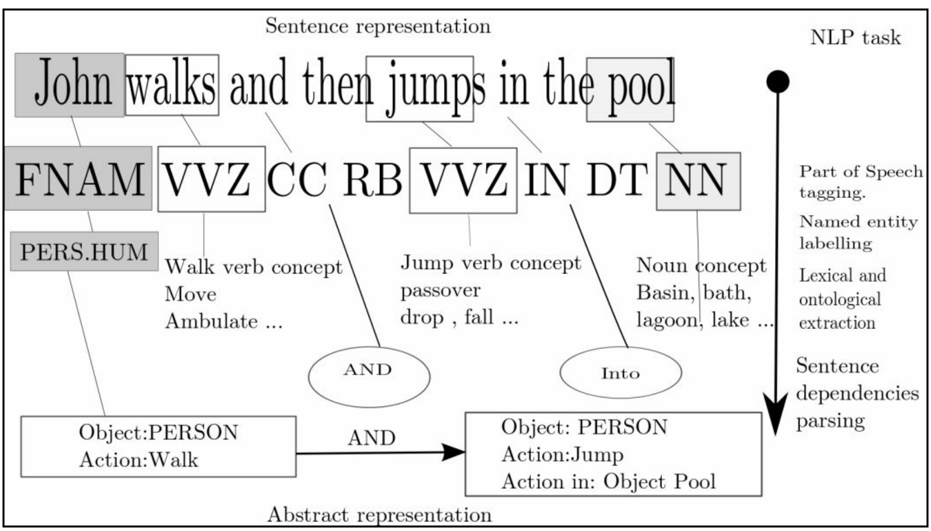


Fig. 2. The Language Engine

6.2 The Gitan Repository

The GITAN repository is the storage containing atomic elements of a Collada file encapsulated with GITAN specific meta-data. The geometries describe the mesh structure of the object, plus GITAN tags (specifying whether this geometry static or animatable is, for instance). The animations follow the Collada animation specifications [9], and contain also GITAN meta-data (like the origin of the Collada segment, its expected input for an activity). It contains all the elements to build a Collada file: cameras, lights,.. It must be noted that we aim to keep this repository as small as possible, by parametrizing as much as possible all the segments. We must be able to morph a cat into a tiger, provided the necessary information. The next section describes how the semantic form of the text is generated.

6.3 The Language Engine

The language engine transform a sentence into a semantic representation. Generic Natural Language Processing methods are used to achieve this transformation. First a labelling process involving a Part of Speech Tagger is applied. Part of speech helps to identify nature of textual information like verbs, adjectives and nouns in the perspective of their animation. Then a Named Entity (NE) labelling step is performed. This step allow identification of specific entities like Persons, Location, Products. Then each NE, verb and nouns are linked with an ontological description. This allows to associate its exact sense with a noun or an entity (i.e. A boat or a building defined in the ontology with their instance word inside the sentence), a specific movement with a verb (i.e. the sense of *course* for the verb *Jump* in a text context *of movement*) or to apply an attribute to a an object (i.e. using adjectives and link their description to their the concerned ontological instance). The labelling process and algorithms to instantiate relation between a word and its ontological representation have been experimented [16]. Final step involve hierarchical representation of terms dependence and identification of semantics sense inside the sentence (i.e. *the cat eats the mouse* or *the mouse is eaten by the cat* have the same sense but use different relation schema). This work have been investigated in [17].

6.4 The Text to Scene Engine

This module converts the semantic representation into an animatable description of the sentence. The major modules are the ones replacing absent entities and converting events into constraints and activities. The GITAN repository is searched in order to find an object or an animation fitting the ontological instance of a word. If present, a tag is created in the file. If absent, the GITAN ontology, using the OWL format [13], will provide a proper link to a potential candidate: if we want to instantiate a tiger in the animation for instance, and if the repository contains a cat, this module will morph the cat into a tiger by scaling it and adding a proper texture to it. The GITAN ontology provides the various parameters to be modified (i.e. size and texture). The same principles apply to the ontological instances of events. If the event is absent from the repository, the ontology will infer a proper parameterized substitute if possible. A walk cycle can potentially be used as a substitute for a jump cycle. Using the GITAN grammar, the TTS engine will link and build the tags in the animate text file. Using constraint calculus [18] , we aim to provide a valid transition for action verbs, and the global positioning of objects in the scene.

6.5 The Graphics Engine

Its goal is to assemble the pre-existing objects (defined as Collada segments in the GITAN repository) into an initial scene and apply the animations to them, according to the definitions from the Animatable Scene input file The most important parts of the engine are the constraint converter and the scene builder. The constraint converter will translate into a Collada format all the constraints and activities found in the Animated Text file. The scene builder will create the Collada scene, using objects stored in the GITAN repository. It will also assign the animations to the objects. The input data used by the graphics engine is made of 2 distinct types: The animatable scene and the Collada segments. The Collada segments are stored in the GITAN repository, and are programmatically included in the final Collada file.

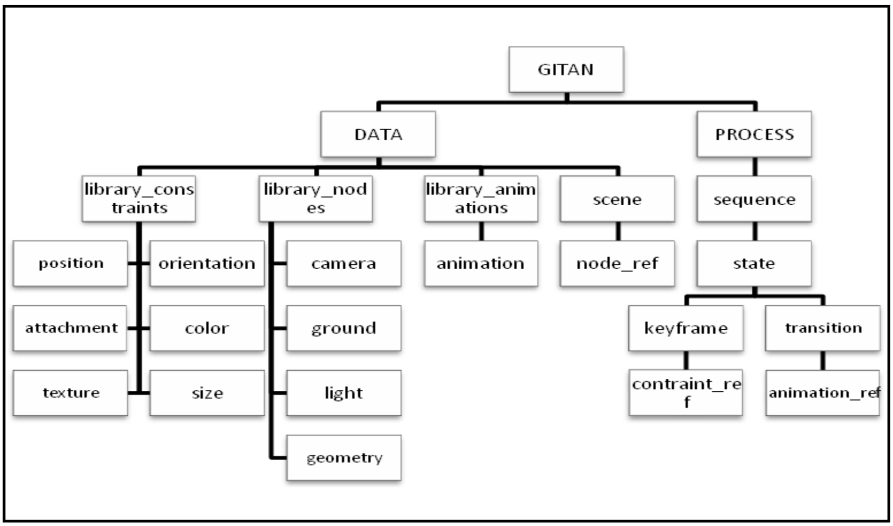


Fig. 3. The animated scene file

*The animated scene file***:** this input contains a high level scene description, based on a XML schema. This is a constraint based key-frame set, defining the various objects needed in the scene, and a set of key-frames, which will trigger animations and constraints. It is made of 2 sub parts: The *data* and the *proces***s** parts. Data describes all the needed objects, animations, for creating the scene. The process part specifies the timing of the animations and the constraints existing between static objects.

The data part contains sets of nodes, animation and constraints. The *nodes* define static objects from the GITAN repository. The *animations* can be a Collada animation; it defines then a predetermined GITAN animation, stored in the repository. A good example is a walk cycle, since such an animation is too complex to be described in a high level animation file. The *animation* can also be defined as a displacement. In this case it will be specified in the animated scene file, by its type (linear, ballistic ...), its speed. This part of the file also contains complete descriptions of the object's *constraints*. We identified 6 major constraint types affecting the position, the color, the size, the orientation, the texture of an object. There is also an attachment constraint, defining a link between 2 objects. An *animation grammar* will be produced which will convert the annotated verbs into animatable constraints. See figure 4 for details.

The process part defines how the scene is to be built and what constitutes the animation. We are using keyframed animation [19], since it's a common standard among DCC and Collada supports it. It contains 2 main components: the *Scene* and the *Sequence*. The Scene is a static description of the scene to be animated. The Sequence is made of *keyframes* and *transitions*. A *keyframe* defines the object's static constraints (positional, color ...) at a certain time. A *transition* defines how the objects change (its position, its shape...) from one keyframe to another. A transition must be made of at least *1 animation*, named the major animation. It can also contain secondary animation. Figure 4 illustrates the link between the various part of the file and the GITAN repository through a specific example ('the man jumps in the pool'). The <*Man*> node is linked to the 'Man' Collada geometry, and is positioned in the Keyframe 1. The <*Pool*> node has the same role. The <*walk-activity*> links the 'Walk cycle' Collada Anim1 to the Transition 1 where it will be executed, like the Translate Displacement, translating the <*Man*>. This creates a translation of the object Man, using a Collada walk cycle. At Keyframe2, the final Man position is set, through the

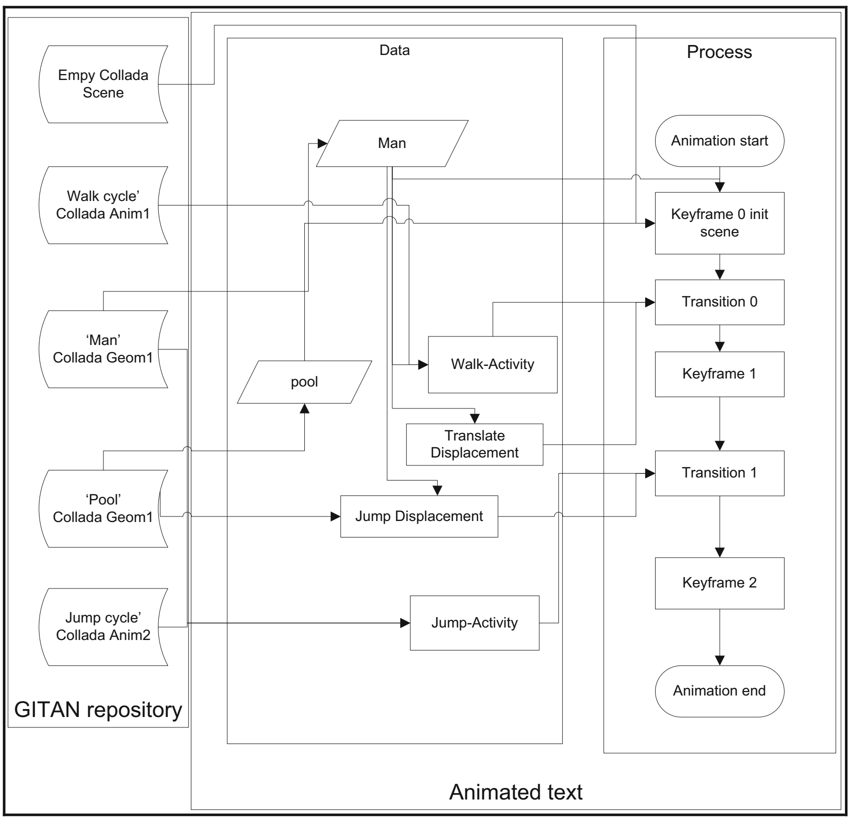


Fig. 4. The animatable scene file, the GITAN repository

<*Man Position Constraint2*>. The < *walk-activity*> is also linked to the <man> object, since this Collada animation must be applied to a valid Collada object, the 'Man Collada Geom1' in this case.

6.6 Development

Since the goal of the project is long term, and there are many areas of research which could influence its development, we will use a staged approach, based on an initial prototype working on a limited set of words. This will help validate one our original hypothesis (regarding the unification of text and image).

6.7 Initial Prototype

The first prototype will revolve around a specific use case: the teacher types a simple sentence. An animation is created. The teacher validates the animation, associated to a bag of word. The student sees the animation, and using the same set of words assembles them, in order to produce the same animation as the one he saw. This prototype will validate the Graphics engine, and the GITAN repository access. For this prototype all words and all the animations representing the verbs will be available in the GITAN repository;

6.8 Preliminary Results

We have validated the principles described by manually annotating all the elements needed by the Graphics engine. Here is

The original sentence: "John walks and then jumps in the pool"

The animated text contains the 2 parts *data* and *process:*

data

- **object 1**, 'john' , URI = c:\data\models\man.gitan
- **object2** , 'pool', URI = c:\data\models\pool.gitan
- **animation-activity1** : 'walk-cycle' , URI=c:\data\animation\biped-walk.gtn.
- **animation-activity2** : 'jump-cycle', URI = c:\data\anmation\jump-cycle.gtn
- **animation--displacement1**: translation, object1 //moves object1 linearly
- **animation-displacement2**: jump, object1 // moves object ballistically
- **constraint1**: 'position1' , object1. // place object1 at position1
- **constraint2**: 'position2', object1. . // place object1 at position2.
- **constraint3**: 'position3', object1. . // place object1 at position3
- **constraint4**: 'position1', object2. . // place object2 at position1

process

- **keyframe1: contraint1, constraint4.**
- **transition1** (10 seconds)
  - **animation-displacement1** (major , 4km/h)
  - **animation-activity1** (minor, 1 cycle/sec)
- **keyframe2: constraint2**
- **transition2** (20 seconds)
  - **animation-displacement2** (major, 6km/h)
  - **animation-activity2**(minor, 1 cycle/sec)
- **keyframe3: constraint4**

The *displacements* (jump, translation. etc...) are all predetermined, having specific parameters. A jump will have the distance and height attribute for instance. The *constraints* are of predetermined types: positional, colour, orientation. All those functionalities will be refined during the progress of the project.

7 Conclusions

With this project we want to liberate the communication between people of the textual or oral language limitations, by using 3D animations. This paper described the various modules which are going to be developed, for the graphical part, which will produce 3D animations. There are open questions that will be solved during the development of this project: how to maintain automatically the integrity of the Collada segments in the repository? Is the 'simple animation' (i.e. not realistic like a video game) paradigm conveying properly the communication intention between 2 persons? Is the proposed keyframe-transition model based on constraint calculus general enough ? the project starts and promises to be very exciting due to all the challenges we face. We want to thank Prompt [20] and the UnimaSoft Company [21], which are funding this project.

References

1. Aristotle, De Anima, 431a, 15-20
2. Kaoru, S., Mizue, N.: Animated Storytelling System via Text. In: ACE 2006, Hollywood, California, USA, June 14-16 (2006)
3. Dupuy, S., Egges, A., Legendre, V., Nugues, P.: Generating a 3D Simulation of a CarAccident From a Written Description in Natural Language: The CarSim System. In: Proc. of the Workshop on Temporal and Spatial Information Processing, ACL 2001 Conference, Toulouse, July 7, pp. 1–8 (2001)
4. Ma, M.: Automatic Conversion of Natural Language to 3D Animation, PhD thesis, University of Ulster
5. Bilasco, I., Villanova-Oliver, M., Gensel, J., Martin, H.: Semantic-based rules for 3D scene adaptation. In: Web3D 2007: Proceedings of the Twelfth International Conference on 3D Web Technology (April 2007)
6. Denis, M.: Imagery and thinking. In: Comoldi, C., McDaniel, M.A. (eds.) Imagery and Cognition. Springer, New York (1991)
7. Frege: F. L. G. Über Sinn und Bedeutung. Zeitschrift für Philosophie und philosophische Kritik, vol. 100, pp. 22–50; Trad. fr. Claude Imbert, Sens et dénotation, in Écrits logiques et philosophiques, pp. 102-126. Seuil, Paris (1971)
8. Wikipedia, `http://en.wikipedia.org/wiki/Animation`
9. COLLADA, `http://www.collada.org`
10. W3C schema definition page site, `http://www.w3.org/standards/xml/schema`
11. Wikipedia, `http://en.wikipedia.org/wiki/Storyboard`
12. Tom Gruber, `http://tomgruber.org/writing/ontology-definition-2007.htm`
13. OWL, `http://www.w3.org/TR/owl2-overview/`
14. AIM@SHAPE, `http://www.aimatshape.net/`
15. Linked Data, `http://linkeddata.org/`

16. Charton, E., Gagnon, M., Ozell, B.: Extension d'un système d'étiquetage d'entités nommées en étiqueteur sémantique. In: Proceedings of TALN 2010, Montréal, Canada (2010)
17. Zouaq, A., Gagnon, M., Ozell, B.: Semantic Analysis using Dependency-based Grammars and Upper-Level Ontologies. In: Proceedings of the 11th International Conference on Intelligent Text Processing and Computational Linguistics (2010)
18. Renz, J., Nebel, B.: Qualitative Spatial Reasoning using Constraint Calculi. In: Aiello, M., Pratt-Hartmann, I., van Benthem, J. (eds.) Handbook of Spatial Logics, pp. 161–215. Springer, Berlin (2007)
19. Dam, F.-V.: Computer Graphics, Principles and Practice, 2nd edn. Addison-Wesley, Reading (1990)
20. PROMPT, `http://www.promptinc.org/`
21. UnimaSoft, `http://www.unimasoft.com`

Visualization and Language Processing for Supporting Analysis across the Biomedical Literature

Carsten Görg[1], Hannah Tipney[2], Karin Verspoor[2], William A. Baumgartner Jr.[2], K. Bretonnel Cohen[2], John Stasko[1], and Lawrence E. Hunter[2]

[1] School of Interactive Computing & GVU Center, Georgia Institute of Technology, Atlanta, GA 30332
[2] Center for Computational Pharmacology, University of Colorado Denver School of Medicine, Aurora, CO 80045
goerg@cc.gatech.edu,
{Hannah.Tipney,Karin.Verspoor}@ucdenver.edu,
William.Baumgartner@ucdenver.edu, kevin.cohen@gmail.com,
stasko@cc.gatech.edu, Larry.Hunter@ucdenver.edu

Abstract. Finding relevant publications in the large and rapidly growing body of biomedical literature is challenging. Search queries on PubMed often return thousands of publications and it can be a tedious task to filter out irrelevant publications and choose a manageable set to read. We have developed a visual analytics system, named Bio-Jigsaw, which acts like a visual index on a document collection and supports biologists in investigating and understanding connections between biological entities. We apply natural language processing techniques to identify biological entities such as genes and pathways and visualize connections among them via multiple representations. Connections are based on co-occurrence in abstracts and also are drawn from ontologies or annotations in digital libraries. We demonstrate how Bio-Jigsaw can be used to analyze a PubMed search query on a gene related to breast cancer resulting in over 1500 primary papers.

Keywords: Visual analytics, investigative analysis, entity identification, language processing, biomedical literature.

1 Introduction

The emergence of biomedical technologies, such as microarrays, genome-wide association studies, and methods exploiting low-cost sequencing, has made the simultaneous observation of all gene products in a genome both easily accessible and routine. The ability to assay biological systems at a genomic scale has enabled the collection and documentation of biomedically relevant information at a level of unprecedented scale and detail. The resulting explosion of knowledge contains information critical to the advancement of biomedical research and the understanding of human health and disorder.

While many aspects of this knowledge is captured in structured form within freely available gene- and protein-centric databases (some 1,170 peer-reviewed databases were cataloged in 2009 [7]), the PubMed bibliographic database housed at the National

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 420–429, 2010.

Center for Biotechnology Information (NCBI) at the National Library of Medicine (NLM), remains the largest, most comprehensive source of biomedically important knowledge [20]. Peer-reviewed biomedical literature is not only the richest and most reliable of data sources, it is also the most overwhelming. Currently PubMed contains references to more than 19 million biomedical articles, and in 2008 expanded at a rate of approximately 2,200 new entries per day (calculated from PubMed 2008 indexed entries) [10].

This flood of knowledge has been accompanied by a breakdown of disciplinary boundaries, which traditionally made it possible to at least keep up with advances in a single field. Genome-scale research is intentionally broad, assaying potentially every gene in a genome, meaning pertinent prior knowledge could come from almost any biomedical discipline. Such blurring of boundaries means it is becoming increasingly common for new and important functions to be discovered for previously characterized genes, for example Relaxin 1 (RLN, NCBI Entrez Gene GeneID: 6013[1]), originally characterized as a cervical ripening hormone in the 1950s [8], has recently been implicated in processes as diverse as osteoarthritis and heart failure [16,19]. Therefore, not only is the volume of available biomedical knowledge captured in the literature growing at an unprecedented rate, but more and more of it is relevant to a larger number of biomedical scientists than ever before.

The challenge facing biomedical researchers is one of how to effectively and efficiently extract and interpret valuable and relevant knowledge trapped within this wealth of biomedical literature. Failure to do so can be extremely costly in terms of the wasted time, effort and money chasing weak leads, inadvertently duplicating already published results and missing important discoveries.

The current strategy most biomedical researchers use to identify articles of interest is to query PubMed using keywords and Boolean search terms via its online interface.[2] Keywords can include gene or protein names (*i.e.*, “ELN” or “Elastin”), diseases or disorders (*i.e.*, “cancer” or “Down Syndrome”), concepts or symptoms (*i.e.*, “pain” or “high glucose”), methodologies or technologies (*i.e.* “gel electrophoresis” or “microarray”), in addition to author names, publication dates and publication types. For each article a title and abstract is returned and the full text article can typically be accessed via a hyperlink. Currently, the only way for a researcher to parse the information returned to them is to manually read each abstract (and maybe even some of the full text articles) in an attempt to identify themes of interest.

Although popular, this strategy is challenging for a variety of reasons. Firstly, the sheer size of many of the document sets returned from these searches is frequently unmanageable; not only is it incredibly difficult and time consuming to manually read hundreds of abstracts, it is also virtually impossible to remember more than a handful of details and themes from such large document sets without taking notes or using other memory aids. Failure to identify and retain key themes and entities can result in knowledge being discarded as unimportant or missed completely. Secondly, within biomedicine the use of synonyms for gene and protein names is rife, which can cause confusion when trying to disambiguate between such entities within a large document

[1] `http://www.ncbi.nlm.nih.gov/gene`
[2] `http://www.ncbi.nlm.nih.gov/pubmed`

set. Thirdly, each time the scientist encounters a new concept or entity (such as a disease or gene) within the document set they must often invest additional time exploring supplementary datasources (*i.e.*, gene- or disease-centric databases such as EntrezGene or OMIM) to determine if these new terms are in fact of interest to them. Not only are these concepts and entities thought of as individual features, the investigating scientist will also be attempting to determine if, and what, relationships may exist between them.

Again, a failure to fully explore such relationships can result in the inappropriate disregard of critically important knowledge. Finally, when manually reading these document sets, the reader is inadvertently biased towards themes and concepts they already deem important, making it tricky to be aware of and receptive to knowledge from far outside their sphere of specialism. When a reader is unaware of just how much they do not know about a biomedical concept or process it is difficult to gauge when something is actually of interest.

In this paper, we describe a visual analytics system, Bio-Jigsaw, which supports biologists in investigating connections between biological entities grounded in the biomedical literature. The system identifies mentions of biological entities in text, specifically genes, using natural language processing strategies, visualizes relationships among those entities based on document co-occurrence, and allows a biological analyst to explore the documents from which those relationships are derived. We present an analysis scenario in which we demonstrate how the system supports biologists in exploring the biomedical literature.

2 Navigation of the Biomedical Literature

There are a number of tools available to the biomedical researcher to aid in navigation of the literature. GoPubMed[3] [5] supports the organization of the abstracts returned from a PubMed query according to Gene Ontology[4] and MeSH (Medical Subject Headings) concepts through recognition of those concepts in the text. Textpresso[5] [17] enables identification of terms from 33 categories, including genes, cells, phenotypes, cellular components, etc. through regular expressions that capture a significant amount of the variation in the surface form of those terms. It further supports document retrieval through queries that can combine categories with specific words, e.g. documents that mention two specific genes and a term belonging to the "regulation" category. The iHOP system[6] [9] provides an interface that links genes based on co-occurrence in sentences in PubMed abstracts, and provides hyperlinks among those sentences for navigation and exploration of literature relationships among genes. Reflect[7] is a Firefox plug-in which recognizes and highlights mentions of proteins and small molecules on any webpage, providing direct access to structured information about the highlighted entities through a pop-up window. Each of these tools plays some role in facilitating interpretation and navigation of the biomedical literature, by providing a more conceptual analysis of the content of the literature.

[3] http://www.gopubmed.org
[4] http://www.geneontology.org
[5] http://www.textpresso.org
[6] http://www.ihop-net.org/UniPub/iHOP
[7] http://www.reflect.ws

These tools are oriented towards supporting the user who is manually exploring the literature, and they do provide important assistance in identifying and organizing relevant literature. They do not, however, emphasize visual exploration of the document space. Out of these systems, only iHop takes any advantage of connections among documents, and none provide analysis over concepts in the document set as a whole. To support such analysis, specific biological facts must be extracted from the literature and represented in computable format. Such information extraction has been the subject of recent research. Protein-protein interactions have been the most common target for biological event extraction from the earliest studies [2,4] to the latest competitions like BioCreative II [15] and II.5 [14]. Research has also extended to other event types including those addressed in the recent BioNLP'09 challenge [13]: gene expression, transcription, protein catabolism, protein localization, binding, phosphorylation, and regulation. Other studies have addressed the extraction of gene-disease relations [3], protein residue annotation [18], among others.

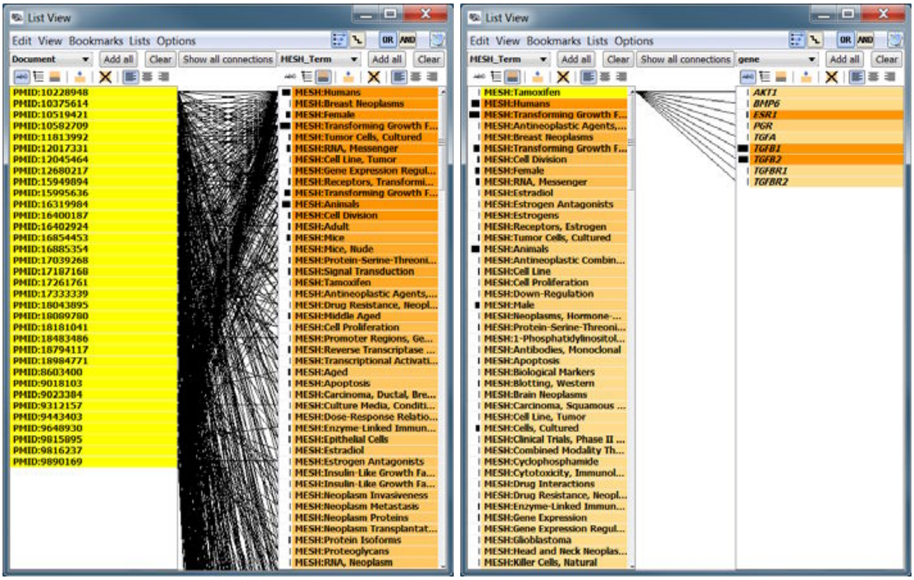
In this work, we treat co-occurrence of two genes within one publication as evidence of a relationship between them, without employing more detailed event or interaction extraction. We perform information extraction, specifically the gene mention detection and gene normalization algorithms of [1], to recognize occurrences of specific genes in PubMed abstracts and to associate them with the appropriate Entrez Gene identifier. It is our intention that in future work we will take advantage of our OpenDMAP concept recognition system [11] to identify more specific relationships among genes. However, with the proper weighting, literature co-occurrence is by itself a valuable indicator that we can take advantage of [6].

3 Visualizing Connections across the Biomedical Literature

We have developed a visual analytics system, Bio-Jigsaw, to support biologists in investigating connections between biological entities or concepts grounded in the biomedical literature. Bio-Jigsaw is a customized version of the Jigsaw system [21,12], tailored to the bio-informatics domain.

Connections in Bio-Jigsaw are based on co-occurrence in abstracts or drawn from ontologies and annotations in digital libraries. Bio-Jigsaw is a multiview system, including a number of different visualizations of the documents in the collection and the entities or concepts (genes, MeSH terms, KEGG pathways, GO biological processes, etc.) within those documents.

The two List Views in Figure 1 show connections between PubMed abstracts and MeSH terms and between MeSH terms and genes, respectively. Lists can be sorted alphabetically, by the frequency of occurrence in the whole document set (the larger an item's bar the more frequently it occurs), or by connection strength (the darker the shade of orange the stronger the connection). Figure 2 shows the Word Tree [22] for Tamoxifen. A Word Tree shows all occurrences of a word or phrase from the documents in the context of the words that follow it. The analyst can navigate through the tree by clicking on its branches. The Graph View in Figure 3 shows connections between documents (white rectangles) and entities (colored circles) using a node link diagram. The Document View in Figure 4 displays documents and highlights identified entities



(a) (b)

Fig. 1. The List View showing connections between PubMed abstracts, MeSH terms, and genes. Selected items are shown in yellow; connected items are in orange.

within them. The word cloud at the top shows the most frequent keywords from the set of currently loaded abstracts.

A textual search query interface allows users to find particular entities and the documents in which they occur. In addition, entities and documents can be explored directly by interacting with those objects in the views. For instance, new entities can be displayed and explored by user interface operations in the views that expand the context of entities and documents. In practice these two approaches are often combined: search queries serve to jump-start an exploration and view interaction then yields richer representations and exploration.

4 Scenario

In this section we walk through an analysis scenario to demonstrate how Bio-Jigsaw supports biologists in exploring the biomedical literature. We also provide a video that demonstrates the scenario actions in more detail.[8]

After looking at clinical breast cancer data our analyst has become interested in the tgfb2 gene. After searching PubMed for 'tgfb2', over 1500 primary papers are returned to her. To start exploring these papers in Bio-Jigsaw, she opens a List View and searches

[8] The video is available at
`http://www.gvu.gatech.edu/ii/jigsaw/BioJigsaw.avi`

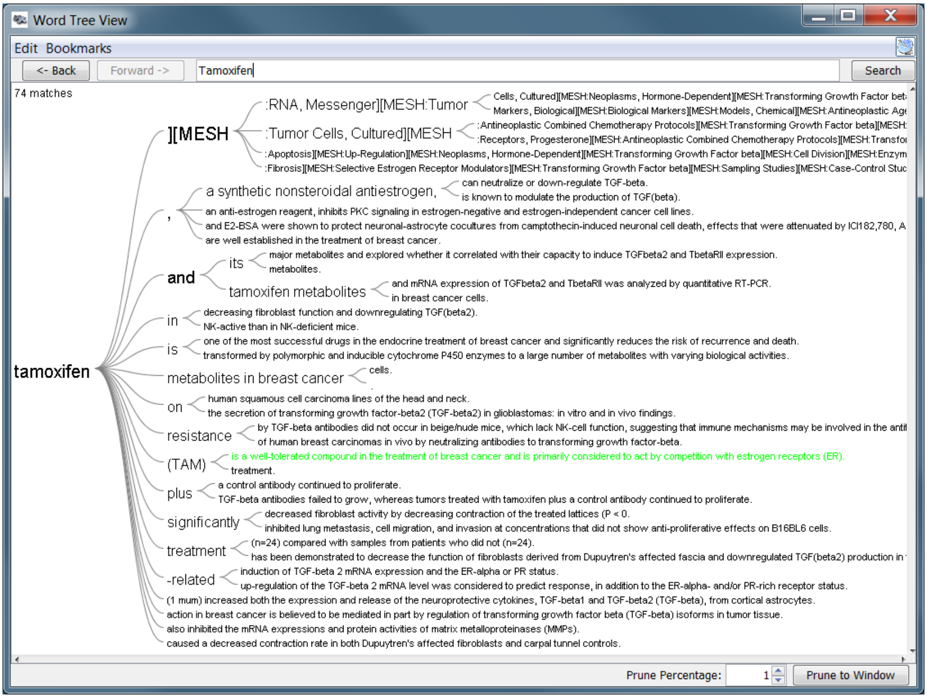


Fig. 2. Word Tree for Tamoxifen across a collection of journal articles. All trailing words are shown, sized by their frequency of occurrence.

for documents containing the phrase ‘breast cancer’; 34 documents are found and displayed in the left column of the List View (see Figure 1(a)). The analyst then views the MeSH terms (a controlled vocabulary used to describe the content of biomedical articles within PubMed[9]) associated with the 34 documents in a second list and sorts them by connection strength. She identifies the term “Tamoxifen” as being of interest because although it is not frequently observed across the document collection as a whole (as illustrated by the short bar), it is well connected to the 34 selected documents (dark color).

The term Tamoxifen is unknown to the analyst. By launching the Word Tree View she investigates this term further (see Figure 2). The Word Tree displays all sentences from the document set containing the word Tamoxifen, grouped by common suffixes. The analyst can quickly see that Tamoxifen is a compound used during the treatment of breast cancer.

By adding an additional list to the List View, the analyst next explores other entities that are associated with the subset of documents identified by the phrase breast cancer. The analyst wants to know if other genes are also affected by Tamoxifen so she displays the gene entity in the new list (see Figure 1(b)). She can see from their highlighting that

---

[9] `http://www.nlm.nih.gov/mesh`

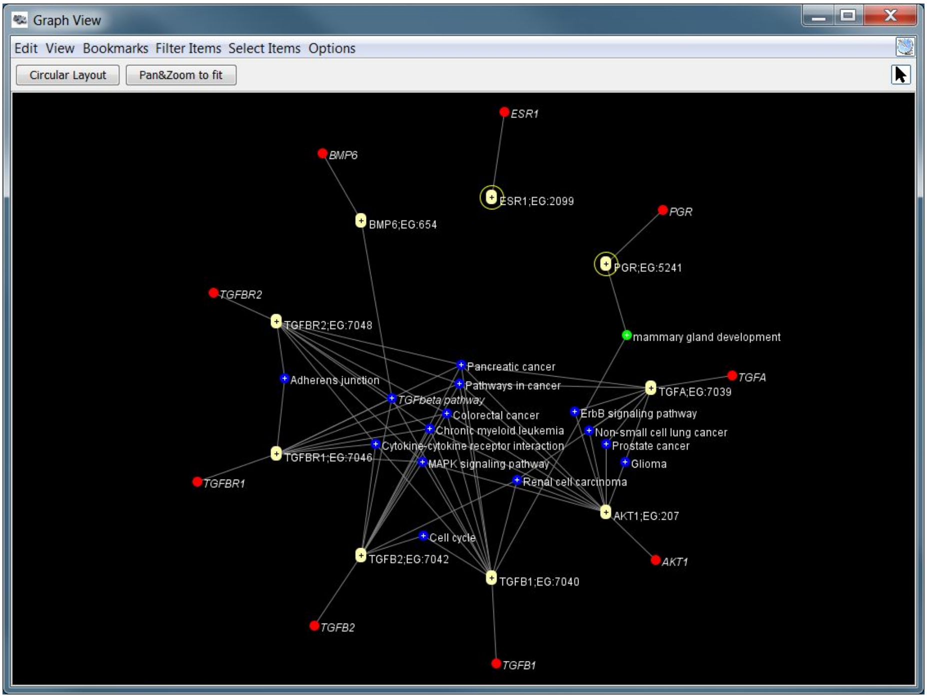


Fig. 3. The Circular Layout in the Graph View showing connections between ontology documents (white rectangles) and KEGG pathways. All entities (colored circles) connecting to more than one document are drawn in the middle making it easier to focus on them.

three genes (tgfb1, tgfb2, and esr1) are particularly well connected to Tamoxifen within this document set.

Now the analyst turns to the Graph View (see Figure 3) and displays all the genes (red nodes) connected to Tamoxifen along with their ontology documents (yellow nodes), as well as connected KEGG pathways (blue nodes) and GO biological processes (green nodes). By applying the circular layout (documents are displayed on a circle), entities connected to only one document are positioned outside the circle, while entities connected to multiple documents are shown inside the circle; the more connections entities have, the closer to the center they are positioned. The analyst then filters out irrelevant entities and notices that although a number of cancer associated KEGG pathways are shared by the tgfb family and their receptors, the progesterone and estrogen receptors (highlighted with yellow circles) do not share these annotations. However, the GO biological process term "mammary gland development" is common to both tgfb and the progesterone receptor and the analyst wonders if the progesterone receptor has a role in breast cancer which is affected by Tamoxifen and if it is mediated by members of the tgfb family.

Now the analyst turns back to the List View to find the documents most connected to the tgfb genes, their receptors, and the progesterone receptor. She displays the eight

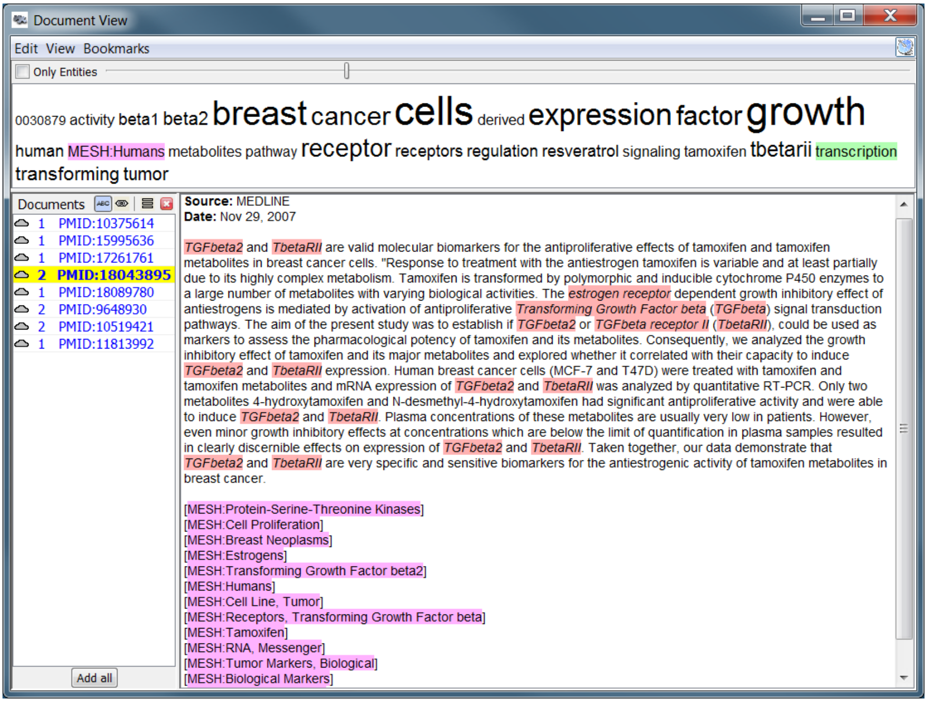


Fig. 4. The Document View showing a set of abstracts relevant to tgfb2 (list on the left side). The word cloud at the top shows the most frequent keywords from the set of abstracts. Document PMID:18043895 is selected and its abstract and related terms are visible. Entities are highlighted: genes in red and MeSH terms in purple.

most connected documents in the Document View (see Figure 4) to read their abstracts. She learns not only that Tamoxifen increases the expression of tgfb2 resulting in inhibition of breast cancer growth, but also that the steroid receptor status of tumors (*i.e.*, if the tumor expresses progesterone receptors or not) can be predictive of the tumors response to Tamoxifen.

5 Conclusion

In this paper, we introduced the Bio-Jigsaw system which integrates natural language processing for entity identification and normalization in the biomedical literature with sophisticated visualization strategies. We provided an example of how the visualization of relationships among entities occurring in a document set can be navigated and explored by a biological analyst to gain new insights into a gene of interest. Through the visualization, the analyst is able to quickly identify new concepts that are relevant to the gene under investigation, and to hone in on concepts that are strongly indicated by the document set. The different views that the system makes available each play an important role in exploring the document set, and the analyst can move smoothly among them as well as manipulate which entities and concepts the visual analysis focuses on. Through the use of this tool, the biologist has the ability to navigate and explore the biomedical literature in such a way that they can much more effectively extract, manipulate, and interpret the knowledge that exists there.

Acknowledgments

This research is based upon work supported in part by the National Science Foundation via Award IIS-0915788 and by the U.S. Department of Homeland Security's VACCINE Center under Award Number 2009-ST-061-CI0001. We also acknowledge the support of NIH National Library of Medicine grants 2R01LM009254 and 2R01LM008111 to Lawrence Hunter, and NIH grant 1R01LM010120-01 to Karin Verspoor.

References

1. Baumgartner Jr., W.A., Lu, Z., Johnson, H.L., Caporaso, J.G., Paquette, J., Lindemann, A., White, E.K., Medvedeva, O., Cohen, K.B., Hunter, L.: Concept recognition for extracting protein interaction relations from biomedical text. Genome Biology 9 (in press)
2. Blaschke, C., Andrade, M.A., Ouzounis, C., Valencia, A.: Automatic extraction of biological information from scientific text: protein-protein interactions. Intelligent Systems for Molecular Biology, 60–67 (1999)
3. Chun, H.W., Tsuruoka, Y., Kim, J.D., Shiba, R., Nagata, N., Hishiki, T., Tsujii, J.: Extraction of gene-disease relations from medline using domain dictionaries and machine learning. In: Pacific Symposium on Biocomputing, pp. 4–5 (2006)
4. Craven, M., Kumlien, J.: Constructing biological knowledge bases by extracting information from text sources. Intelligent Systems for Molecular Biology, 77–86 (1999)
5. Doms, A., Schroeder, M.: GoPubMed: exploring PubMed with the Gene Ontology. Nucleic Acids Research 33, 783–786 (2005), GoPubMed
6. Gabow, A., Leach, S.M., Baumgartner Jr., W.A., Hunter, L.E., Goldberg, D.S.: Improving protein function prediction methods with integrated literature data. BMC Bioinformatics 9(198) (2008)
7. Galperin, M.Y., Cochrane, G.R.: Nucleic acids research annual database issue and the nar online molecular biology database collection in 2009. Nucleic Acids Research 37, D1–D4 (2009)
8. Graham, E.F., Dracy, A.E.: The effect of relaxin – mechanical di-latation of the bovine cervix. Journal of Dairy Science 36, 772–777 (1953)
9. Hoffmann, R., Valencia, A.: A gene network for navigating the literature. Nat. Genet. 736(7) (July 2004)
10. Hunter, L., Cohen, B.: Biomedical language processing: what's beyond PubMed? Mol. Cell. 21(5), 589–594 (2006)
11. Hunter, L., Lu, Z., Firby, J., Baumgartner Jr., W.A., Johnson, H.L., Ogren, P.V., Cohen, K.B.: OpenDMAP: An open-source, ontology-driven concept analysis engine, with applications to capturing knowledge regarding protein transport, protein interactions and cell-specific gene expression. BMC Bioinformatics 9(78) (2008)
12. Kang, Y., Görg, C., Stasko, J.: The evaluation of visual analytics systems for investigative analysis: Deriving design principles from a case study. In: IEEE VAST, pp. 139–146 (October 2009)

13. Kim, J.-D., Ohta, T., Pyysalo, S., Kano, Y., Tsujii, J.: Overview of BioNLP 2009 shared task on event extraction. In: BioNLP 2009 Companion Volume: Shared Task on Entity Extraction, pp. 1–9 (2009)
14. Krallinger, M., Leitner, F., Valencia, A.: The BioCreative II.5 challenge overview. In: Proceedings of the BioCreative II.5 Workshop 2009 on Digital Annotations (2009)
15. Krallinger, M., Morgan, A., Smith, L., Leitner, F., Tanabe, L., Wilbur, J., Hirschman, L., Valencia, A.: Evaluation of text-mining systems for biology: overview of the second biocreative community challenge. Genome Biology 9(suppl. 2), S1 (2008)
16. Kupari, M., Mikkola, T.S., Turto, H., Lommi, J.: Is the pregnancy hormone relaxin an important player in human heart failure? Eur. J. Heart Fail 7, 195–198 (2005)
17. Muller, H.M., Kenny, E.E., Sternberg, P.W.: Textpresso: an ontology-based information retrieval and extraction system for biological literature. PLoS Biol. 2(11) (2004)
18. Nagel, K., Jimeno-Yepes, A., Rebholz-Schuhmann, D.: Annotation of protein residues based on a literature analysis: cross-validation against uniprotkb. BMC Bioinformatics 10(suppl. 8), S4 (2009)
19. Santora, K., Rasa, C., Visco, D., Steinetz, B.G., Bagnell, C.A.: Antiarthritic effects of relaxin, in combination with estrogen, in rat adjuvant-induced arthritis. J. Pharmacol. Exp. Ther. 322, 887–893 (2007)
20. Sayers, E.W., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., DiCuccio, M., Edgar, R., Federhen, S., Feolo, M., Geer, L.Y., Helmberg, W., Kapustin, Y., Landsman, D., Lipman, D.J., Madden, T.L., Maglott, D.R., Miller, V., Mizrachi, I., Ostell, J., Pruitt, K.D., Schuler, G.D., Sequeira, E., Sherry, S.T., Shumway, M., Sirotkin, K., Souvorov, A., Starchenko, G., Tatusova, T.A., Wagner, L., Yaschenko, E., Ye, J.: Database resources of the National Center for Biotechnology Information. Nucl. Acids Res. 37(suppl. 1),D5–D15 (2009)
21. Stasko, J., Görg, C., Liu, Z.: Jigsaw: supporting investigative analysis through interactive visualization. Information Visualization 7(2), 118–132 (2008)
22. Wattenberg, M., Viégas, F.B.: The word tree, an interactive visual concordance. IEEE Transactions on Visualization and Computer Graphics 14(6), 1221–1228 (2008)

SceneMaker: Multimodal Visualisation of Natural Language Film Scripts

Eva Hanser[1], Paul Mc Kevitt[1], Tom Lunney[1], Joan Condell[1], and Minhua Ma[2]

[1] School of Computing & Intelligent Systems
Faculty of Computing & Engineering
University of Ulster, Magee
Derry/Londonderry BT48 7JL, Northern Ireland
[2] School of Computing
Faculty of Business, Computing and Law
University of Derby
Derby DE22 1GB, England
hanser-e@email.ulster.ac.uk,
{p.mckevitt,tf.lunney,j.condell}@ulster.ac.uk,
m.ma@derby.ac.uk

Abstract. Producing plays, films or animations is a complex and expensive process involving various professionals and media. Our proposed software system, *SceneMaker*, aims to facilitate this creative process by automatically interpreting natural language film scripts and generating multimodal, animated scenes from them. During the generation of the story content, SceneMaker gives particular attention to emotional aspects and their reflection in fluency and manner of actions, body posture, facial expressions, speech, scene composition, timing, lighting, music and camera work. Related literature and software on Natural Language Processing, in particular textual affect sensing, affective embodied agents, visualisation of 3D scenes and digital cinematography are reviewed. In relation to other work, SceneMaker follows a genre-specific text-to-animation methodology which combines all relevant expressive modalities and is made accessible via web-based and mobile platforms. In conclusion, SceneMaker will enhance the communication of creative ideas providing quick pre-visualisations of scenes.

Keywords: Natural Language Processing, Text Layout Analysis, Intelligent Multimodal Interfaces, Affective Agents, Genre Specification, Automatic 3D Visualisation, Affective Cinematography, SceneMaker.

1 Introduction

The production of movies is an expensive process involving planning, rehearsal time, actors and technical equipment for lighting, sound and special effects. It is also a creative act which requires experimentation, visualisation of ideas and

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 430–439, 2010.

their communication between everyone involved, e.g., playwrights, directors, actors, cameramen, orchestra, managers and costume and set designers. We are developing a software system, *SceneMaker*, which will assist in this production process. SceneMaker will provide a facility to pre-visualise scenes before putting them into action. Users input a natural language (NL) text scene and automatically receive multimodal 3D visualisations. The objective is to give directors or animators a reasonable idea of what a scene will look like. The user can refine the automatically created output through a script and 3D editing interface, accessible over the internet and on mobile devices. Such technology could be applied in the training of those involved in scene production without having to utilise expensive actors and studios. Additionally, SceneMaker could be used for rapid visualisation of ideas and concepts in advertising agencies. SceneMaker will extend an existing software prototype, CONFUCIUS [1], which provides automated conversion of single natural language sentences to multimodal 3D animation of characters' actions and camera placement. SceneMaker will focus on the precise representation of emotional expression in all modalities available for scene production and genre-sensitive art direction. To achieve this, SceneMaker will include new tools for text layout analysis of screenplays, commonsense and affective knowledge bases for context understanding, affective reasoning and automatic genre specification. This research focuses on three research questions: How can emotional information be computationally recognised in screenplays and structured for visualisation? How can emotional states be synchronised in presenting all relevant modalities? Can compelling, life-like and believable animations be achieved?

Section 2 gives an overview of current research on computational, multimodal and affective scene production. In section 3, the design of SceneMaker is discussed. SceneMaker is compared to related multimodal work in section 4 and Section 5 gives the conclusion and future work.

2 Background

Automatic and intelligent production of film/theatre scenes with characters expressing emotional states involves four development stages:

1. Detecting personality traits and emotions in the film script
2. Modelling affective 3D characters, their expressions and actions
3. Visualisation of scene environments according to emotional findings
4. Development of a multi-modal user interface and mobile application.

This section reviews state-of-the-art advances in these areas.

2.1 Detecting Personality and Emotions in Film Scripts

All modalities of human interaction express personality and emotional states namely voice, word choice, gestures, body posture and facial expression. In order to recognise emotions in text and to create life-like characters, psychological theories for emotion, mood, personality and social status are translated into computable methods, e.g., Ekman's 6 basic emotions [2], the Pleasure-Dominance-Arousal model (PAD) [3] with intensity values or the OCC model (Ortony-Clore-Collins) [4] with cognitive grounding and appraisal rules. Word choice is a useful indicator for the personality of a story character, their social situation, emotional state and attitude. Different approaches to textual affect sensing are able to recognise explicit affect words such as keyword spotting and lexical affinity [5], machine learning methods [6], hand-crafted rules and fuzzy logic systems [7] and statistical models [6]. Commonsense knowledge-based approaches [8,9] and a cognitive inspired model [10] include emotional context evaluation of non-affective words and concepts. Strict formatting of screenplays eases machine parsing of scripts and facilitates detection of semantic context information for visualisation. Through text layout analysis of capitalisation, indentation and parentheses, elements such as dialog, location, time, present actors, actions and sound cues can be visually recognised and directly mapped into XML-presentations [11].

2.2 Modelling Affective Embodied Agents

Research aiming to automatically model and animate virtual humans with natural expressions faces challenges not only in automatic 3D character transformation, synchronisation of face expressions, e.g., lips and gestures with speech, path finding and collision detection, but furthermore in the refined sensitive execution of each action. The exact manner of an affective action depends on intensity, fluency, scale and timing and impacts on the viewer's interpretation of the behaviour. Various scripting languages specifically cater for the modelling of the detected emotions and affective behaviour characteristics. Non-verbal behaviour of avatars is automatically modelled from conversational text with the Behaviour Expression Animation Toolkit (BEAT) [12]. Based on the analysis of linguistics and context of dialogue scripts appropriate Multimodal Presentation Mark-up Language (MPML) [13] annotations are automatically added to model speech synthesis, facial and body animations of 3D agents. SCREAM (Scripting Emotion-based Agent Minds) [14] is a web-based scripting tool for multiple characters which computes affective states based on the OCC-Model [4] of appraisal and intensity of emotions, as well as social context. ALMA (A Layered Model of Affect) [15] implements AffectML, an XML based modelling language which incorporates the concept of short-term emotions, medium-term moods and long-term personality profiles. In [7] the OCEAN (Openness, Conscientiousness, Extroversion, Agreeableness and Neuroticism) personality model [16], Ekman's basic [2] emotions and a model of story character roles are combined through a fuzzy rule-based system to decode the meaning of scene descriptions and to control the affective state and body language of the characters. Su et al. [7] map personality and emotion output to graphics and animations. Postural values for four main body areas manipulate the shape, geometry and motion of the character model considering physical characteristics of space, timing, velocity, position, weight and portion of the body. Embodied Conversational Agents (ECA) are capable of real-time face-to-face conversations with human users or other agents, generating and understanding NL and body movement. The virtual human, Max [17], engages museum visitors in small talk. Max listens while the users type their input, reasons about actions to take, has intention and goal plans, reacts emotionally and gives verbal and non-verbal feedback. Greta [18] is modelled as an expressive multimodal ECA. Affective Presentation Markup Language (APML) defines her facial expressions, hand and arm gestures for different communicational functions and with varying degrees of expressivity (manner). The behaviours are synchronised with the duration of phonemes in speech. Multimodal annotation coding of video or motion captured data specific to emotion collects data in publicly available facial expression or body gesture databases [19]. The captured animation data can be mapped to 3D models, which is useful for instructing characters precisely on how to perform desired actions.

2.3 Visualisation of 3D Scenes and Virtual Theatre

Visual and acoustic elements involved in composing a virtual story scene, the construction of the 3D environment or set, scene composition, automated cinematography and the effect of genre styles are addressed in complete text-to-visual systems and scene directing systems. Scene visualisation requires consideration of the positioning and interaction of actors and objects, the camera view, light sources and audio like background noises or music. SONAS [20] constructs a three-dimensional virtual town according to the verbal descriptions of a human user. WordsEye [21] depicts non-animated 3D scenes with characters, objects, actions and environments. A database of graphical objects holds 3D models, their attributes, poses, kinematics and spatial relations. In CONFUCIUS [1], multimodal 3D animations of single sentences are produced. 3D models perform actions, dialogues are synthesised and basic cinematic principles determine the camera placement. Another modality, cinematography, can assist in conveying themes and moods in animations. Film techniques are automatically applied to existing animations in [22]. Reasoning about plot, theme, character actions, motivations and emotions, cinematic rules are followed, which define appropriate placement and movement of camera, lighting, colour schemes and the pacing of shots. A high-level synchronised Expression Mark-up Language (EML) [23] integrates environmental expressions like cinematography, illumination and music as a new modality into the emotion synthesis of virtual humans. ScriptViz [24] renders 3D scenes from NL screenplays immediately during the writing process, extracting verbs and adverbs to interpret events and states in sentences.

The time and environment, the theme and the emotional tone of a film's story classify different genres with distinguishable presentation styles. Commonly, genres are categorised into, e.g., action, comedy, drama, horror and romance. Genre is reflected in the detail of a production, exaggeration and fluency of movements, pace (shot length), lighting, colour and camerawork. These parameters are responsible for an appropriate affective viewer impression. Cinematic principles in different genres are investigated in [25]. Dramas and romantic movies are slower paced with longer dialogues, whereas action movies have rapidly changing, shorter shot length. Comedies tend to be presented in a large spectrum of bright colours, whereas horror films adopt mostly darker hues. The automatic 3D animation production system, CAMEO [26], incorporates direction knowledge, like genre and cinematography, as computer algorithms and data to control camera, light, audio and character motions. A system which automatically recommends music based on emotion is proposed by [27]. Associations between emotions and music features in movies are discovered by extracting chords, rhythm and tempo of songs.

2.4 Multimodal Interfaces and Mobile Applications

Multimodal human-computer interaction exists on mobile devices, such as SmartKom Mobile [28]. Script writing tools assist the writing process of screenplays, like ScriptRight [29] for mobile devices. The Virtual Theatre Interface project [30] offers a web-based user interface to manipulate actors' positions on stage, lighting and audience view points. Editing and manipulating virtual characters on mobile devices is tested on the 'mobile animator' interface [31]. 3D objects and camera are directly controllable on the handheld display by selecting and dragging body parts using virtual manipulators with a stylus. To maintain the level of detail of the 3D characters for display on bigger screens, handheld users are provided with a simplified representation. Bounding boxes roughly section the 3D model. The user can zoom into the simplified sections to access the full set of joints.

This wide range of approaches to modelling emotions, moods and personality aspects in virtual humans and scene environments along with first attempts to bring multi-modal agents onto mobile devices provide a sound basis for SceneMaker.

3 Design of SceneMaker

Going beyond the animation of explicit events, SceneMaker will apply Natural Language Processing (NLP) methods to screenplays to automatically extract and visualise emotions, moods and film genre. SceneMaker will be tested by augmenting short 3D scenes with affective influences on the body language of actors and environmental expression, like illumination, timing, camera work, music and sound automatically directed according to the genre style.

3.1 SceneMaker Architecture

SceneMakers's architecture is shown in Fig. 1. The key component is the *scene production module* including modules for understanding, reasoning and multimodal visualisation situated on a server. The *understanding module* performs natural language processing and text layout analysis of the input text. The *reasoning module* interprets the context based on common, affective and cinematic knowledge bases, updates emotional states and creates plans for actions, their manners and the representation of the set environment. The *visualisation module* maps these plans to 3D animation data, selects appropriate 3D models from

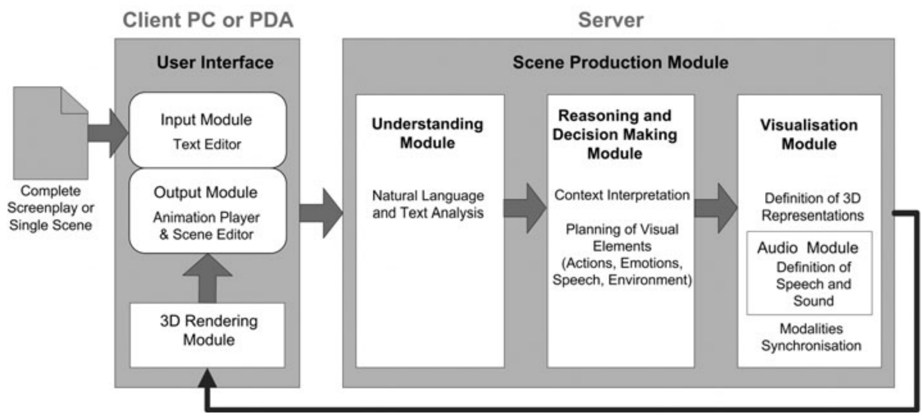


Fig. 1. SceneMaker architecture

the graphics database, defines their body motion transitions, instructs speech synthesis, selects sound and music files from the audio database and assigns values to camera and lighting parameters. The visualisation module synchronises all modalities into an animation script. The online user interface, available via computers and mobile devices, consists of two parts. The input module provides assistance for film script writing and editing and the output module renders the 3D scene according to the manuscript and allows manual scene editing to fine-tune the automatically created animations.

3.2 Implementation of SceneMaker

Multimodal systems automatically mapping text to visuals face challenges in interpreting human language which is variable, ambiguous, imprecise and relies on common knowledge between the communicators. Enabling a machine to understand a natural language text involves feeding the machine with grammatical structures, semantic relations and visual descriptions to be able to match suitable graphics. Existing software tools fulfilling sub-tasks will be modified, combined and extended for the implementation of SceneMaker. For the interpretation of the input scripts, SceneMaker will build upon the NLP module of CONFUCIUS [1], but a pre-processing tool will first decompose the layout structure of the input screenplay script. In the NLP module of CONFUCIUS the syntactic knowledge base parses the input text and identifies grammatical word types, e.g., noun, verb, adjective or other, with the Connexor Part of Speech Tagger [32] and determines their relation in a sentence, e.g., subject, verb and object with Functional Dependency Grammars [33]. The Semantic knowledge base (WordNet [34] and LCS database [35]) and temporal language relations will be extended by an emotional knowledge base, e.g., WordNet-Affect [36], and context reasoning with ConceptNet [9] to enable an understanding of the deeper meaning of the context and emotions. In order to automatically recognise genre, SceneMaker will identify keyword co-occurences and term frequencies and determine the length of dialogues, sentences and scenes/shots. Cinematic knowledge will then adjust the cinematic settings according to the determined genre. For instance, the light will be brighter in scenes of a comedy movie as compared to the scenes of a horror movie.

In SceneMaker, the visual knowledge of CONFUCIUS, such as object models and event models, will be related to emotional cues. CONFUCIUS' basic cinematic principles will be extended and classified into expressive and genre-specific categories. EML [23] appears to be a comprehensive XML-based scripting language to model expressive modalities including body language as well as cinematic annotations. Resources for 3D models are H-Anim models [37] which include geometric or physical, functional and spatial properties. As in the Mobile Animator [31], the animation editor will provide virtual manipulators to modify the automatically animated H-Anim models manually. The user will be able to choose characters, objects, light sources and camera for the manipulation of geometry and animation in a selected time frame. Further fine tuning can be achieved through re-editing the input script. For the speech generation from dialogue text, the speech synthesis module used in CONFUCIUS, FreeTTS [38], will be tested for its suitability in SceneMaker with regard to mobile applications and the effectiveness of emotional prosody. An automatic audio selection tool, as in [27], will be added for intelligent, affective selection of sound and music according the theme and mood of a scene. Test scenarios will be developed based on screenplays of different genres and animation styles, e.g., drama films, which include precise descriptions of set layout and props versus comedy, which employs techniques of exaggeration for expression. The effectiveness and appeal of the scenes created in SceneMaker will be evaluated against hand-animated scenes and existing feature film scenes. The functionality and usability of SceneMaker's components and the GUI will be tested in cooperation with professional film directors, comparing the process of directing a scene traditionally with actors or with SceneMaker.

4 Relation to Other Work

Research implementing various aspects of modelling affective virtual actors, narrative systems and film-making applications relates to SceneMaker. CONFUCIUS [1] and ScriptViz [24] realise text-to-animation systems from natural language text input, but they do not enhance the visualisation through affective aspects, the agent's personality, emotional cognition or genre specific styling. Their animation is built from well-formed single sentences and does not consider the wider context. SceneMaker will facilitate animation modelling of sentences, scenes or whole scripts. Single sentences require more reasoning about default settings and more precision will be achieved from collecting context information from longer passages of text. SceneMaker will introduce text layout analysis to derive semantic content from the particular format of screenplays scripts. Emotion cognition and display will be related to commonsense knowledge. No previous storytelling system controls agent behaviour through integrating all of personality, social status, narrative roles and emotions. Only EML [23] combines multimodal character animation with film making practices based on an emotional model, but it does not consider personality types or genre. CAMEO [26] is the only system relating specific cinematic direction, for character animation, lighting and camera work, to the genre of a given story, but genre types are explicitly selected by the user. SceneMaker will introduce a new approach to automatically recognise genre from script text with keyword co-occurrence, term frequency and calculation of dialogue and scene length. SceneMaker will bring all relevant techniques together to form a software system for believable affective computational animation production from NL scene scripts. SceneMaker will present a web/mobile based user interface for directors or animators to directly edit scenes.

5 Conclusion and Future Work

The software system, SceneMaker, which automatically visualises affective expressions of screenplays, aims to advance knowledge in the areas of affective computing, digital storytelling and expressive multimodal systems. SceneMaker, contributes to believability and artistic quality of automatically produced animated, multimedia scenes. Existing systems solve partial aspects of NLP, emotion modelling and multimodal storytelling. SceneMaker focuses on semantic interpretation of screenplays scripts, the computational processing of emotions, virtual agents with affective behaviour and expressive scene composition including emotion-based audio selection. In relation to other work, SceneMaker will incorporate an expressive model for multiple modalities, including prosody, body language, acoustics, illumination, staging and camera work. Emotions will be inferred from context. Genre types will be automatically derived from the scene scripts and influence the design style of the output animation. SceneMaker's 3D output will be available to users for edit. SceneMaker's mobile, web-based user interface will assist directors, drama students, writers and animators in the testing of their ideas. Accuracy of animation content, believability and effectiveness of expression and usability of the interface will be evaluated in empirical tests comparing manual animation, feature film scenes and real-life directing with SceneMaker. In conclusion, this research intends to automatically produce multimodal animations with heightened expressivity and visual quality from screenplay script input.

References

1. Ma, M.: Automatic Conversion of Natural Language to 3D Animation. PhD Thesis, School of Computing and Intelligent Systems, University of Ulster (2006)
2. Ekman, P., Rosenberg, E.L.: What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system. Oxford University Press, Oxford (1997)

3. Mehrabian, A.: Framework for a Comprehensive Description and Measurement of Emotional States. Genetic, Social, and General Psychology Monographs 121(3), 339–361 (1995)
4. Ortony, A., Clore, G.L., Collins, A.: The Cognitive Structure of Emotions. Cambridge University Press, Cambridge (1988)
5. Francisco, V., Hervás, R., Gervás, P.: Two Different Approaches to Automated Mark Up of Emotions. In: Sattar, A., Kang, B.-h. (eds.) AI 2006. LNCS (LNAI), vol. 4304, pp. 101–114. Springer, Heidelberg (2006)
6. Strapparava, C., Mihalcea, R.: Learning to identify emotions in text. In: Proceedings of the 2008 ACM Symposium on Applied Computing. SAC 2008, pp. 1556–1560. ACM, New York (2008)
7. Su, W.-P., Pham, B., Wardhani, A.: Personality and Emotion-Based High-Level Control of Affective Story Characters. IEEE Transactions on Visualization and Computer Graphics 13(2), 281–293 (2007)
8. Liu, H., Lieberman, H., Selker, T.: A model of textual affect sensing using real-world knowledge. In: Proceedings of the 8th International Conference on Intelligent User Interfaces. IUI 2003, pp. 125–132. ACM, New York (2003)
9. Liu, H., Singh, P.: ConceptNet: A practical commonsense reasoning toolkit. BT Technology Journal 22(4), 211–226 (2004)
10. Shaikh, M.A.M., Prendinger, H., Ishizuka, M.: A Linguistic Interpretation of the OCC Emotion Model for Affect Sensing from Text. In: Affective Information Processing, pp. 45–73. Springer, London (2009)
11. Choujaa, D., Dulay, N.: Using screenplays as a source of context data. In: Proceeding of the 2nd ACM International Workshop on Story Representation, Mechanism and Context. SRMC 2008, pp. 13–20. ACM, New York (2008)
12. Cassell, J., Vilhjálmsson, H.H., Bickmore, T.: BEAT: the Behavior Expression Animation Toolkit. In: Proceedings of the 28th Annual Conference on Computer Graphics and interactive Techniques SIGGRAPH 2001, pp. 477–486. ACM, New York (2001)
13. Breitfuss, W., Prendinger, H., Ishizuka, M.: Automated generation of non-verbal behavior for virtual embodied characters. In: Proceedings of the 9th International Conference on Multimodal Interfaces. ICMI 2007, pp. 319–322. ACM, New York (2007)
14. Prendinger, H., Ishizuka, M.: SCREAM: scripting emotion-based agent minds. In: Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems: Part 1. AAMAS 2002, pp. 350–351. ACM, New York (2002)
15. Gebhard, P.: ALMA - Layered Model of Affect. In: Proceedings of the 4th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2005), Utrecht University, Netherlands, pp. 29–36. ACM, New York (2005)
16. De Raad, B.: The Big Five Personality Factors. In: The Psycholexical Approach to Personality. Hogrefe & Huber (2000)
17. Kopp, S., Allwood, J., Grammer, K., Ahlsen, E., Stocksmeier, T.: Modeling Embodied Feedback with Virtual Humans. In: Wachsmuth, I., Knoblich, G. (eds.) ZiF Research Group International Workshop. LNCS (LNAI), vol. 4930, pp. 18–37. Springer, Heidelberg (2008)
18. Pelachaud, C.: Multimodal expressive embodied conversational agents. In: Proceedings of the 13th Annual ACM International Conference on Multimedia. MULTIMEDIA 2005, pp. 683–689. ACM, New York (2005)

19. Gunes, H., Piccardi, M.: A Bimodal Face and Body Gesture Database for Automatic Analysis of Human Nonverbal Affective Behavior. In: 18th International Conference on Pattern Recognition, ICPR, vol. 1, pp. 1148–1153. IEEE Computer Society, Washington (2006)
20. Kelleher, J., Doris, T., Hussain, Q., Ó Nualláin, S.: SONAS: Multimodal, Multi-User Interaction with a Modelled Environment. In: Spatial Cognition, pp. 171–184. John Benjamins Publishing Company, Amsterdam (2000)
21. Coyne, B., Sproat, R.: WordsEye: an automatic text-to-scene conversion system. In: Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, pp. 487–496. ACM Press, Los Angeles (2001)
22. Kennedy, K., Mercer, R.E.: Planning animation cinematography and shot structure to communicate theme and mood. In: Proceedings of the 2nd International Symposium on Smart Graphics. SMARTGRAPH 2002, vol. 24, pp. 1–8. ACM, New York (2002)
23. De Melo, C., Paiva, A.: Multimodal Expression in Virtual Humans. Computer Animation and Virtual Worlds 2006 17(3-4), 239–348 (2006)
24. Liu, Z., Leung, K.: Script visualization (ScriptViz): a smart system that makes writing fun. Soft Computing 10(1), 34–40 (2006)
25. Rasheed, Z., Sheikh, Y., Shah, M.: On the use of computable features for film classification. IEEE Transactions on Circuits and Systems for Video Technology 15(1), 52–64 (2005)
26. Shim, H., Kang, B.G.: CAMEO - camera, audio and motion with emotion orchestration for immersive cinematography. In: Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology. ACE 2008, vol. 352, pp. 115–118. ACM, New York (2008)
27. Kuo, F., Chiang, M., Shan, M., Lee, S.: Emotion-based music recommendation by association discovery from film music. In: Proceedings of the 13th Annual ACM International Conference on Multimedia, MULTIMEDIA 2005, pp. 507–510. ACM, New York (2005)
28. Wahlster, W.: Smartkom: Foundations of Multimodal Dialogue Systems. Springer, Heidelberg (2006)
29. ScriptRight, `http://www.scriptright.com`
30. Virtual Theatre Interface, `http://www.accad.ohio-state.edu/VT`
31. Gutiérrez, M., Vexo, F., Thalmann, D.: The Mobile Animator: Interactive Character Animation in Collaborative Virtual Environments. In: Proceedings of the IEEE Virtual Reality 2004, VR 2004, pp. 125–132. IEEE Computer Society, Washington (2004)
32. Connexor, `http://www.connexor.eu/technology/machinese`
33. Tesniere, L.: Elements de syntaxe structurale. Klincksieck, Paris (1959)
34. Fellbaum, C.: WordNet: An Electronic Lexical Database. MIT Press, Cambridge (1998)
35. Lexical Conceptual Structure Database,
`http://www.umiacs.umd.edu/~bonnie/LCS_Database_Documentation.html`
36. Strapparava, C., Valitutti, A.: WordNet-Affect: an affective extension of WordNet. In: Proceedings of the 4th International Conference on Language Resources and Evaluation, LREC 2004, pp. 1083–1086 (April 2004)
37. Humanoid Animation Working Group, `http://www.h-anim.org`
38. FreeTTS 1.2 - A speech synthesizer written entirely in the JavaTM programming language, `http://freetts.sourceforge.net/docs/index.php`

Interaction Technique Combining Gripping and Pen Pressures

Yu Suzuki, Kazuo Misue, and Jiro Tanaka

Department of Computer Science, University of Tsukuba,
1-1-1 Tennodai, Tsukuba, Ibaraki, 305-8573 Japan
{suzuki,misue,jiro}@iplab.cs.tsukuba.ac.jp

Abstract. We propose an interaction technique called "gripping" to improve the operation of input with pen-based interfaces. This operation involves a strong grip when holding a pen. It can provide new input operations while maintaining the usability of the pen because users can grip it without having to lift their fingers off it. There is an interaction technique called pen pressure that users can use without having to lift their fingers off the pen. We introduce a novel interaction technique that combines gripping and pen pressures. This combination enables users to simultaneously input new two values in addition to common pen interactions like those with tapping or stroking. By applying this technique to a paint tool, for instance, novel drawing operations become possible and artists can create new artworks. Two experiments were conducted to investigate whether users could perform an operation that combined gripping and pen pressures. As a result, they confirmed that combined interaction was possible and there was an optimal range of pressures that enabled combined interaction. We also implemented an application software for this combined interaction.

1 Introduction

Pens are one of the most familiar tools for creative activity because almost everyone has used them from an early age. Therefore, a pen-based interface, which is a pen-shaped device, would be more suitable for creative activities than devices only used for computers (like mice and keyboards). However, current pen-based interfaces have limited operability due to their lack of input operations. Therefore, we developed "gripping" [7], which is an interaction technique that can increase the number of input operations without users having to lose the usability of the pen. Gripping is an operation that involves a firm grip when holding a pen. It can maintain the usability of the pen because it maintains the pen idiom, which simply means that the pen is being physically gripped.

There is an interaction technique called pen pressure that users can use without having to lift their fingers off the pen. This pen pressure is used in commercial software and is useful for our pen-based interface. We considered combining gripping and pen pressure. This combination provides a pen-based interface with multi-stream input operations. That is, users can carry out common pen interactions like those in tapping or stroking while inputing two values simultaneously.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 440–448, 2010.

Applying this technique to a paint tool would enable them to undertake novel drawing operations and would offer new perspectives for artwork.

However, it is not clear whether users can independently apply gripping and pen pressure, and combine both operations. We attempted to answer this question by conducting two experiments that concerned the relationship between gripping and pen pressure.

2 Related Work

There have been many researchers who have improved the operability of pen-based interfaces. Here, we review some studies that have dealt with multi-stream input operation by using these interfaces.

We previously developed [6] an interaction technique using hand motions in the air. We used three motions; rolling, shaking, and swinging. Bi et al. [1] also conducted research that used rolling interactions. These researchers succeeded in increasing the number of input operations while maintaining the usability of the pen. Miura et al. [2] developed an interaction technique in which the stylus was rotated and slid. Siio et al. [5] proposed an interaction technique using the metaphor of a paperweight. The status of the user's palm — whether it was touching the bottom of a PDA or not — determined how modes were switched. A tablet developed by Wacom could detect not only the coordinates of a pen but also the pressure and tilting applied to it. Some research[3], [4], [8] that has used pen pressure and tilting has been conducted.

These researchers aimed to increase the number of input operations of pen-based interfaces. This aspect is common to our research. We accomplished multi-stream input by using pen gripping, which is a crucial motion when using pen-based interfaces.

3 Combining Gripping and Pen Pressure

3.1 Basic Idea

Gripping [7] is an interaction technique, which we proposed for a pen-based interface, that utilizes essential motion when using a pen. Gripping is a simple and easy operation that involves a strong grip when a pen is held. Pen pressure, on the other hand, is a common interaction technique in pen-based interfaces. Humans can control pen pressure naturally when using a pen. For instance, they can control line strength by varying pen pressure. A common aspect of both interaction techniques is to apply them without having to lift onefs fingers off the pen. Therefore, we came up with the idea of combining gripping and pen pressure.

3.2 Advantages of Combining Gripping and Pen Pressures in Creative Activities

A pen-based interface has recently been used and found to be suitable for creative activities like painting. However, the only change in creative activities by using the pen-based interface is that the canvas has been changed from paper to an LCD screen. Because the functions of the pen-based interface are basically the same as those with a pen, the role of the pen does not essentially change. Thus, creative activities using the pen-based interface do not receive favors of using it.

By combining gripping and pen pressure, we can achieve multi-stream input and simultaneously control two parameters. In other words, pen-based interfaces provide new modality for inputs. For instance, this can be applied to line-drawing operations while simultaneously changing the widths and colors of lines. Therefore, this combined interaction technique would enable users to benefit from pen-based interfaces during their creative activities.

3.3 Implementation

We developed a pressure-sensitive (PS) stylus that could detect gripping pressure [7]. The PS stylus was equipped with three pressure sensors to detect three degrees of finger strength. We designed the PS stylus to detect gripping pressures from 30 to 500 g, which was the average for the three sensors. The PS stylus recognized gripping pressure in 1024 steps. An application could attain a value from 0 to 1023 (the higher the value, the greater the strength). The relation between the output value of the sensor and the actual pressure was not linear but logarithmic. Hence, we modified the logarithmic characteristics to be linear through software.

We used a tablet made by Wacom to detect pen pressure.

4 Experiment 1: Range of Gripping and Pen Pressures for Simultaneous Control

4.1 Purpose

The combination of gripping and pen pressures could provide a novel interaction style. However, humans cannot always simultaneously control gripping and pen pressure. For instance, although it is clear that we cannot exert strong pressure on the pen by using weak gripping pressure, we still need to find the maximum strength of pen pressure exerted by weak pen pressure. Thus, we did an experiment to investigate what range of gripping and pen pressures could be used to enable simultaneous control.

4.2 Participants and Apparatus

Six male volunteers whose ages ranged from 22–26 participated in the experiment. Five of them were right-handed and one of them was left-handed. All of them correctly grasped the pen with their two forefingers and thumb. We adjusted the positions of the sensors according to the length of their fingers and asked them to grip the PS stylus naturally. The participants could use the PS stylus without having to worry about the sensors because of this adjustment.

We used a pressure sensor, which was the same as the ones attached to the PS stylus, to measure the pen pressure. We measured the pen pressure by placing the sensor on a desk and getting the participants to press it with the tip of the PS stylus. We used a pen tablet to measure the pen pressure. However, it was difficult to analyze the relationship between gripping pressure and pen pressure because the sensing specifications of the pen tablet were not clear. We could match the sensing specifications by using a sensor that was the same as that attached to the PS stylus.

The participants in this experiment sat down on a chair and pressed the sensor on the desk by using the tip of the PS stylus. The experimental software was run on an Intel Core 2 Duo 3.16-GHz PC with Windows Vista.

4.3 Tasks and Measurements

We measured the maximum value of pen pressure exerted with minimum pen pressure. First, participants held the pen lightly. Measurement began when participants touched the tip of the pen onto the pressure sensor. They exerted slight gripping pressure and pressed the pen with maximum pressure. Then, they increased the pen pressure step by step until it reached the maximum value (1023). This experiment was used to find the strength of the range of gripping and pen pressures that humans could simultaneously control because gripping pressure intensifies depending on increasing pen pressure. Participants repeated this trial five times. We measured both the gripping and pen pressures in this experiment.

4.4 Results

Fig. 1 is a scatter diagram where all the measured values have been plotted. The X-axis represents the gripping pressure and the Y-axis represents the pen pressure. A–F represent the six participants.

The results from analyzing all the measured values revealed that the regression relationship between gripping and pen pressures was not linear but logarithmic ($R^2 = 0.754$). Participant F had a tendency to hold the pen with the strongest grip. The regression curve of participant F was expressed by $y = 168.4 \log x - 283.97$.

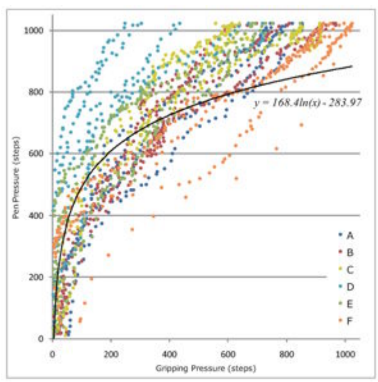


Fig. 1. Scatter diagram. X-axis represents gripping pressure and Y-axis represents pen pressure. A–F represent six participants.

4.5 Discussion

As participant F tended to hold the pen with the strongest grip, we decided to use the measured values to obtain the relation between gripping and pen pressures. When the gripping pressure was equivalent to x, the maximum value of pen pressure y could stand for $y = 168.4 \log x - 283.97$. Therefore, we found that it was possible for humans to simultaneously control gripping and pen pressures if the pressure range was $y >= 168.4 \log x - 283.97$.

We also found that there was a correlationship between gripping and pen pressures, which indicates that gripping can be an alternative to pen pressure. For instance, users can carry out continuous inputs like those with pen pressure when using devices that cannot detect pen pressure.

5 Experiment 2: Simultaneously Independent Control of Gripping and Pen Pressures

5.1 Purpose

We observed the relationship between gripping and pen pressures from the previous experiment and we could specify the pressure range that could be obtained. However, it was not clear whether humans could easily perform both operations simultaneously throughout all the available pressure range. We investigated what range could easily attain simultaneous control of gripping and pen pressures in this experiment.

5.2 Participants and Apparatus

The participants were nine volunteers (seven males and two females), who were 22–26 years old. Eight of them were right-handed and one of them was left-handed.

We used a pressure sensor that was the same as that attached to the PS stylus in Experiment 1 to measure the pen pressure. We measured the pen pressure by placing the sensor on a desk and getting participants to press it with the tip of the PS stylus.

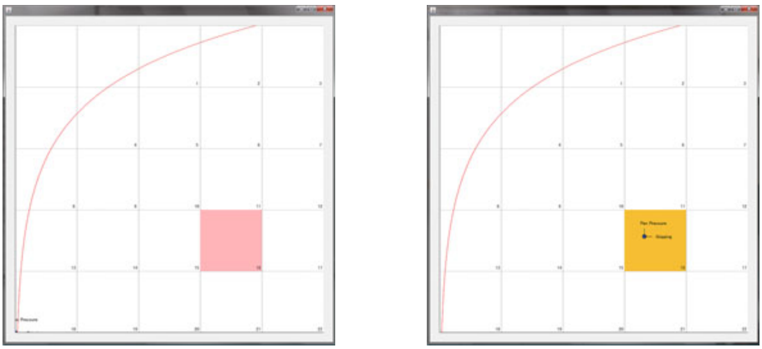
The participants in this experiment sat down on a chair and pressed the sensor on the desk by using the tip of the PS stylus. We used a 20-inch LCD with a resolution of 1280$\times$1024 pixels. The experimental software was run on an Intel Core 2 Quad 2.83-GHz PC with Windows Vista.

5.3 Tasks

One task was where participants simultaneously controlled gripping and pen pressures and then selected a target.

A pressure map (Fig. 2) was presented to the participants, which represented the pressure range. The X-axis represented the gripping pressure and the Y-axis represented the pen pressure. A blue cursor represented the pressure exerted by a participant. The point of origin of the pressure map was the bottom left corner. The cursor was located on the point of origin when both the gripping and pen pressures were 0 (Fig. 2(a)). The pressure map was 800 $\times$ 800 pixels. There were 1024 steps of output for gripping and pen pressures that were mapped uniformly to 800 pixels. There were 25 rectangles because the horizontal and vertical lines on the pressure map were divided into five. That is, the pressure range was divided into 25 partial ranges. These rectangles were targets that the participants selected. The target rectangles were pink. When the cursor was inside the target, the target changed to orange. A red curve was plotted on the pressure map in experiment 2, which represented the relation between gripping and pen pressures that humans can simultaneously control.

We excluded three rectangles because rectangles outside the curve were difficult to select. Over half the area of the excluded rectangles were outside the curve. We assigned ID numbers from 1–22 to the rest of the rectangles in the left top corner. These IDs were used when displaying the results.



(a) Screenshot when cursor is located on point of origin.

(b) Screenhost when cursor is inside target.

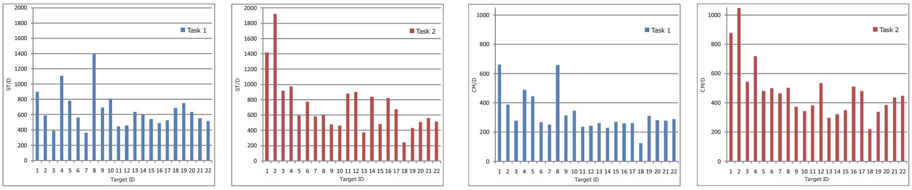
Fig. 2. Screenshot of pressure map. Cursor is blue circle and target is pink rectangle. Target changes to orange when cursor is inside target.

There were two kinds of tasks. The main difference was in the position of the cursor when the trial began. The initial position of the cursor in task 1 was the point of origin and the initial position in task 2 was the top right corner, i.e., where gripping and pen pressures had maximum values. When gripping and pen pressures were applied simultaneously, the initial strength of these two pressures was not always constant. Therefore, we prepared the two tasks to simulate the effects of the difference in initial pressures.

The nine participants started the experiment by moving the cursor to the initial point. Then, a target was presented and they moved the cursor to the target by exerting gripping and pen pressures. The target was selected by using the *Keeping* operation inside the target. Each participant selected the target 22 times to match the number of targets. The order in which targets were presented was balanced using a Latin square. Participants repeated this target selection three times. In summary, 1188 correct target selections (9 participants $\times$ 22 partial pressure ranges $\times$ 3 repetitions $\times$ 2 tasks) were performed in this experiment.

5.4 Performance Measurements

We evaluated performance from two points of view, i.e., target selection time (ST) and cursor movement distance (CM). We used ST as an index of rapidity



(a) ST (task 1) (b) ST (task 2) (c) CM (task 1) (d) CM(task 2)

Fig. 3. Results for target selection time (ST) and cursor movement distance (CM)

and CM as an index of ease. We could not use ST and CM without modifications because the distance between the initial cursor position and target varied from the target. Therefore, we defined the distance, D, between the cursor position and the target. Then, we minimized the effects of distance by dividing ST and CM by D. We adopted a Euclidean distance in task 1 from the point of origin to the bottom left corner of the target as the distance. We adopted a Euclidean distance in task 2 from the top right corner of the pressure map to the top right corner of the target as the distance. Then, we defined distance D as the Euclidean distance when the length of the target rectangle was one, plus one. For instance, the distance for target 14 in task 1 was $D = 1 + \sqrt{2}$ and that for target 20 in task 1 was $D = 3$.

5.5 Results

We removed 61 measurements that were beyond two standard deviations from the mean value.

Figs. 3(a) and 3(b) show the results for ST, which indicate that the participants took a lot of time to select targets 1, 4, and 8 in task 1. They also took a lot of time to select targets 1 and 2 in task 2.

Figs. 3(c) and 3(d) show the results for CM, which indicate that participants took a lot of time to select targets 1 and 8 in task 1. They also took a lot of time to select targets 1, 2, and 4 in task 2.

5.6 Discussion

Our analyses of ST and CM revealed that it was difficult to select targets 1, 2, 4, and 8. Fig. 4 shows these four target positions on the pressure map, where these four targets are located along the curve that we obtained from experiment 1. This indicates that areas along the curve are difficult to select. We only examined two initial cursor positions in this experiment and areas that were difficult to select were biased. Therefore, we concluded that areas along the curve were difficult to select independently of the initial cursor position. That is, it was not optimal to use the pressure ranges of the areas for interaction. In contrast, participants could select other areas easily and quickly on some levels. Therefore, we concluded that it was possible for them to use these areas for interaction.

We set the selection equation at $y = 0.667x+1.333$ to divide the pressure map into areas that were easy and difficult to select. This equation was an expression connecting the bottom right corners of targets 2 and 8. The gripping and pen pressures were equivalent to x and y, and areas that were easy to select could be defined as $y < 0.667x + 1.333$. We believe that this definition would work well as a guide in design applications. Multi-stream input operations that combine gripping with pen pressure can be accomplished by using this pressure range optimally.

It is clear that gripping and pen pressures are in a proportional relation in our setup because we used the same sensors for measuring both pressures. Hence, every designer can apply our results when he or she uses other sensors whose specifications are known.

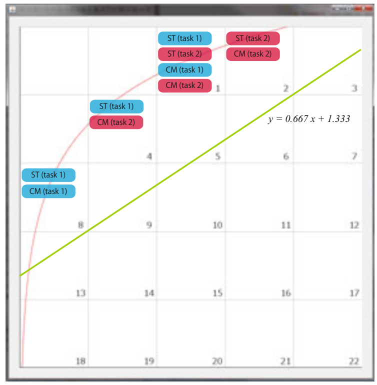


Fig. 4. Four difficult targets to select on pressure map. All four targets are along curve we obtained from experiment 1.

6 Applications

We found that gripping and pen pressures could be used simultaneously from the experiments. By combining gripping and pen pressures, a user can simultaneously change two parameters without interrupting operations. Here, we discuss applications using a combination of gripping and pen pressures.

Gripping Brush. Some paint tools support pen pressure, where a user can change the line width by varying pen pressure while drawing a line. The line width in in our paint tool could also be changed by varying pen pressure. The stronger the pen pressure became, the thicker the line became. In addition, the color of the line could be changed by gripping. The color saturation decreased when the pen was grasped strongly and increased when the pen was grasped weakly. Various kinds of lines like those in Fig. 5 could be drawn by combining gripping and pen pressures. The upper line in Fig. 5 was drawn while only controlling pen pressure, the middle line was drawn while only controlling gripping pressure, and the lower line was

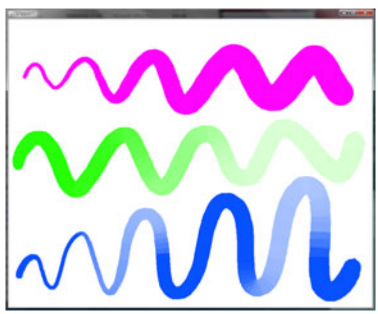


Fig. 5. Screenshot when using gripping and pen pressures. Upper, middle, and lower lines are drawn while only controlling pen pressure, gripping pressure, and both gripping and pen pressure, respectively.

drawn while controlling both gripping and pen pressures. The user could freely change pressures in all assignments like they did in controlling gradation or transparency.

Drawing operation while simultaneously changing two parameters have not been achieved in current pen-based interfaces. However, the combination of gripping and pen pressures should enable users to draw novel representations of lines. In addition, this combination should also enable users to create new drawing techniques because they can freely assign functions to the modality of input. This combination would be useful for creative activities like those in painting because it can help artists create novel artworks.

7 Conclusions

We proposed an interaction technique combining gripping and pen pressures, which enabled users to carry out common pen interactions while simultaneously inputting two parameters. That is, this combination provided users with four DOFs in input operation combined with two-dimensional input of X-Y coordinates. We conducted two experiments to investigate whether or not users could do operations with combined gripping and pen pressures. As a result, we found that combined interaction was feasible and there was an optimal pressure range.

References

1. Bi, X., Moscovich, T., Ramos, G., Balakrishnan, R., Hinckley, K.: An exploration of pen rolling for pen-based interaction. In: Proc. of UIST 2008, pp. 191–200 (2008)
2. Miura, M., Kunifuji, S.: RodDirect: Two-Dimensional Input with Stylus Knob. In: Proc. of MobileHCI 2006, pp. 113–120 (2006)
3. Ramos, G., Boulos, M., Balakrishnan, R.: Pressure Widgets. In: Proc. of CHI 2004, pp. 487–494 (2004)
4. Ren, X., Yin, J., Zhao, S., Li, Y.: The adaptive hybrid cursor: A pressure-based target selection technique for pen-based user interfaces. In: Baranauskas, C., Palanque, P., Abascal, J., Barbosa, S.D.J. (eds.) INTERACT 2007. LNCS, vol. 4662, pp. 310–323. Springer, Heidelberg (2007)
5. Siio, I., Tsujita, H.: Mobile Interaction Using Paperweight Metaphor. In: Proc. of UIST 2006, pp. 111–114 (2006)
6. Suzuki, Y., Misue, K., Tanaka, J.: Pen-based interface using hand motions in the air. The IEICE Transactions on Information and Systems E91-D, 2647–2654 (2008)
7. Suzuki, Y., Misue, K., Tanaka, J.: Interaction technique for a pen-based interface using finger motions. In: Jacko, J.A. (ed.) HCII 2009. LNCS, vol. 5611, pp. 503–512. Springer, Heidelberg (2009)
8. Tian, F., Xu, L., Wang, H., Zhang, X., Liu, Y., Setlur, V., Dai, G.: Tilt menu: Using the 3d orientation information of pen devices to extend the selection capability of pen-based user interfaces. In: Proc. of CHI 2008, pp. 1371–1380 (2008)

Destination Board System Based on Photographs

Toyohisa Nakada

Niigata University of International and Information Studies
nakada@nuis.ac.jp

Abstract. A destination board is a board that is often placed on the door of a room to notify visitors of the user's current location. The system proposed in this paper is as follows: A user who wants to notify other people of his/her location takes a photograph of his/her current location. The photograph is automatically transmitted to the computer that is mounted on the door of his/her room. Visitors are shown the photograph and recognize the location shown in the photograph. Visitors can also leave some graffiti on the photo. I constructed a prototype system and performed the initial experiment. The results show that visitors often recognize the user's current location from the photograph and sometimes leave some graffiti on the image.

Keywords: destination board, photograph, location, graffiti.

1 Introduction

There are many ways of notifying other people of one's location. For example, a magnet on a paper that shows several places is often placed on one's room door. A visitor watches the position of the magnet on the paper and infers the user's current location. When the magnet points to the word "lunch," the user is probably in the cafeteria. Such a legacy system is useful; however, it is not flexible. For example, it is not possible to remotely update the user's current location. Moreover, the user has to set the magnet's position before going to some place. Systems that use a mobile device to remotely update the user's current location have already been proposed. The problem of such a system is usability. If the user selects the current location from among some registered places, the initial setting in which he/she registers several places is required. On the other hand, if the user writes his/her current location as a string, it is difficult to input the string in such a mobile device.

In this paper, I propose a destination board system that is easy to use and makes it possible to remotely update the user's location. An overview of the system is shown in Fig. 1. This system consists of a camera device with wireless communication and a touch-screen computer. A user who wants to notify visitors of his/her location takes a photograph of his/her current location; the photograph is immediately transmitted to the computer screen that is mounted on the user's front door. Visitors can know the user's current location from

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 449–456, 2010.

Fig. 1. Overview of the destination board system based on photographs

the photograph. If a visitor has information about the environment of the user, he/she will be able to identify the user's current location. The advantage of this system is that only one operation is required to notify other people of the user's current location. Recently, digital cameras with wireless connection have been distributed. An SD-card-type Wi-Fi connection (Eye-Fi card) is also useful in order to add network access to digital cameras. iPhone has a camera device and wireless connection. These products enable us to easily upload a photograph by pressing the shutter button. The device enables us to create the system; when a user takes a photograph, the photograph is shown on the screen mounted on the user's front door. This is advantageous for a user who wants to notify other people of his/her current location. On the other hand, the advantage for visitors is that they can leave some graffiti on the photograph. When a visitor realizes that the person that he/she is looking for is not around, he/she can leave a meaningful message and not just graffiti. Graffiti is also a great motivator to a person who wants to notify other people of his/her current location because he/she recognizes the existence of someone who saw the captured photograph.

In this paper, I will describe the advantages and the problems of a destination board in section 2. The proposed system architecture is discussed in section 3. Details regarding the initial experiment are presented in section 4. Related works and conclusion are described in sections 5 and 6, respectively.

2 Destination Board

In general, there are two users of a destination board system; a user who wants to notify other people of his/her current location and a visitor who is looking for the user. The advantages of the proposed system with respect to the former are follows:

1. Because in order to publish his/her own current location, the user is required to only press the shutter button, it is absolutely easy for him/her to use the system.
2. It is not needed to prepare before using the system, places where the user often goes to are not registered on the system, because all places can be taken by the photograph.
3. The system is adapted to a situation in which the user wants to update his/her current location after going out.
4. Because the user can know whether visitors came to the door by looking at the graffiti that they have left, he/she is motivated to use the system.

In contrast, the advantage of the proposed system with respect to the visitors and the following points are not yet clear.

Issue 1. Is it possible to understand where a user is by looking at a photograph?
Issue 2. Do visitors leave some graffiti?

The purpose of the paper is to examine the abovementioned two issues by using a prototype of the destination board system.

3 Prototype

3.1 Hardware Configuration

Fig. 2 shows the hardware configuration of the prototype. The system consists of an iPhone, a small touch-screen computer, and a Wi-Fi access point.

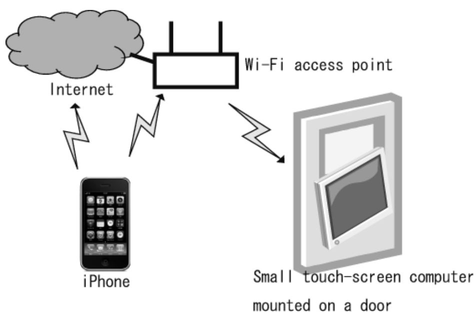


Fig. 2. Hardware configuration of the prototype

3.2 Photographs are Captured Using iPhone

I used an Eye-Fi card to take photographs in the first prototype; however, I could not upload the captured photograph outside of the wireless network. Therefore, I selected the iPhone for the second prototype. Because it has a wireless and a mobile network, I can upload the captured photograph from almost any place. The application that takes a photograph and immediately sends an e-mail with the photograph to a computer is installed onto the iPhone. A small computer in front of the door receives the e-mail and displays the photograph. The operation is very simple; click the application's icon, press the shutter button, and press the "use" button to determine whether the captured photograph is selected.

3.3 Displaying Transmitted Photograph and Leaving Graffiti

An overview of the prototype is shown in Fig. 3. The transmitted photograph is displayed on a small computer mounted on the user's front door. The computer has a touch screen. Visitors can leave graffiti on the displayed photograph by using a stylus. The software used for receiving the transmitted photograph, displaying the photograph, and drawing lines on the photograph by moving the stylus is created in Java.

A visitor can undo the left graffiti in a short time. Because it is possible that graffiti is drawn by many people, a visitor can erase graffiti drawn by others by using the undo function. Therefore, graffiti that is drawn before 5 min is not cleared in this prototype. The drawn graffiti is saved to a local disk when a new photograph is transmitted, and the graffiti is cleared.

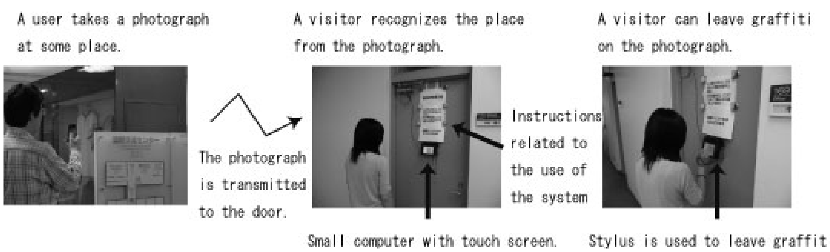


Fig. 3. Overview of the prototype system. Left figure shows a user taking a photograph of his current location.

4 Initial Experiment

4.1 Overview of the Experiment

The purpose of the initial experiment is to clarify issues and check how convenient it is for the visitors to use this system. The author of this paper who is a lecturer at a college is the user who wants to notify other people of his/her current location. His visitors are students, staff members, other faculty members of the college, and so on. The door on which the small computer is mounted is personal space's door. Therefore, visitors can understand that the displayed photograph is taken by the author. Visitors get to know of the system only when they see the small display screen on the front of the door; they are not given a prior explanation of the prototype. The only description on the door is "The photograph shows where the author's name currently is."

Fig. 4 shows the number of captured photographs and the number of photographs with graffiti on them. The X-axis and the Y-axis represent the date and the number of photographs, respectively. The white bar and the black bar indicate the number of photographs taken by the author and the number of photographs on which graffiti is drawn. For example, on June 29, eight photographs were taken, and graffiti was left on two photographs.

Nine people visited the author's room without appointment during the experiment period. These people were asked to take a questionnaire survey in June 2009. The purpose of this survey was to figure out whether the visitors could identify the location of the author when they visited him by watching the displayed photograph.

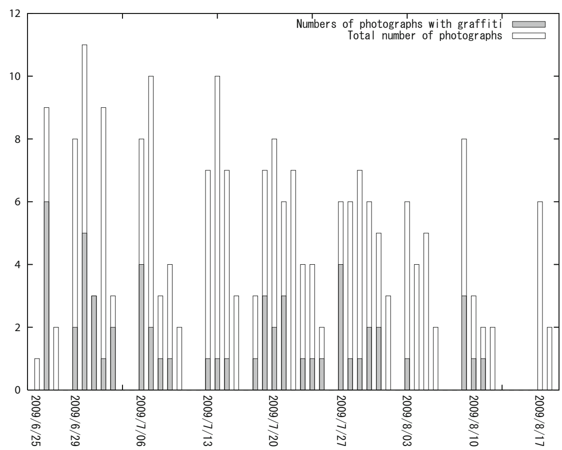


Fig. 4. Total number of photos and number of photos with graffiti on them in the initial experiment

4.2 Can Visitors Identify the Location of a Person by Looking at a Photograph?

Fig. 5 shows example photographs that I took and the recognition of my current location by visitors. I often took a photograph that gave an overview of the current place such as the picture on the top left top of Fig. 5; this picture was taken at an early stage of the experiment. Because the visitors informed me that they could not clearly infer my location from the photograph, I took more descriptive photographs such as the left middle and top right photographs shown in Fig. 5. The left middle photograph shows the identifier of the room (352). The room number is useful for identifying my location. The right middle photograph was taken before going out of the Wi-Fi network. A visitor was looking for me outside; however, he could not find me because I had already gone outside of the college.

The results of the questionnaire survey reveal that in the proposed system, the current locations often were unclear as compared to the legacy destination board system, which uses a magnet and a place displayed on paper.

Fig. 5. Sample photographs taken at several places and recognition of the place by visitors

4.3 Graffiti

Of the 214 photographs that were taken in the initial experiment, 58 photos had graffiti drawn on them. A user who wants to notify other people of his/her current location could watch some graffiti in 29 days per 41 working days in the experiment. The ratio was approximately 70%.

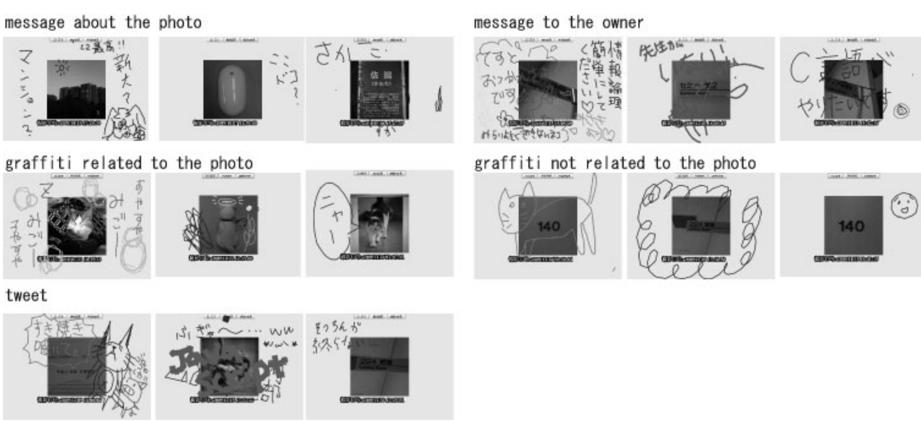


Fig. 6. Examples of graffiti

Fig. 6 shows examples of actual graffiti. I classified graffiti into the following five types: message about the photograph, message for the user who takes the photograph, graffiti related to the photograph, graffiti not related to the photograph, and tweet. Some visitors left comments about the photographs when they could not identify the places shown in the photographs.

5 Related Works

Many position detection systems have been proposed such as a combination of ultrasonic and RFID systems [1], an array of RFID put on environment [2], and a sensor fusion method [3]. Moreover, many applications based on the detected position have been proposed. Hightower et al. [4] have proposed an estimation method for the methods of transport from a sequence of detected positions. Sashima et al. [5] have proposed a supporting system for the communication between people in an academic meeting and those at an event site.

Some types of position detection systems that involve the use of photographs have already been proposed. A user can be navigated by using a photograph that shows a landmark and is sent from a navigator [6]. The navigated user sees the landmark in the photograph and looks for it. Positioning systems such as GPS are not required for the navigation; however, many photos that show the landmark must be prepared. In our system, no initial preparation is required. An automatic detection of position from a photo has also been proposed [7]. This method is based on supervised machine learning. Therefore, a considerable amount of training data, which are sets of photographs and the descriptions of the locations, is needed. Our system does not require such learning; however, the scalability of the proposed system is relatively small because the recognition of the user's current location depends upon the visitor's knowledge of the place. For example, if the captured picture may show a location that is familiar to a student from the same college but not to a student from a different college.

Graffiti on a picture is often used for supporting communications such as PhotoChat [8]. Communication support is also considered in the proposed system; however, at this stage of my study, I have focused on the system as a position detection system. Communication support for the proposed system is a future work.

6 Conclusion and Future Works

In this paper, I described a destination board system that uses photographs. I created a prototype system and performed the initial experiment. Some problems such as lack of recognition of the user's current location from the captured photograph were encountered. In the future, I want to enhance this system.

In the prototype, I sought to answer the following two questions: "Is it possible to identify the location of the user by looking at a photograph?" and "Do visitors leave some graffiti?" With respect to the first question, the visitors could not identify the user's location on the basis of some photographs. Further evaluation is required. With respect to the second question, a considerable amount of graffiti was left. In fact, I was motivated by this graffiti. However, other participants who upload a photograph are required for future works.

Evaluation from the viewpoint of communication support may be useful because the captured photograph and graffiti have some meaning. The evaluation and the development of a measure of communication support are future works.

References

1. Nakada, T., Kanai, H., Kunifuji, S.: A Support System for Finding Lost Objects using Spotlight. In: 7th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI 2005), pp. 321–322 (2005)
2. Liu, Y., Chen, L., Pei, J., Chen, Q., Zhao, Y.: Mining Frequent Trajectory Patterns for Activity Monitoring Using Radio Frequency Tag Arrays. In: Proceedings of the Fifth IEEE International Conference on Pervasive Computing and Communications (PerCom 2007), pp. 37–46 (2007)
3. Hightower, J., Brumitt, B., Borriello, G.: The Location Stack: A Layered Model for Location in Ubiquitous Computing. In: 4th IEEE Workshop on Mobile Computing Systems & Applications (WMCSA 2002), pp. 22–28 (2002)
4. Hightower, J., Consolvo, S., LaMarca, A., Smith, I., Hughes, J.: Learning and recognizing the places we go. In: Beigl, M., Intille, S.S., Rekimoto, J., Tokuda, H. (eds.) UbiComp 2005. LNCS, vol. 3660, pp. 159–176. Springer, Heidelberg (2005)
5. Sashima, A., Kurumatani, K., Izumi, N.: Location-mediated service coordination in ubiquitous computing. In: Proc. of the Workshop on Ontologies in Agent Systems, 2nd International Joint Conference on Autonomous Agents and Multi-Agent Systems, Deutsche Bibliothek (2003), http://CEUR-WS.org/Vol-73/
6. Beeharee, A.K., Steed, A.: A natural wayfinding exploiting photos in pedestrian navigation systems. In: Proceedings of the 8th Conference on Human-Computer interaction with Mobile Devices and Services (MobileHCI 2006), vol. 159 (2006)
7. Zheng, Y., Zhao, M., Song, Y., Adam, H., Buddemeier, U., Bissacco, A., Brucher, F., Chua, T., Neven, H.: Tour the World: building a web-scale landmark recognition engine. In: Proceedings of International Conference on Computer Vision and Pattern Recognition, CVPR 2009 (2009)
8. Sumi, Y., Ito, J., Nishida, T.: Photochat: communication support system based on sharing photos and notes. In: CHI 2008 Extended Abstracts, Works in progress, pp. 3237–3242 (2008)

Practicing on Stage: Increasing Transparency and Interaction in Class Activity with Digital Pen system

Taro Sugihara[1], Motoki Miura[2], and Susumu Kunifuji[1]

[1] Japan Advanced Institute of Science and Technology
{sugihara,kuni}@jaist.ac.jp
[2] Kyushu Institute of Technology
miuramo@mns.kyutech.ac.jp

Abstract. In this paper we described the concept of developing an interactive education system and understanding the actual activity inside the classroom. We developed an Anoto-based digital pen system and conducted a case study on three mathematics lectures in a high school, with participatory observation. After these practice lectures, focus group interviews were scheduled for the students in addition to a separate interview with a teacher; these interviews revealed the effects of our system. We concluded from the series of the investigations that the system was able to enhance interactive area of individual students, increase the transparency of class activities, raise students' motivation levels, and fortify the educational effects. Finally, a "stage" concept was generated from these results.

Keywords: Educational technology, case study, digital pen system, Anoto pen, class activity, and field research.

1 Introduction

Mobile computing and wireless networking technologies have been applied to facilitate computer-supported collaborative learning (CSCL) based on mobile learning or ubiquitous learning [1][2][3]. Most of the CSCL projects employ mobile/handheld devices such as tablet PCs, PDAs, and cellular phones as a main interface for students to track and collate their activities. These systems with pen tablets or note PCs have been applied not only in university and college courses such as StuPad [4], LiveNotes [5], Classroom feedback system [6] and Classroom Presenter [7], but also in the context of the learning activities of younger students in elementary schools [8], outdoor environments [2] and museums [3]. PDA-based systems have also been developed for similar reasons [9][10][11][12].

The interfaces and media should be carefully designed to assist learning, especially in the case of low-performing students. Students and teachers may not always be familiar with these systems, which may require specific training on the part of the users. Students' thinking process was disturbed by tasks they have to learn for system use. Oviatt, Arthur and Cohen indicated that students' performances of using tablet interface were worse than pen interface one [13].

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 457–464, 2010.

We have previously developed "AirTransNote (ATN)," a student note sharing system, to facilitate collaborative and interactive learning during regular lectures in a conventional classroom environment [14][15]. In our studies, we had chosen a digital pen that captured student notes written on regular paper, and a PDA [14] and Anoto-based system [15] that could transmit the notes immediately to a teacher's PC. The teacher could browse the notes, and show them to the students. Moreover, our ATN could provide feedback to students based on such notes. When using ATN, no computer literacy was required since the writing task was quite popular. However, our conventional systems did not implement the feature that ensured interactive response to students despite the fact that written data from digital pens were dynamically input to the system.

In these studies, however, relatively little attention was paid to the question of how the system affected actual class activity. It is quite important for real users to develop interactive systems, and to take into account how the interactive system should be. Christensen pointed out that systems that support learner-centered education will assume significance in the near future [16]. The role for the interactive system and its contribution to real educational fields will be enhanced considerably.

We carried out a case study to clarify the effects of an Anoto-based digital pen system on high school education. After applying our system thrice during a mathematics class, we observed students' behaviors. We also investigated the effects through two kinds of interviews—focus group interviews with the students and a conventional interview with a teacher. We also tried to build a concept for developing interactive systems and understanding the real dynamics of the classroom.

2 Anoto-Based Digital Pen System

2.1 Overview

In order to minimize the hurdles, our architecture adopts Anoto-based digital pens and a wireless data gathering system that is termed as the "digital pen gateway." An Anoto-based digital pen (Anoto pen) has the capability to store and send handwritten notes made on a specific dotted paper sheet. A camera embedded in the pen tip continuously captures images of the small dots on the sheets. The captured image is used to calculate the position of the pen tip. With the Anoto pen, the student is free of the setup and calibration processes. Further, the teacher can collect accurate and stable student notes. The Anoto pen can detect different paper sheets if a distinct dot pattern is printed and registered. Thus, the student can handle multiple sheets without any additional operations. As is apparent, setup procedures such as connecting physical cables are eliminated.

Anoto pen data is usually transmitted to a device such as a PC or a PDA through a USB cradle or a point-to-point Bluetooth connection. It is possible to use a student PDA to mediate data transmission, but the number of units of intermediate hardware should be minimized to reduce the chances of malfunction. Thus, we chose a digital pen gateway system solution developed by NTT Comware Corp. This gateway system is designed to receive pen data through Bluetooth and can handle about forty pens simultaneously. Our system's configuration is shown in Fig. 1.

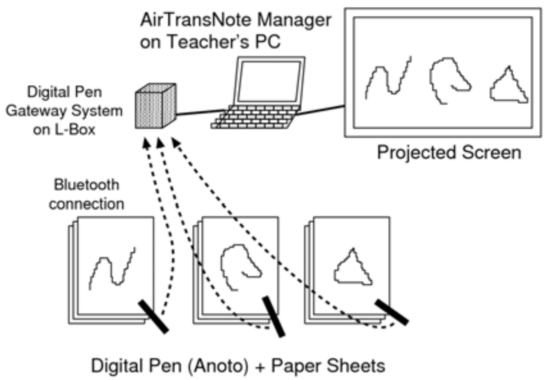


Fig. 1. Overview of the system

The Anoto pen only transmits data when a student explicitly taps a specific region of the paper that is registered as "Send." This limitation may cause some delay in notes updates, which is unfavorable to our initial objectives. However, we regard the merits of accuracy, stability of notes, and easy setup as more important than immediate and automatic updates. If the students can be made to be naturally aware of the sending status, they will not be disturbed because the sending operation in itself is quite simple.

2.2 Correct/Incorrect Feedback to the Students

It is quite difficult for teachers to provide constant feedback on their students' responses in the course of conventional lectures owing to the limitations of human resource. Students, in most cases, have to wait for the teacher's feedback, until he/she comes to their seat or until the designated time for explanations.

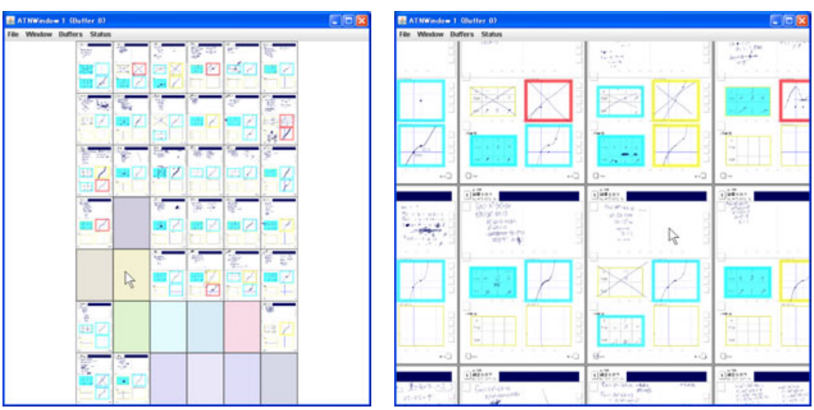


Fig. 2. View of all students' progress (left) and the zoomed view (right)

To solve this problem, we implemented a function to disseminate correct/incorrect feedback, and software for the recognition of on-line handwritten Japanese characters [17]. The system requires that teachers prepare response sheets, which have pre-defined calculation and answer areas, and that they register the correct answers onto the system. After this initial preparation, when students write down their answers, the system is equipped to generate instantaneous correct/incorrect judgments, as shown in Fig. 2. When the answer is correct, the answer area is painted over in blue. Red was applied to indicate incorrect answers. Students are thus saved a considerable amount of time, and they do not have to spend fruitless hours in wait for the teacher's verdict.

3 Methods

3.1 Interview Procedure and Practice in Math Classes

Participatory observation, focus groups interview, and conventional interview were c by a researcher who was not involved in the development of this system. The system was adopted in the course of three mathematics classes for high school sophomores; topic of these classes was "the maximum value and minimum value" of differential approach. Thirty-one students participated in the classes. They got down to their problem-solving exercises after the teacher had finished explaining the concepts and examples on the blackboard in the initial phase of the class. The system was used as the students responded to the set exercises.

After this practice, we announced that we would be paying students to participate in our subsequent activity; fifteen students offered to join in. We interviewed three focus groups for approximately one hour each. These three groups comprised five informants each. During the initial phases of these interviews, we informed the interviewees about their rights and received their informed consent. The interviews were fashioned in a semi-structured style. The interviewer asked the students to find differentiations between these practice lectures and ordinary one. We also conducted an interview with the teacher along the same standpoints. An IC recorder was used to tape both interviews.

3.2 Interview Procedure and Practice in Math Classes

Our analysis was focused on particulars that were related to the effects on classroom interaction and the influence of our system on the class activity. We made transcripts to analyze these interviews. Next, we classified the transcripts on the basis of Grounded Theory Approach, and named them for the sake of coding (Step1). We categorized the codes according to certain similarities, in a way that would eventually assist our conceptualization (Step2). Step1 and Step2 were repeated until all new concepts had been covered. Finally, concepts concerning the aspects of interaction and education were selected for analysis of the relationship between them.

4 Results

4.1 Interviews' Results: Increasing Transparency

One of the typical responses was related to the change of in the participatory attitude of the class owing to the increasing transparency of processes in the classroom. One reason for this, as was proffered by the students, was that competition among classmates was enhanced and made more enjoyable by the function of generating instantaneous judgments as to whether a submitted answer is correct or incorrect. Students could easily notice the progress that their classmates were making.

Moreover, students reported that they did not like their mistakes to be exposed in front of their peers and rivals. In the course of the ordinary math classes, rivals cannot detect fellow-students' errors or the degree of their progress, owing to the physical distance between students. At these practice lectures, however, all students were afforded the opportunity to check their rivals' progress and the number of right answers. In a conventional classroom activity, such transparency is not possible to achieve. Students who are convinced of the correctness of their answers usually go up to the blackboard and write their answers on it, while the rest of the class checks the answers; this leads to a one-sided presentation style which is the most traditional style in Japan. The other students' calculating processes and answers remain untested. The usual lecture style thus has the loophole of students losing opportunities to know of the mistakes they are making in their calculations; the conventional approach thus entails a risk that students may flounder in the face of significant challenges such as entrance examinations.

On the other hand, as per the responses of students, there were certain negative effects of the system as well. Although the system has a name withheld function, students were hesitant to write and send in their answers to the system. In fact, we observed that one student had sent his answers to the system after calculating on his notes, using them as drafts. Students, further, did not wish to get into situations wherein it would appear that they were not solving their exercises owing to the difficulty levels. They felt that the name withheld function did not work well, because students' curiosities were invariably driven to find out the author of any given answer.

4.2 Interviews' Result and an Episode: Increasing Interactive Behavior

Our system could enhance the interactive action in the classroom. When we were observing the class activity, a remarkable incident occurred. A student had finished his own task and was browsing the displayed processes and the answers of his classmates. He suddenly pointed out a mistake in the work of a classmate. The student who had made the error had not noticed it until then; the two students were far away, and there was no way they could have checked their answers directly had it not been for the facilities of the system.

They felt that the system enhanced students' activities. For example, a typically answer was that the possibility of pointing out the errors of classmates, which is not possible in an ordinary class because of distance and a lack of direction or focus. The other representative topic was to check their condition. The interviewees desired to confirm their classmate's progress whether their stage of work completion was superior or inferior to that of the others. They did their task, all the while keeping an eye on the display (e.g., after finishing tasks or sending incomplete calculations for the system to register). The teacher told us that someone pointing out a wrong answer was an unusual occurrence in an ordinary class.

4.3 Interviews' Results: Raising Motivation

Our system led to raising the students' motivation levels of participating in lectures. Almost all students said that the practice lectures with the system in operation were attractive and enjoyable, and that they were aroused by the interactive environment spawned by the system. When a group discussed this topic, the use of the system was regarded as being akin to the use of a new stationery product. Students were refreshed and motivated by it. In fact, some students who often slept during class woke up for these practice lectures and tried gamely to solve their tasks. The teacher pointed out this change in atmosphere. Students were very lively and joyful in the practice classes in contrast to the ordinary lectures. Consequently, our system may be said to engender positive effects in terms of student motivation, helping them engage in lectures.

5 Insight and Conclusion

In response to the question regarding what the biggest changes in the classroom under the system were, to the respondents picked increased interaction and transparency. The overview of an ordinary class is illustrated in figure 2. Students concentrate on the surface of the blackboard in conventional lectures. When someone is assigned to be an answerer, the selected student who is obviously sufficiently prepared (having worked at home, or at his/her desk) goes on to write his/her calculations as though partaking in a ritualized ceremony. No one comes to know of the typical mistakes that ought to be avoided, and the risks of making basic mistakes in significant examinations remains as acute as ever. Turning to theme of interaction among students, a conventional lecture stifles the interactive area and keeps it narrow; discussions or consultations are held only between neighbors.

We came up with a "stage" analogy from the results. That is, when students got down to exercises, they practiced hard in the auditorium or outside the theater until their play were perfect. After that, the hero/heroine who wished to be a star appeared on stage which, in truth, was in front of the blackboard, and hogged the spotlight. The other students were just looking at the play as audiences.

On the other hand, there was no auditorium in our practice lectures (Fig. 3). The fact that students could not hide their calculating processes meant that they had to practice on equal terms on the stage itself. Everyone became either a player or a member in the audience. The interactive area was enhanced until it took up the entire classroom. The tedium that had been such a feature of ordinary classes had changed into a joyous atmosphere as a result of the interaction.

The teacher who was a playwright and a stage director could figure out the player condition, take adequate action on students, or assign well-suited roles to the various players. He could show them typical patterns of wrong answers and induce them to take care of such mistakes.

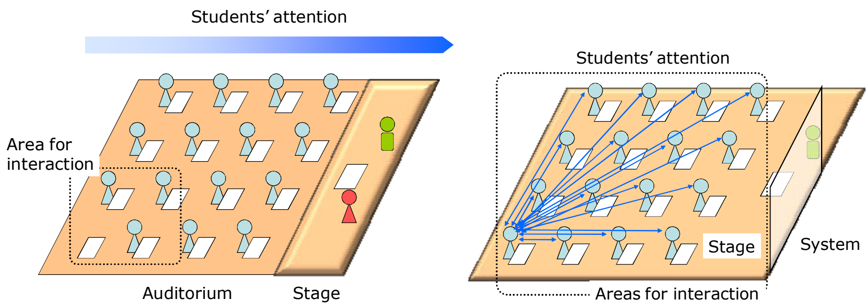


Fig. 3. Interaction and attention in ordinary styled classroom (left) and changes in interaction and attention with the system (right)

In order to develop educational technology, this concept plays an important role for both academically and practically. The standpoint may support the understanding of class activities in terms of an analogy of a "stage." If the various functions of educational technology are implemented along with the concept, it is possible to motivate students and to bring about a change in conventional styled classroom lectures. However, our attention has also been drawn to the novel concern of violations of privacy. This subject remains as a theme of serious work in future.

In conclusion, the implementation of the system led to an enhancement in the interactive area for individual students, increased the transparency of class activities, and raised student motivation levels. We proposed a "stage" concept to find out more on the influence of the system. This analogy was useful in understanding the class activity once the system had been adopted.

Acknowledgments. Our research is supported by Grant-in-Aid for Scientific Research (20680036) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

References

[1] Ogata, H., Yano, Y.: Supporting Knowledge Awareness for a Ubiquitous CSCL. In: Proc. eLearn 2003, pp. 2362–2369 (2003)
[2] Yang, J.C., Chen, C.H.: Design of Inquiry Learning Activity Using Wireless and Mobile Technologies. In: Proc. ICALT 2006, pp. 398–402 (2006)
[3] Yatani, K., Sugimoto, M., Kusunoki, F.: Musex: A System for Supporting Children's Collaborative Learning in a Museum with PDAs. In: Proc. WMTE 2004, pp. 109–113 (2004)
[4] Truong, K.N., Abowd, G.D.: StuPad: Integrating Student Notes with Class Lectures. In: Extended Abstracts of CHI 1999, 208–209 (1999)
[5] Kam, M., Wang, J., Iles, A., Tse, E., Chiu, J., Glaser, D., Tarshish, O., Canny, J.: Livenotes: A System for Cooperative and Augmented Note-Taking in Lectures. In: Proc. CHI 2005, pp. 531–540 (2005)

[6] Anderson, R.J., Anderson, R., Vandegrift, T., Wolfman, S., Yasuhara, K.: Promoting Interaction in Large Classes with Computer-Mediated Feedback. In: Proc. CSCL 2003 (2003)

[7] Anderson, R.J., Hoyer, C., Wolfman, S.A., Anderson, R.: A Study of Digital Ink in Lecture Presentation. In: Proc. CHI 2004, pp. 567–574 (2004)

[8] Iwayama, N., Akiyama, K., Tanaka, H., Tamura, H., Ishigaki, K.: Handwriting-Based Learning Materials on a Tablet PC: A Prototype and Its Practical Studies in an Elementary School. In: Proc. IWFHR 2004, pp. 533–538 (2004)

[9] Davis, R.C., Landay, J.A., Chen, V., Huang, J., Lee, R.B., Li, F., Lin, J., Charles III, B.M., Schleimer, B., Price, M.N., Schilit, B.N.: NotePals: Lightweight Note Sharing by the Group, for the Group. In: Proc. CHI 1999, pp. 338–345 (1999)

[10] Ratto, M., Shapiro, R.B., Truong, T.M., Griswold, W.G.: The ActiveClass Project: Experiments in Encouraging Classroom Participation. In: Proc. CSCL 2003 (2003)

[11] Yoshino, T., Munemori, J.: SEGODON: Learning Support System that can be Applied to Various Forms. In: Ghaoui, C. (ed.) E-Education Applications: Human Factors and Innovative Approaches, pp. 132–152. Information Science Publishing (2004)

[12] Singh, G., Denoue, L., Das, A.: Collaborative Note Taking. In: Proc. WMTE 2004, pp. 163–167 (2004)

[13] Oviatt, S., Arthur, A., Cohen, J.: Quiet Interfaces that Help Students Think. In: Proc. UIST 2006, pp. 191–200 (2006)

[14] Miura, M., Kunifuji, S., Shizuki, B., Tanaka, J.: Augmented Classroom: A Paper-Centric Approach for Collaborative Learning System. In: Murakami, H., Nakashima, H., Tokuda, H., Yasumura, M. (eds.) UCS 2004. LNCS, vol. 3598, pp. 104–116. Springer, Heidelberg (2005)

[15] Miura, M., Kunifuji, S., Sakamoto, Y.: Practical Environment for Realizing Augmented Classroom with Wireless Digital Pens. In: Apolloni, B., Howlett, R.J., Jain, L. (eds.) KES 2007, Part III. LNCS (LNAI), vol. 4694, pp. 777–785. Springer, Heidelberg (2007)

[16] Christensen, C.M., Johnson, C., Horn, M.: Disrupting Class: How Disruptive Innovation Will Change the Way the World Learns. Mcgraw-Hill, New York (2008)

[17] Nakagawa, M., Akiyama, K., Tu, L.V., Homma, A., Higashiyama, T.: Robust and Highly Customizable Recognition of On-Line Handwritten Japanese Characters. In: Proc. ICPR 1996, vol. 3 , pp. 269–273 (1996)

Extracting a Keyword Network of Flood Disaster Measures

Motoki Miura[1], Mitsuhiro Tokuda[2], and Daiki Kuwahara[2]

[1] Department of Basic Sciences,
[2] Department of Civil and Architectural Engineering,
Faculty of Engineering, Kyushu Institute of Technology
miuramo@mns.kyutech.ac.jp, tokuda@civil.kyutech.ac.jp

Abstract. For rapid and effective recovery from a flood disaster, an anti-disaster headquarters must not only assess the extent of damage, but also possess overall knowledge of the possible anti-disaster measures. However, in an anti-disaster headquarters, the personnel possess specialized knowledge, and the prospective measures and services helpful to people affected by disasters tend to go unnoticed. In order to provide overall knowledge of the possible anti-disaster measures, we analyzed the measures undertaken in previous flood disasters and visualized the structure of these measures using networked graphs. In this paper, we describe our analysis method and our implementation of a visualizing system. Our analysis helps the personnel to grasp the underlying knowledge of flood disasters by clarifying a few categories of the undertaken measures and the links between them.

Keywords: Visualization, Graph Drawing, Flood Disaster Measures.

1 Introduction

IPCC (Intergovernmental Panel on Climate Change)[1] and ICHARM (International Centre for Water Hazard and Risk Management)[2] have reported that because the frequency of extra-tropical cyclones has increased due to the greenhouse effect, the risks of flood disasters and localized heavy rains have increased. Thus, over the last few decades, the demand for techniques that prevent or reduce flood damage has increased worldwide. To prevent or reduce the damage caused by flooding, an anti-disaster headquarters must possess a overall knowledge of the possible anti-disaster measures. However, the anti-disaster headquarters' personnel have specialized knowledge, and prospective measures and services for people affected by disasters tend to go unnoticed.

In order to provide overall knowledge about the possible anti-disaster measures, we analyzed the measures undertaken during previous flood disasters, and

[1] http://www.ipcc.ch/
[2] http://www.icharm.pwri.go.jp/index.html

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 465–474, 2010.

Table 1. Source text sentences and keywords

ID	Source Text (Sentence)	Keywords
1	Do not hesitate to counsel refugees. Safety of human life is most important.	Life, Refuge, RefugeesCounsel
3	We observed that even in an emergency, people do not evacuate. They tend to avoid danger, and prefer a peaceful environment. When timing an evacuation announcement, counseling refugees is also important. However, it is more crucial to make people aware of the need for evacuation.	Refuge, RefugeesCounsel, Publication, TenseSituation
4	A volunteer center should be established immediately. A volunteer is not only a worker, but is also a hope for recovery. They are often a source of encouragement to the victims.	VolunteerCenter, Volunteer, Victims, Encouragement
8	A huge amount of garbage is thrown away. A temporary garbage dump should be constructed immediately. For quick processing, separation of garbage should be promoted.	Garbage, TemporaryPlace, GarbageSeparation
9	Do not hesitate to ask for relief even if it costs money. Any financial issue can be settled afterwards. The head of municipality must prove that we can afford the refugee costs.	Budget, ImmediatelyAfter, HeadOfMunicipality

visualized the structure of these measures using networked graphs. In this paper, we describe our analysis method and our implementation of a visualizing system. In order to construct overall knowledge on flood disaster measures, data of these measures should be collected not only from the city center staff, but also from victims of such disasters across various stricken sites. We collected real-life data, but to complement this information, we referred to a book titled "Know-hows of prevention, reduction, and recovering of flood disaster—message from the stricken areas" [1]. The book contains case studies and guidelines in the event of flood disasters. The contents are divided into three categories: usual, happening, and recovering. While the book structure was well designed for anti-disaster headquarters personnel, who need quick reference to such measures, we believe that the book could be restructured to a different point of view, for example, people and places. Basically, the content of the book is restricted by a "tree" structure (e.g., chapters and sections), which cannot adequately represent the complicated structure of flood disaster measures. Therefore, we analyzed the contents of the book and compiled an underlying structure as a network graph.

2 Related Work

Many researchers investigated and reported the cause analysis and measures of flood disasters [2,3,4]. Each study is effective for accumulating information on the prevention and reduction of damage, and under similar conditions, each local case can be adapted to other areas. Based on convergent thinking, our research aims to develop a analysis method using various flood disaster measures.

Table 2. Keyword extracted from source text (frequency, keyword)

76 CityCenterStaff	76 Publication	69 RelatedOrganization
59 Municipality	59 Residents	57 RefugePlace
56 Victims	52 Cooperation	50 Support
48 Distribute	47 Institution	47 Supply
44 Budget	41 ResidentsAssociation	40 Prefecture
38 AntiDisasterHeadquarters	38 Helpees	34 CommunicationMethod
33 Sharing	32 Earth&Sand	32 Garbage
29 Disposal	29 InfoGathering	27 Advance
27 DamageInvestigation	27 Occurrence	27 TemporaryHousing
25 Consultation	25 DisasterDrill	23 Health
23 Inundation	22 Foods	22 Nonofficial
22 Request	21 DamageLevel	21 PrivateDisasterPreventionOrganization
21 Storing	21 Volunteer	20 Consciousness
19 Acceptance	19 RadioBroadcasting	17 Disinfecting
17 Efficiency	17 RefugeesCounsel	17 Refuge
16 Disinfectant	16 Emergency	16 HeadOfMunicipality
16 Law&Institution	16 ListOfCitizen	16 MedicalStaff
16 Rivers	16 Road	16 TemporaryPlace
15 HazardMap	15 RiversWaterLevel	14 ElectricAppliance
14 Infrastructure	14 Nation	14 Preparation
14 StrickenArea	13 Administration	13 Explanation
13 GarbageSeparation	13 Reinforcement	13 SafetyReference
13 Window	12 CommunicationSystem	12 Donation
12 FireDefenceTeam	12 House	12 IndustrialWaste
12 MassMedia	12 Usually	11 Agreement
11 Certification	11 Driftwood	11 Manual
11 MovingMethod	11 Recycle	10 FinancialAssistance
10 LocalGovernment	10 Undertaking	10 Vehicles
10 VolunteerCenter	9 Confusion	9 DisasterRecord
9 Exempt	9 MaintainHygiene	9 PreventionAct
9 ReceivingMessage	9 Register	9 Standard
9 Toilet	9 WeatherInfo	8 GarbageAccumulationPlace
8 Investigation	8 Rainfall	8 TreatmentPlant
8 Webpage	7 AfterDisaster	7 Cleaning
7 Estimation	7 Knowledge	7 RedCrossSociety
7 RefugeGuidance	7 Transport	7 Unify
7 WaterSupply	6 Contract	6 Firehouse
6 ImmediatelyAfter	6 Life	6 MakingResource
6 MedicalAppliances	6 Police	6 PowerSource&Fuel
6 PreventionPlan	6 PublicOffice	6 SceneOfDisaster
6 SelfDefenceForce	5 Child	5 Encouragement
5 FloodControl	5 GiftOfMoney	5 Limitation
5 Sandbag	5 System	5 TenseSituation
5 Water	5 WelfareCouncil	4 Agriculture
4 ChildCare	4 CleaningTool	4 Communication
4 Dam	4 Decision	4 DisasterReliefAct
4 DrinkingWater	4 Flooding	4 Infection
4 OrderOfPriority	4 PersonData	4 SufferFromDisaster
4 WWW	4 Watching	3 ComplaintProcedure
3 Education	3 FireFighter	3 InfoExchange
3 InfoManage	3 NurserySchool	3 Patient
3 PressConference	3 Rescue	3 Safety
3 Siren	3 Telephone	2 Anxiety
2 BadSmell	2 HelpeesToilet	2 Inspection
2 RefugeGoods	2 Repair	2 Shortage
2 Store	2 SupplyMeals	2 TrafficRegulation
1 Arrange	1 Clinic	1 Community
1 DangerousObject	1 Fire	1 FirstAidStation
1 Hill	1 HousingLand	1 Measures
1 Minimum	1 Pet	1 Prevention
1 Privacy	1 Process	1 Property
1 Rain	1 SecondaryDamage	1 Sewer
1 SupplyMeal	1 Trust	

A geographic information system (GIS) [5] is often utilized for creating hazard maps and decision support systems. Data obtained from satellites [6] can also be effective in monitoring and mining the risks. Kobayashi et al. developed a tangible interface for the collaborative simulation of flood disasters [7]. Utilizing such solutions can help reduce the potential damage. We also plan to consider and support the decisions of anti-disaster headquarters, but with the intention of updating the headquarters of the overall structure of flood disaster measures, especially in terms of long-term recovery.

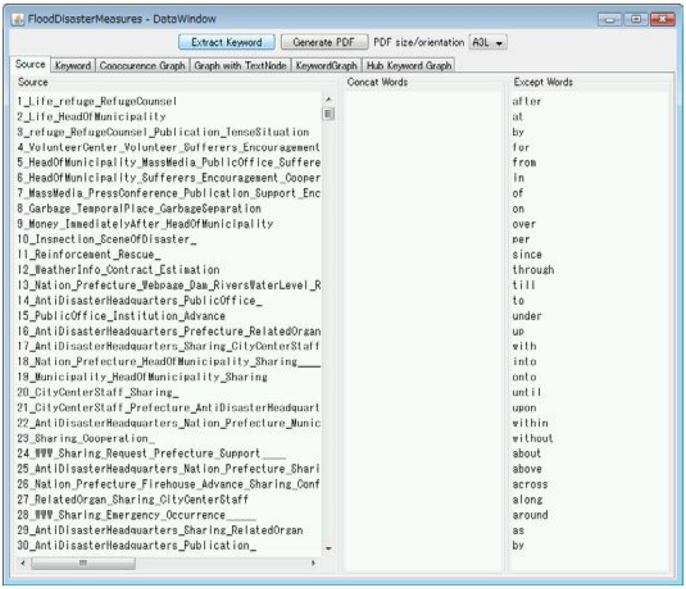


Fig. 1. Data window of flood disaster measures

Graph drawing techniques, such as the Bipartite Graph [8] and its three-dimensional representation [9], were conducted and visualization methods of social networking [10,11] were also proposed. This paper describes the adoption of such graph drawings and analysis techniques for mining flood disaster measures.

3 Procedure

To extract the underlying structure of the book, we created a network graph of the words and visualized it. The visualized network graph provided an alternative perspective of the target domain and helped in the acquisition of overall knowledge.

We chose 828 sentences that contain concrete cases and measures from the know-how book. In this paper, we refer to the text as *source text*. Because the source text involves unnecessary elements such as prepositions, brackets and punctuation, and also because the text is written in Japanese, we first separated the words in the sentences using MeCab[3], which is a fast Japanese morphological analyzer. We then selected idioms that should be concatenated.

The source text was suitable for retaining the original context of measures. However, we found that the original terms, especially, those which appeared in Japanese text, made the network clusters sparse because of their ambiguity and variety. Hence, we manually labeled several keywords to correlate the original

[3] http://mecab.sourceforge.net/

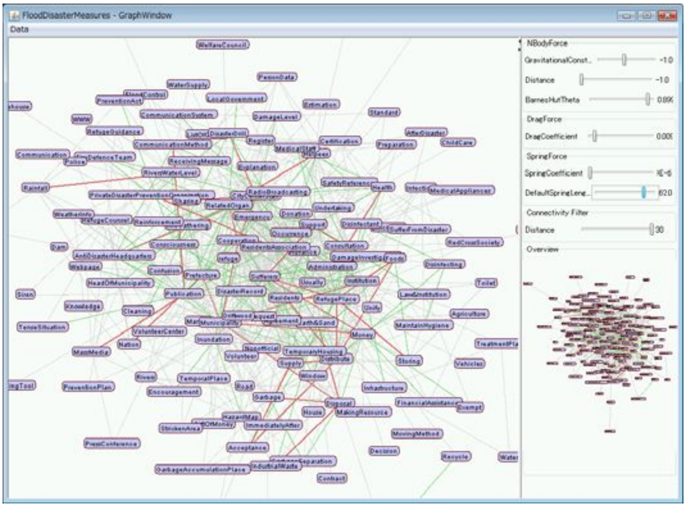


Fig. 2. Graph window of flood disaster measures

context to the source text. Table 1 lists some examples of the keywords corresponding to the original sentences. Note that the sentences in this example were translated by the authors and are not the original text in the know-how book. Labeling can be performed by introducing keyword mapping rules, which automate the labeling for further target texts. Table 2 lists the keywords and their frequency of use.

Next, we visualized the relationships of the know-hows for extracting knowledge about flood disaster measures. For that purpose, we developed a system for drawing a network graph. The Data window (Figure 1) consists of six tabs. The Source and Keyword tabs manage the source text and extracted keywords, respectively. The remaining four tabs correspond to a control panel for each method in the construction of the network.

The user first inserts source text into the Source tab. Each line of the source text contains the know-how ID and keywords regarding the know-how, separated by '_'. After clicking the Extract Keyword button, the user can construct the network and browse through the graph using the Graph window (Figure 2). The graph drawing function is powered by the prefuse toolkit [12]. We employed the Graph View demo to visualize the networks. Using the Graph window, the user can control the coefficients of the spring graph algorithm such as the spring force and default edge length. The connectivity filter allows the user to focus on a particular node and its surrounding local network by detaching the peripheral links.

3.1 Keyword Network with Know-How Text Node

The following section describes our experiences in extracting knowledge through the graph drawing system. First, we visualized the know-how keyword network

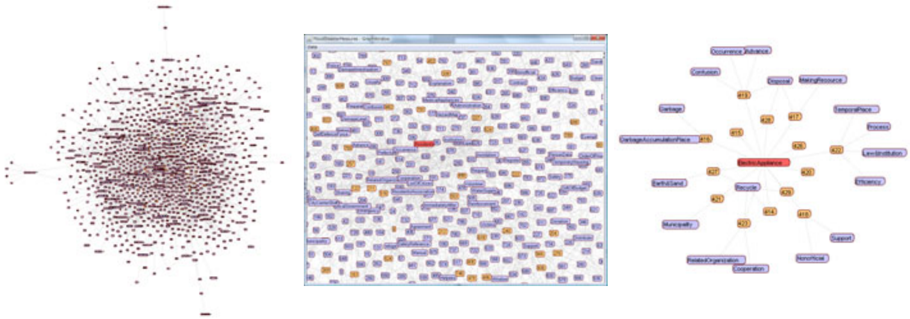


Fig. 3. Keyword network graph with know-how text node

in the most straightforward way: keywords and know-how IDs as nodes and relationships among them as links. Figure 3 shows a network built using this simple method. The figure on the left is the entire graph, which includes 2,539 links among 828 know-how nodes and 188 keyword nodes. The magnified figure is in the center. On clicking a node, the system highlights the connectivity. However, it was difficult to comprehend such a complicated structure. The figure on the right is the local view of the graph created by skipping over two hops from "Electric Appliance." From the graph, we can observe that the keyword "Recycle" shares five know-hows with "Electric Appliance." However, such data can also be observed by calculating the co-occurrence of keywords.

3.2 Keyword Network without Know-How Text Node

To focus on the structure of "keywords," we considered two methods. The first method was to eliminate the know-how ID nodes and to join all neighboring

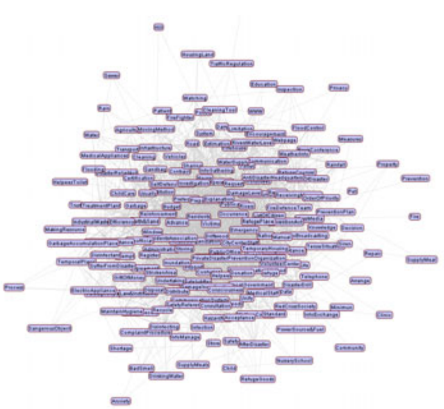


Fig. 4. Keyword network graph with all know-how text nodes eliminated

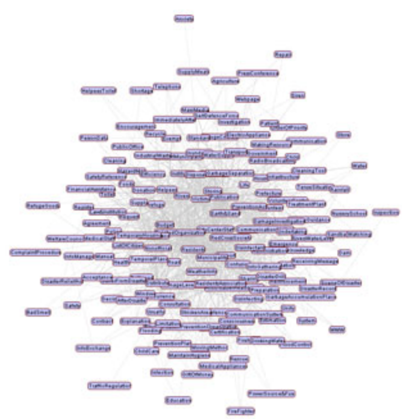


Fig. 5. Keyword network graph (with reduced links)

Table 3. Frequency of co-occurrence degree in the know-how text

Co-occurrence Degree	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16–23	24
Keyword Pairs	1374	366	148	58	32	32	13	10	3	0	1	1	2	1	2	0	1

keyword nodes with links. Figure 4 shows the resulting graph, which involves 3,450 links among 188 nodes. Because of the frequent occurrence of keywords, the network was too dense, and the structure could not be observed. The second method was similar to the first one, except, only the most frequent keyword in the know-how text could link to another keyword appearing in the same know-how text. Figure 5 shows the resulting graph, which consists of 1,691 links. Although this method reduces the links by half, it is still dense.

3.3 Co-occurrence Keyword Network

To reduce the links in a more appropriate manner, we considered a method that utilizes the co-occurrence of the keywords in the know-how text. The method first calculates the frequency of co-occurrence for all the keywords. Table 3 lists the number of keyword pairs, categorized according to their degree of co-occurrence. Figure 6 also plots the same data on a logarithmic y-axis. This method allows the user to estimate the number of links and highlight them by selecting the degree of co-occurrence. Figure 2 shows the resulting graph with colored links, which involves 670 links over one co-occurrence.

Figure 7 shows the graph of the keywords with 6–24 co-occurrence degrees. Table 4 lists keywords with 7–24 co-occurrence degrees; their links are shown in red. From the graph, we can observe some clusters that involve similar words. For example, the keywords “Disposal,” “Earth & Sand,” “Industrial Waste,” “Garbage,” and “Garbage Separation,” construct a cluster of “Garbage.” The “Publication” cluster relates to “Mass Media,” “Refugee Counsel,” “Radio Broadcasting,” and “Communication Method.” Similarly, “Victims,” “Distribute,”

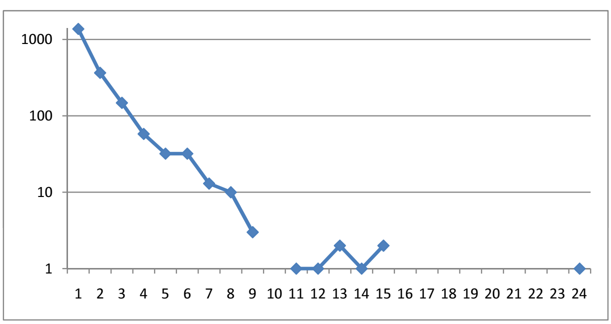


Fig. 6. Frequency of co-occurrence degree in the know-how text

Table 4. Keywords over six co-occurrence degrees

Co-occurrence	Keyword Pair
24	Cooperation–RelatedOrganization
15	Acceptance–Supply, Municipality–Prefecture
14	Distribute–Supply
13	AntiDisasterHeadquarters–CityCenterStaff, Consultation–Health
12	Distribute–Foods
11	Nation–Prefecture
9	CityCenterStaff–RelatedOrganization, Distribute–RefugePlace RefugePlace–Victims
8	CityCenterStaff–Cooperation, CommunicationMethod–Publication Cooperation–Municipality, Disposal–Garbage Disposal–IndustrialWaste, MassMedia–Publication Rainfall–RiversWaterLevel, RefugePlace–Residents RefugePlace–Storing, Victims–Support
7	Acceptance–Distribute, CityCenterStaff–DisasterDrill CityCenterStaff–Sharing, Consultation–Support Disposal–Earth&Sand, Disposal–Municipality Distribute–Residents, Foods–RefugePlace Helpees–RelatedOrganization, Institution–Municipality Budget–Municipality, Budget–Victims PrivateDisasterPreventionOrganization–ResidentsAssociation

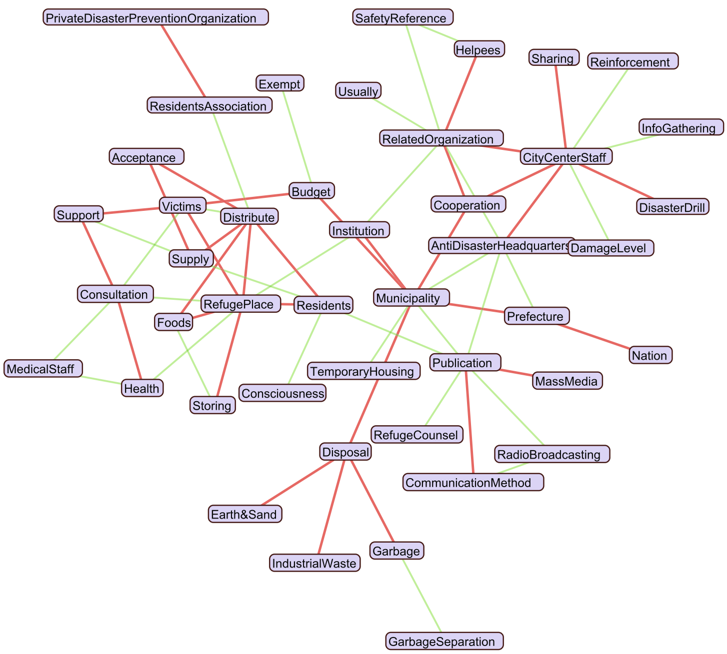


Fig. 7. Keyword network over five co-occurrence degrees

"Supply," "Refugee Place," "Foods," and "Storing" appear in the same cluster. We also believe that the other connections and distances in the graph are reasonable, where each connection has a logical reason. We observed that the data in Table 4 is also helpful, but the spatial layout of Figure 7 provided us with a good representation of the network structure of flood disaster measures.

4 Conclusion and Future Work

In this paper, we discussed our trials to analyze the network structure of flood disaster measures through the interactive visualization of graphs. Although the methods employed were simple and straightforward, we confirmed that these techniques can help highlight the meaningful data for flood disaster measures. In the future, we will improve the system to accept new measures posted by the anti-disaster headquarters' personnel and citizens, and dynamically construct a feasible network diagram. We hope that the overall knowledge provided by the system assists not only the anti-disaster headquarters but also citizens, and improve the quality of disaster-recovery services.

Acknowledgment

Our research is partly supported by a grant-in-aid for Scientific Research (20760414) and a fund from the Japan Institute of Construction Engineering.

References

1. Flood Disaster Summit Committee (ed.): Know-hows of prevention, reduction and recovering of flood disaster—message from the striken areas. Gyosei Corporation (June 2007) (in Japanese), ISBN 978-4-324-08224-9
2. Jonkman, S.N., Kelman, I.: An Analysis of the Causes and Circumstances of Flood Disaster Deaths. Disasters 29(1), 75–97 (2005)
3. Yin, H., Li, C.: Human impact on floods and flood disasters on the Yangtze River. Geomorphology 41(2-3), 105–109 (2001)
4. Imamura, F., Van To, D.: Flood and Typhoon Disasters in Viet Nam in the Half Century Since 1950. Natural Hazards 15(1), 71–87 (1997)
5. Chang, M.-S., Hsueh, C.-F.: Developing geographic information system for flood emergency logistics planning. In: SCSC: Proceedings of the 2007 Summer Computer Simulation Conference, pp. 652–659 (2007)
6. Brakenridge, G.R., Anderson, E., Nghiem, S.V., Caquard, S., Shabaneh, T.B.: Flood Warnings, Flood Disaster Assessments, and Flood Hazard Reduction: The Roles of Orbital Remote Sensing. In: Proceedings of the 30th International Symposium on Remote Sensing of Environment (November 2003)
7. Kobayashi, K., Narita, A., Hirano, M., Kase, I., Tsuchida, S., Omi, T., Kakizaki, T., Hosokawa, T.: Collaborative Simulation Interface for Planning Disaster Measures. In: CHI 2006 extended abstracts (Work-in-progress), pp. 977–982 (2006)
8. Misue, K.: Drawing Bipartite Graphs as Anchored Maps. In: Proceedings of Asia-Pacific Symposium on Information Visualization (APVIS 2006), pp. 169–177 (February 2006)
9. Ito, T., Misue, K., Tanaka, J.: Sphere Anchored Map: A Visualization Technique for Bipartite Graphs in 3D. In: Jacko, J.A. (ed.) HCII 2009, July 2009. LNCS, vol. 5611, pp. 811–820. Springer, Heidelberg (2009)

10. de Nooy, W., Mrvar, A., Batagelj, V. (eds.): Exploratory social network analysis with Pajek, vol. 27. Cambridge University Press, Cambridge (2006), ISBN 0-521-84173-9
11. Gao, J., Misue, K., Tanaka, J.: A Multiple-aspects Visualization Tool for Exploring Social Networks. In: Smith, M.J., Salvendy, G. (eds.) HCII 2009, July 2009. LNCS, vol. 5618, pp. 277–286. Springer, Heidelberg (2009)
12. The prefuse visualization toolkit, `http://prefuse.org/`

Study on Classification of the Tacit Knowledge Using an Eye-Tracking Interface: Experiment of Observation Pictures and Assessment of Reproduction Drawing

Yuta Watanabe[1], Issei Kodama[2], Hidehiko Hayashi[3], and Akinori Minaduki[4]

[1] Kushiro Public University, School of Business Management, 1, 1, 4, Ashino, Kushiro-shi, Hokkaido, Japan 085-8585
s072115@kushiro-pu.ac.jp
[2] Tohoku University, School of Engineering, 6, 6, Aoba, Aramaki, Aoba-ku, Sendai-shi, Miyagi, Japan, 980-8579
a8tb1086@cs.he.tohoku.ac.jp
[3] Naruto University of Education, 748, Nakashima,Takashima, Naruto-cho, Naruto-shi, Tokushima, Japan 772-8502
hhayashi@naruto-u.ac.jp
[4] Kushiro Public University, Center for Information and Technology and System, 1, 1, 4, Ashino, Kushiro-shi, Hokkaido, Japan 085-8585
minazuki@kushiro-pu.ac.jp

Abstract. When a person arrives at knowledge-creation by a sequence of knowledge conversion that who finds Knowledge from Information, Information from Data, a group may undergo knowledge-conversion similar to a person, too. On the other hand, knowledge-creation process may be complexed by the knowledge-creation, resulted from individual knowledge of group member, in the knowledge conversion by group. That is to say, it's important for practicing effectively group-knowledge-creation to disclose these knowledge conversion process and order measuring method for externalized knowledge. In this study, we propose the group formation method when we want to activate knowledge of group, and analyze it by the experiment.

Keywords: Knowledge-creation, Tacit Knowledge, Eye tracking, Group formation.

1 Introduction

In organization with knowledge works, creation of new knowledge or their activity always start from individual indiscernible knowledge rather than organizational discernible knowledge. One of their works for knowledge creation which is called the group-knowledge-creation converts individual knowledge into organizational one through dialogue, discussion and experience sharing in group [1],[2],[3],[4].

The group-knowledge-creation is needed for any phase, such as situation of a suggestion, a project and a practice in corporate activities and also important for any type

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 475–484, 2010.

of business, such as university and administration. In order to make the group formation for their success, we have developed a new method which extracts and estimates tacit knowledge [1],[2] in the knowledge conversion [3],[4] through a process of reproduction experiment and observation pictures. Our study based on focusing extraction and estimation of tacit knowledge has significant implications to discover effective group formation method in knowledge creation study. Our developed method is the feature which practices eye-tracking measurement and analysis for tacit knowledge in order to make the group formation in knowledge creation research.

The constitution of this paper is as follows. First of all, knowledge conversion process and the method to extract knowledge of group are given an outline in second paragraph. Next, basic notion of eye tracking and the first experimental demonstration are mentioned in third paragraph. In fourth paragraph, we put the second experimental demonstration into effect using developed method and consider group-knowledge-creation. Last fifth paragraph is summary of this paper.

2 Person's Tacit Knowledge which Acts Group-Knowledge-Creation

Knowledge has two notions on Knowledge management, one is "Tacit knowledge" and the other is "Explicit knowledge". Of these two notions, the former is individual knowledge about particular situation and it is difficult to formalize or transmit to others. On the other hand, the latter is explicit knowledge and can transmit to others with formal logical languages. Nonaka [1] says that the above-mentioned two notions of knowledge are not different but supplements or operates each other. Then, he advocated the dynamic knowledge creation model (SECI model), create and expand knowledge, through exchanging for tacit knowledge and explicit knowledge.

For organization to share a useful knowledge, it has to change knowledge universal through turning tacit knowledge into explicit knowledge. On this organizational knowledge creation process, the importance is said to create new vision through dialogue, discussion and dissension with various opinions is the factor which advances the conversion of individual knowledge into group knowledge. In other words, it is necessary for active knowledge creation to concentrate various opinions and sense of values, large experience (various tacit knowledge). There are two side of tacit knowledge, the cognitive side (schema, frame, perspective and so on) and technical side (specific know-how, skill and so on), and we pay attention to the cognitive side of tacit knowledge. In suppositions that concentrating various tacit knowledge activates organizational knowledge creation, we propose and consider the experimental method which extracts tacit knowledge and supposes its properties.

2.1 Definition of Data, Information and Knowledge

From the previous study [4], the definition of Data, Information, and Knowledge, treated by this text, is arranged. Data is true that is actually, and before it is defined. Data sublimes to Information when it catches receiver's interest and concern, and the definition is performed to each Data. And, a series of Information which is systematized in a specific purpose, and integrated becomes Knowledge (Figure 1).

It is thought that person is practicing the above-mentioned knowledge transformation under unconsciousness.

Data	Facts in actuality. Each data isn't related and It doesn't have the meaning. Data is the beginning for creating intellectual capital, such as information and knowledge.
Information	It's information to define data in some purpose. When a receiver's attention is engaged by a data, the data is interpreted information.
Knowledge	Knowledge is a series of information which is systematized and integrated in specific purpose. Person's knowledge creating activity gets only it, many services and products are developed.

Fig. 1. Linear model of Data, Information and Knowledge

2.2 Development of a Method to Extract Knowledge

For this study, we develop a new method to extract and measure the amount of Group-knowledge, externalized from individual knowledge in the group.

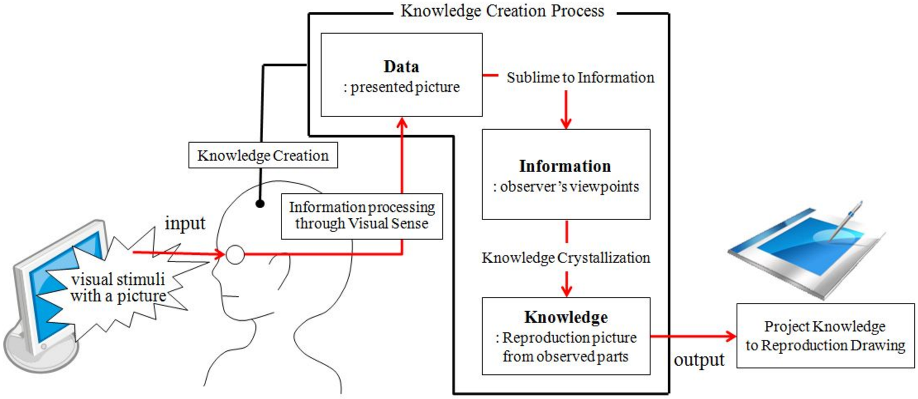


Fig. 2. Person's Information Processing and Experimental Process in this study

According to the process for showing in Figure 2, a person makes a sequence of knowledge conversion. Trough reproduction experiment with pictures, created knowledge is projected to the reproduced picture.

In this study, we extract the glance data in observation pictures to presume the cognitive side of tacit knowledge. It is thought that a person unconsciously makes a sequence of knowledge conversion (Information from Data, Knowledge from Information) and to watch freely is leading meaningful information (feature parts in the picture) for a person from a host of data (the whole picture). Eye-gaze points, that is to say, derived points as information are appeared a difference among individuals and reflected the cognitive side of tacit knowledge. From this, our method can embody the person's tacit knowledge elements by extraction the glance data in observation pictures.

3 Presumption of Tacit Knowledge by Eye-Tracking System

The tendency of how to receive a data, come originally from tacit knowledge of an individual, was decided by producing a difference of eye-gaze through observation landscape pictures. Ordinarily, it is clear that a person gazes into a high point of interest and cast it a glance [5]. Then, it is able to suppose that eyes contain a mentality, such as subject, and experience by effect of tacit knowledge. Therefore, if it were to extract the factor of mentality by effect of tacit knowledge from the tendency of eyes, it could judge the process of "The conversion of Data to Information" in an individual (Figure 3).

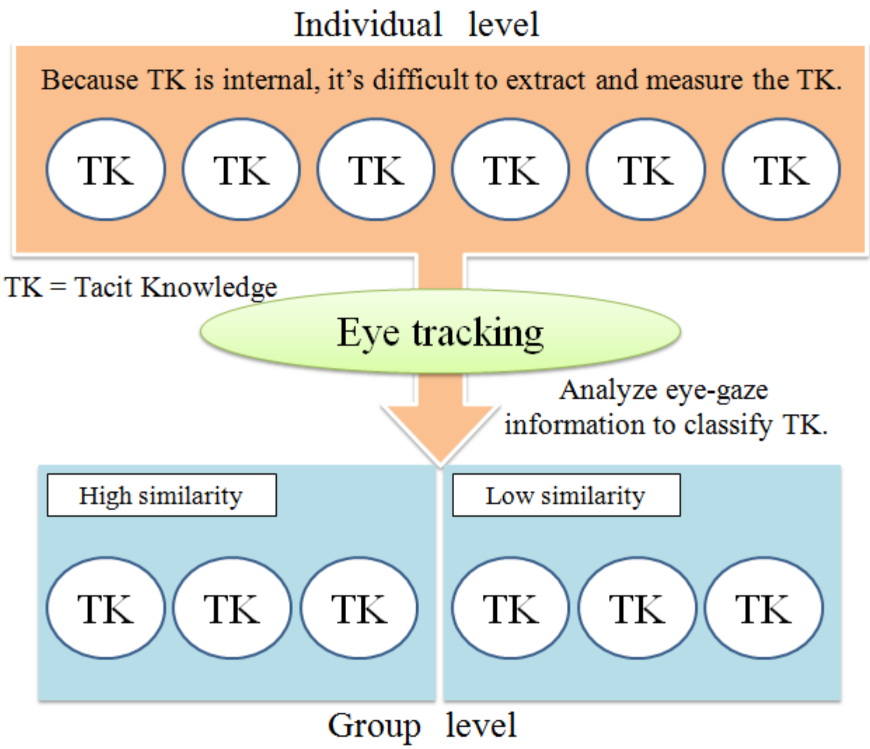


Fig. 3. Classify tacit knowledge by eye tracking

3.1 Eye-Tracking Measurement Experiment

Subjects with Eye-mark-recorder observe six pictures of street scenery all over the place (Figure 4). There were 18 as an object of analytical transaction and the observation time was set in 2 seconds. It is thought that humans instantly obtain essential information from the moment to 2 seconds [6], and observation time was set as short as possible to make handily of analytical transaction and reduce the load of subjects.

Fig. 4. The scene of the measurement eye-action experiment

3.2 Resemble Ratio Judgement in Eye-Gaze Points Using Cos Similarity

To judge resemble ratio of eye-gaze for forming groups was calculated, got the eye-fixation coordinates per subject by picture observation, as an element of vector v_1 and v_2 by the formula (1) of cos similarlity [Sim]. “v_1 v_2” is inner product, $|v_1|$ and $|v_2|$ are intervals of vector v_1 and v_2.

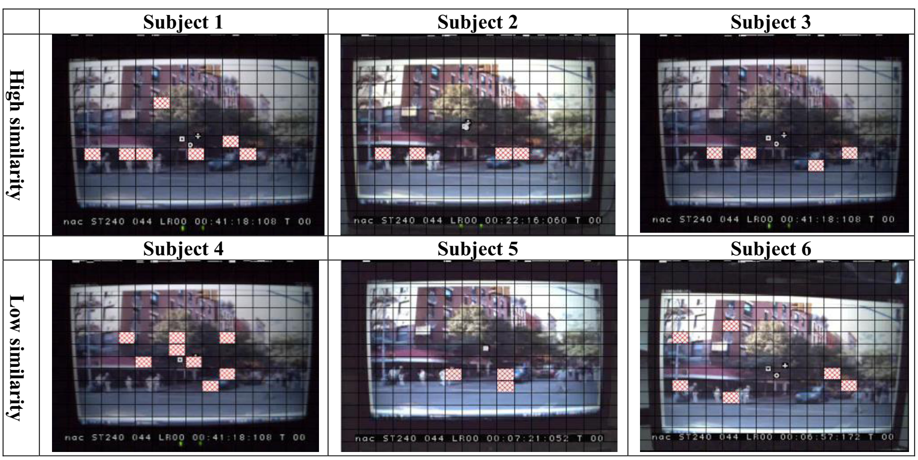


Fig. 5. Analytical images of eye-gaze point

$$\boldsymbol{Sim} = \frac{\boldsymbol{v_1} \cdot \boldsymbol{v_2}}{|\boldsymbol{v_1}||\boldsymbol{v_2}|} \quad , \quad (\boldsymbol{0} < \boldsymbol{Sim} < \boldsymbol{1}) \tag{1}$$

In collective decision making, it is thought that a factor like group dynamics is functioning because collective decision making is complexed as numbers of group increase like four and five compared with individual decision making [7]. From this, for this study: three people are set as group members, the smallest unit which forms a group, held a factor like group dynamics, appeared problems as acquisition of the floor and so on. We formed groups which chooses maximum value and minimum value in each category, cos similarity's combination of vertical direction and horizontal direction in three subjects (Figure 5). And, Figure 6 is the scene of eye- fixation points according to area in each group by Glance analysis software (EMR-dFactory ver.2). Formed groups are high similarity (Sim^H) and low similarity (Sim^L) in Resemble Ratio of Eye-gaze points.

cos similarity	**Sim^HGroup**	**Sim^L Group**
Horizontal	Group A [0.56]	Group D [0.22]
Vertical	Group B [0.61]	Group E [0.21]
Mixture	Group C [0.62]	Group F [0.23]

Fig. 6. Each group formed by using cos similarity

Fig. 7. The scene of the reproduction experiment with pictures

4 Reproduction Experiment with Pictures

Using cos similarity, we conducted reproduction experiment with pictures according to Figure 2 to investigate group-knowledge-creation of in each group from section 3.2 (Figure 6). Before we conducted this experiment, we had administered preliminary questionnaires about presentation picture, and sampled characteristic points of picture. So we made its points evaluation items of questionnaires (**Research A**). After **Research A**, we administered questionnaires to evaluated recall ratio of picture in each group and compared it from the results (**Research B**).

4.1 Characterize Picture (Research A)

Comparing recall ratio, we defined recall ratio as reflect degree of characteristic points of picture, and we had administered preliminary questionnaires to sample it.

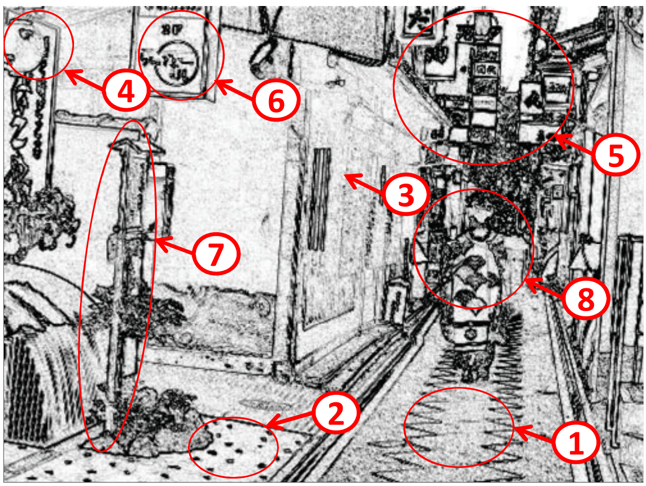


Fig. 8. The displayed picture and characteristic numbers

(1) (2)

Fig. 9. The image drawn by (1) group Sim^H and (2) group Sim^L

We had respondents to marking and describe the points which they thought characteristic or attracting to characterize the picture objectively. This work has the intension of removing arbitrary view of an inspector. From this result, we computed answer ratio to divide number of answers of characteristic points by answer times of all. And to make height of recognition degree clear, the characteristic numbers shown in Figure 7 is defined that the bigger it is, the higher answer ratio is.

4.2 Comparison of the Recall Ratio from Each Characteristic Point (Research B)

We made questionnaires which examines recall ratio from characteristic points defined by **Research A**. In **Research B**, we showed reproduction experiment with pictures to 12 judges in order of random and get answers. From its answers, we calculated each group's recall ratio of pictures (Figure 9) and recall ratio of characteristic points (Figure 10), and analyzed synthetic recall ratio.

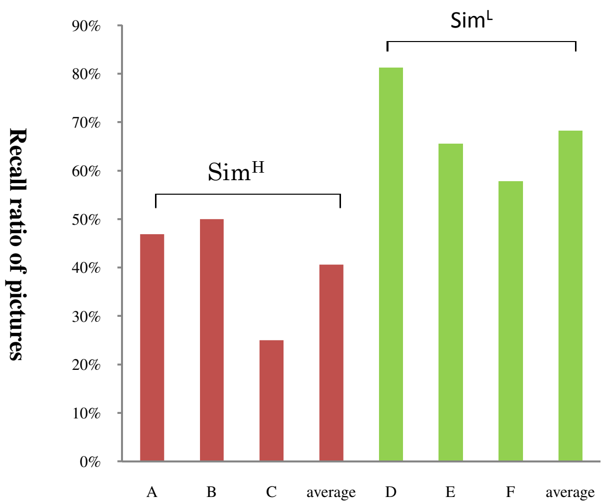


Fig. 10. Each group's recall ratio

Figure 11 is the graph of each group's recall ratio of pictures. This graph says that recall ratio of group Sim^L (D,E,F) is higher than that of group Sim^H (A,B,C). The highest is 81% of group D, the second is 66% of group E, next is 58% of group F, 50% of group B, 47% of group A, and the lowest is 25% of group C which is the highest in cos similarity.

Figure 11 is the graph which shows two groups of recall ratio of characteristic points. It says that in many characteristic points, group Sim^L is higher than Sim^H. Moreover, in characteristic points whose recall ratio is low, group Sim^L has high recall ratio. Its gap is especially remarkable in characteristic points 1. Group Sim^H's recall ratio is only 8%, while Sim^L is 92%.

And Judges also said that "group D and E drew in details.", "group C drew only conspicuous points." and so on, so almost all comments agree to these graphs.

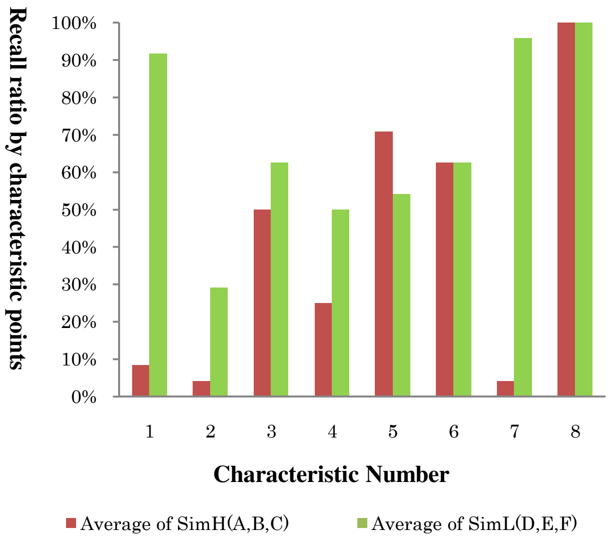


Fig. 11. Recall ratio of every characteristic points

4.3 Consideration

This experiment objectively proved that group Sim^L's recall ratio of pictures is high. And group Sim^L has high recall ratio even in characteristic points which is hardly recognized. Low recall ratio points is what observers little could pay attention to, so group Sim^L could observe well. It's because member's attention dispersed well in group Sim^L. An eye-action arisen from observation pictures is supposed that it refers to "The conversion of Data to Information" in knowledge conversion process and connotes an element of tacit knowledge which contains sensibility information. As a result, we discovered that the group of alien tacit knowledge can practice active knowledge-creation by getting various data.

On the other hand, there is possibility that those results were affected by difference of communication degree in each group. However, in this experiment, there is little factor that is affected by how they communicated. So we can say that we can activate the knowledge creation of group to choose members who have different tacit knowledge.

5 Conclusion

Our study is an unparalleled and high originality in a framework of Organizational Knowledge Creation Study, because the method embodies the group formation that effectively brings out the tacit knowledge of the group. In the past, there were some reports arguing qualitatively the importance of tacit knowledge of groups with a specific case [7] in Organizational Knowledge Creation Study. However, for example, almost all the studies like these, a certain report argued and generated concepts between an organization and another one, and other reports argued consensus building. These studies didn't embody fixed quantity of the tacit knowledge of groups. And also, there were nothing but qualitative methodologies based on using explicit knowledge, such as the mind map study [8] and KJ method study [9] in a framework of Organizational Knowledge Creation Study.

On the other hand, there are a few basic reports of images which targeted individuals on the field of the regeneration memory of recognition science. However, these studies are different from ours who embody the tacit knowledge of groups and argue fixed quantity. So far these studies, such as ours have scarcely researched, because it's difficult to extract tacit knowledge and re-create it.

Therefore the new possibility has come into view, because it makes easily to extract a human response and suppose the inside state of a human by the progress of humans measurement technology in recent years. Our study is a result by being blessed with the progress of measurement technology and leading the analysis methods of measuring data which filled into the purposes. This method enables us to make effective group by observation observing pictures in a short time, so we will continue this study in order to make the method more stable and practical.

References

1. Nonaka, I., Takeuchi, H.: The Knowledge-Creating Company. Oxford University Press, USA (1995)
2. Parent, M., Gallupe, R.B., Salisbury, W.D., Handelman, J.M.: Knowledge creation in focus groups: can group technologies help? Information & Management 38(6), 47–58 (2007)
3. Mitchell, R., Boyle, B.: Knowledge creation measurement methods. Journal of Knowledge Management 14(1), 67–82 (2010)
4. Yukawa, K.: Towards the Knowledge Creating Society through the Use of IT- Community as a new Platform. Economic review 7(1), 108–134 (2003)
5. Ohno, T.: Gaze-Based Interface System. Journal of Information Processing Society of Japan 44(7), 726–732 (2003)
6. Kaiho, H.: Syunkan Zyohosyori no Shinrigaku. Fukumura Shuppan Inc. (2005)
7. Akiyama, M.: Cooperative Decision-Making Process: An Analysis of Information Search Processes and Recall of Alternatives after Choice. Cognitive Studies: Bulletin of the Japanese Cognitive Science Society 5(4), 65–77 (1998)
8. Junhong, Z., Hirano, M., Kubo, T., Feng, L., Peiqian, L.: Tacit Knowledge Management in Craft-Art Manufacturing Enterprises: Cases in China and Japan. Journal of Japan Association for Performance Excellence Online 2(1), 22–29 (2007)
9. Lin, C.: A Knowledge Sharing Support System Based on Mind-Map for Group-Based Software Development Course. IPSJ, Nationwide rally lecture thesis collection the 70th (4), 4–141 (2008)
10. Misue, K.: On Adding Groupware Facilities to D-ABDUCTOR. IPSJ SIG Notes 93(56), 17–24 (1993)
11. Baddeley, A.D.: Human memory: theory and practice. Allyn and Bacon (1990)

Evolution of the Catalogue of Life Architecture

Andrew C. Jones, Richard J. White, Jonathan Giddy,
Alex Hardisty, and Hardik Raja

Cardiff University, School of Computer Science & Informatics,
Queen's Buildings, 5 The Parade, Cardiff CF24 3AA, UK
{Andrew.C.Jones,R.J.White,J.P.Giddy,Alex.Hardisty,H.Raja}@cs.cf.ac.uk

Abstract. The Species 2000 and ITIS Catalogue of Life aims to create and deliver a catalogue of all known species, using a distributed set of data sources. The current Species 2000 software has developed over a number of years, and the system requirements have evolved substantially over the same period. In this paper we discuss the current Catalogue of Life software, the way the requirements are evolving, and major elements of a planned new architectural design being developed as part of the 4D4Life EU e-Infrastructure project. Of particular importance in the new design is to be able to maintain the catalogue, dealing with potential overlaps between supplier databases; to keep it up to date and manage revisions that arise out of changes of scientific opinion; to be able to map between different taxonomies within and outside the catalogue; to be able to provide a wider range of services to other electronic systems which need the catalogue as their "taxonomic backbone"; and to support third-party applications by means of an open platform architecture.

1 Introduction

Since the mid-1990s, the Species 2000 project[1] has been assembling a catalogue of all known species of living organisms, a task begun by Linnaeus in the 18$^{\text{th}}$ Century and abandoned as data grew until computing made it feasible again. When planning to assemble a list of the world's known species (estimated to be around 1.8 million [5]), two obvious strategies could be adopted. One is to obtain lists of species in all the main groups in the taxonomic classification of organisms (mammals, fish, insects, fungi, bacteria, etc.) and just append them "end-to-end". The advantage of this approach is that there should be little or no overlap between these lists; it is rare for classification uncertainty to cause a species to be counted more than once by being present in more than one of these groups. Alternatively we could collect lists of all species known from all regions, countries or other geographical areas and put them together "side-by-side". The difficulty with this approach is that many species will occur in more than one area; they may even be known by different names in different places. Hence the original concept of the catalogue, as initially developed by Species 2000, was to use the former approach, obtaining and integrating lists of species belonging

[1] http://www.sp2000.org/

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 485–496, 2010.

to each major taxonomic group into which living organisms are classified. The term "Global Species Databases" (GSDs) was coined for these lists, as each was required to include all species in its group of organisms for the whole world.

Subsequently the Catalogue of Life partnership[2] was formed between Species 2000 and ITIS (Integrated Taxonomic Information System).[3] Products of the Catalogue of Life (CoL) include an *Annual Checklist* (distributed on CD-ROM every year as well as being available online) and the *Dynamic Checklist* (a regularly-updated catalogue obtained directly from the contributing databases). The original intention was that a Web resource would be made available, which users could consult in order to find out the accepted names for species, their position in the taxonomic hierarchy adopted by Species 2000, etc.

The original SPICE for Species 2000 software [7] was designed around this and other quite restrictive requirements (see Section 2). Meanwhile, other biodiversity informatics resources emerged, including GBIF[4], providing among other things a portal for access to specimen records, and Encyclopedia of Life[5], providing rich descriptive and other species-related information. The need for a Catalogue of Life supporting the organisation of the species information in these resources has been increasingly recognised. Indeed the CoL has the potential to be a key source of taxon concepts for the emerging Global Names Architecture[6]. Similarly, in specific applications such as ecological niche modelling (e.g. [10]) the species distribution information used may include conflicting scientific names for the same or related species; a way to address this problem is needed. In particular, because scientific opinion is continuously changing and developing, the accepted scientific names and the precise species concepts to which they refer are subject to change and sometimes to disagreement. The "synonymic index" provided by the Catalogue is the key to accessing this information, since information about other scientific names is available, linked to names of the species accepted in the Catalogue, as is information about common (vernacular) names [9,8]. Indeed, in the EC-funded i4Life project, which is about to commence, the main aim is to construct better links between species-data projects and communities, including the CoL. The SPICE system has evolved through a number of projects: SPICE for Species 2000 (BBSRC/EPSRC), then Species 2000 europa (EC), and currently 4D4Life (EC). In 4D4Life one of the objectives is to create a new architecture which satisfies the requirements that have emerged recently. An important additional task – introduction of globally unique identifiers into the Catalogue – was commenced with funding from the TDWG infrastructure project. In an important related project ("LITCHI" [6]), techniques have been developed to deal with issues of catalogue consistency. Cardiff University's role throughout these projects has been primarily to design and implement appropriate standards, procedures and software in order to fulfil the CoL requirements.

[2] `http://www.catalogueoflife.org/`
[3] `http://www.itis.gov/`
[4] `http://www.gbif.org/`
[5] `http://www.eol.org/`
[6] `http://www.globalnames.org/`

In this paper we shall discuss the design of the original SPICE system, and then consider how new requirements have emerged which would be difficult to satisfy without a revised architecture. We shall then consider the main elements of a new architecture currently being designed, explaining how it will be specified so as to be maintainable, robust and extensible in the foreseeable future.

SPICE is an example of a distributed, heterogeneous system comprising a large number of contributing databases. In its original form it had a fairly simple specification. We shall discuss the influence this specification had on the design, and reflect on whether it is appropriate to assume these specifications will remain static, given the increasing user expectations that are especially evident in relation to Web-based applications. Another important issue is the contribution the CoL can make within a service-based infrastructure as a source of scientific naming knowledge, supporting "intelligent" searching and linking of data.

2 The Current Catalogue of Life Software

The Catalogue was originally conceived as an entirely dynamic on-line system, in which users would use the Species 2000 web-site interactively as a gateway to obtain species information directly from the providers in real time.

An early prototype was made available on the Web using a system called "CAS2" (Common Access System version 2, described briefly in [7]). This was based on the use of HTML frames. Users expressed their search queries via a form, in which the species name of interest was typed, possibly including "wild-card" characters to search for a range of species simultaneously. The server generated a response in the form of a web page containing a number of frames, arranged side by side. Each frame contained the web address (URL) of one of the GSD's web sites, expressed as an HTTP GET request with a parameter representing the search, copied from the user's request form submission.

Although this demonstrated the general principle of simultaneous search of a number of data providers, it suffered from serious limitations. It relied on the data provider's software responding in the correct way to the HTTP request and providing an HTML response with a format suitable for display in a frame. Some providers simply made available their standard search page, which did not necessarily interpret wild-card characters in the way assumed by the system, nor was the response always formatted for convenient viewing in a narrow column. As the number of providers grew, the frames became narrower and more difficult to manipulate. If a data provider's web-site became unavailable, or they changed their software, the CGI calls would fail to obtain a response. There was no integration of the separate responses, so that users had to inspect and possibly interact with each frame separately. Also, in most cases, a positive response would likely be obtained from only one provider, the one which contained the species in question, the other frames containing blank responses.

To address these issues, the Species 2000 and ITIS Catalogue of Life then developed the SPICE system to harvest species data from GSD data providers. A "Common Data Model" (CDM) was specified, defining the protocol and content of communications both between data providers and the SPICE CAS (or "hub") and between the hub and users or client software. The CDM dissected communications into six fine-grained HTTP GET requests with defined responses [11]. There were two requests to obtain basic metadata about the GSD, one to search for species matching a supplied query, one to return a set of standard data about a single species, and two to navigate up and down the taxonomic hierarchy. This set of simple step-by-step requests was designed to assist with integrating the resulting data and to be compatible with any reasonable data structure which the GSD might use in its database – there was thus no requirement for the GSD to hold its data in a rigidly defined format. Instead, it was assumed that GSD providers would implement an Adapter ("wrapper") which provided protocol and data structure compatibility with the CAS. A wrapper accepts requests from the CAS, searches the GSD for the information requested, and returns it as an XML document in the specified format. As an alternative, it was initially possible to communicate with wrapped GSDs using CORBA. However, although this provided significant efficiency benefits in some cases [12], its use was discontinued due to difficulties in using CORBA to communicate between systems behind firewalls, especially when proxy servers were employed.

The same communication mechanism was implemented to allow users or client software to send requests to the SPICE hub, to search for species information. The SPICE hub can thus be viewed as an aggregator presenting the same interface to clients as the GSDs provided to the hub. Some of the reasons for doing this were to allow other organisations to act as intermediate aggregators between the GSDs and the SPICE hub, to allow the SPICE hub to act as a provider to further aggregators, and to act as a web service for clients – the user interface on the Web used this web service to obtain data from the hub.

For the reasons explained above, this system delivered a product known as the Dynamic Checklist, intended to provide an immediate view of the latest state of knowledge as maintained by the GSDs. However, it became clear that this system had a number of deficiencies. There was a barrier to entry into the system: a GSD provider had to maintain an on-line database at all times, and to write and maintain a wrapper for their data, which also had to remain accessible. Many wrappers and GSDs could not support the efficient use of wild-card characters. All wrappers and GSDs were queried every time a user made a search request, even though most searches would obtain a positive response from only one GSD. The SPICE software had to wait until all wrappers had responded, even if the response was negative, before presenting its response to users. The mechanism worked this way because of an initial agreement in the Species 2000 project preventing Species 2000 from holding copies of GSD data.

However, it became clear that it was necessary to solve the problems with the load on the wrappers and the difficulties with flexible wild-card searches by introducing a central index of which species were included in each GSD. Also, the lack of continuous GSD availability was a major problem, which could be solved by maintaining a central copy of *all* the species data from each GSD. Thus a central cache of all the GSDs standard data was introduced into later versions of the SPICE software. This substantially improved the availability, scalability and functionality of the system. For example, to provide a complete response to the user it was no longer necessary for all the GSDs to be available on-line. Similarly, GSDs and their wrappers were not subject to heavy loads by having to respond separately to every user request; instead, they only had to support provision of their data to "refresh" the cache when they were updated. Further benefits arose because the cache could be tailored for more efficient, flexible use. For example, searching operations became more efficient because the cache could be optimised to support search requests, whereas the GSD databases might have been optimised for other goals. Furthermore, Searches involving more than one GSD became straightforward because all the data was in one cache database – no distributed querying was required. Indeed, more complex and flexible searches also became possible, as these no longer relied on the required functionality being implemented at the GSD level. More general use of wildcard characters, fuzzy searching where the exact spelling of a name is unknown, and boolean combinations of search criteria all became more feasible, although not all these features have been implemented.

Experience with operating the CoL over several years in the form described above, with the SPICE version 5 software, has shown that many further improvements could be made. Some can be achieved merely by user interface redesign, e.g. to improve search criteria flexibility; others would need more radical changes, e.g. to enhance management of taxonomic content and improve sustainability.

3 Revised Requirements

A key purpose of the 4D4Life project is to put the CoL onto a robust, reliable and sustainable footing to serve the needs of biologists and other users as an authoritative source of species name information. Thus the requirements list for the new CoL system architecture serves this purpose. The list also takes into account the deficiencies in the current system as described above, and the additional flexibility required of the new architecture, to be easier to maintain and manage and to support a variety of modes of data delivery. To make the requirements intelligible to project partners who are not computer scientists, we summarised them in a series of scenarios or "user stories". The scenarios outlined below emphasise potential extensions to the current functionality which we wish to support in the design of the new system architecture.

1. Searching for species
   - At present one can navigate up or down the taxonomic hierarchy to a particular higher taxon of interest, or search for species by name, but these species location methods cannot be combined. It is useful to be able to search for a species by name within a selected higher taxon group.
   - As additional data types are planned to be included in the standard data set for species, including geographical data, it is desirable to support full boolean combinations of search conditions.

  - Users should be able to view the additional data relating to the species set currently being considered, e.g. to help in selecting species of interest.
  - Users should be able to narrow down or broaden a set of species retrieved, by performing further searches within or outside the subset already retrieved, resulting in a union or intersection of search result sets.
2. Services for using species data
  - Subsets of taxa identified by the mechanisms above should be available for use in various ways, including downloading as a report formatted for printing, as a data set suitable for further processing, or as a self-contained database for local use.
  - A software client should be able to present a string believed to represent a species name and receive a list of species which have been referred to by that name. Subsequent actions by the client or a human user may select one of more of these taxa for further processing.
  - Having obtained a name or unique identifier of a species of interest, a client may request a list of names which have been used to refer to that species, to enhance queries to other sources of species data and thus avoid the danger of missing information which has been recorded under an alternative name.
  - A client may request additional information about selected species, including geographical distribution, web links to other on-line resources, references to authoritative sources, etc.
3. User and client interactions with the data. Some of the following user actions will be facilitated if users are able to register with the CoL and customise their preferences, which are stored for their subsequent logins.
  - A user registers an interest in a particular higher taxon, and is notified whenever the data for this taxon changes.
  - A user may annotate a taxon record, so that the annotation is displayed only when they return to that record themselves, or can be seen or amended by a specified group of users, or by anyone, with the annotation being marked as having been created by the first user.
  - A user reports what he or she perceives as an error or omission in one or more taxon entries, and this is fed back automatically to the GSD provider. A management tool tracks progress in resolving the issue, and reports to the original user, data provider and a manager of the CoL.
4. Comparing regional data sets. The basic CoL concept is to maintain a complete Catalogue in a "global hub," co-ordinating caching and provision of Catalogue data. However, views of taxonomy may vary from one region of the world to another and the "hub" concept can be extended accordingly by introducing "regional hubs" to maintain such distinctive views.
  - A user may wish to compare data about one or more selected species in two hubs (one of which may be the global hub), perhaps by viewing the taxonomic hierarchies side by side or requesting a report on the differences between the treatment of the species in the two hubs. Such comparisons might be used for various purposes, including creating a merged consensus database (e.g. as a candidate for further development

to form a new GSD) or establishing a "cross-map" (e.g. to facilitate navigating between hubs, so a user viewing species in one regional data set can easily find the corresponding species in another regional view).

5. Managing the Catalogue
   - Although the SPICE version 5 design reduced the load on the GSDs by requiring them only to respond to requests for re-caching their data, rather than to respond to every user request individually, the re-caching process was inefficient. Using the Common Data Model v. 1.21 [11], a GSD can only indicate the date and time of its last update, not the particular data which have been updated. Thus when a change occurs in the GSD, however small, the entire GSD database has to be re-cached. For some GSDs, this presents a significant load, or is even infeasible due to limitations of their database, software or network bandwidth available. Hence a mechanism is needed for updating the cache incrementally by requesting and receiving only the taxa which have changed.
   - A means is required to propagate alerts both to other components of the system (e.g. to notify that a database has been revised, so it will need to be re-imported to the CAS cache) and to systems managers (e.g. to notify of failure, inconsistency, etc).
6. Synchronisation of data sets within and beyond the Catalogue
   - An organisation, e.g. the CoL itself, which maintains several copies of the CoL database for efficiency, may wish to propagate changes from an updated copy to other copies in an efficient and timely manner.
   - A user may wish to set up their own database of species-related information, using a copy of the CoL as a framework for organising the species. A facility is required to download a copy of a selected part of the global or a regional database in a format appropriate for importing into an existing database (e.g. a delimited text file) or for creating a new database (e.g. an SQL dump).
   - A user who has previously created a database using the CoL as a template, as described above, may have entered further data of their own. Realising that the CoL has since been updated, they require the ability to propagate recent changes from the CoL into their own database.
   - Similar scenarios may occur where the databases are not identical, but changes made to one need to be propagated appropriately to another, maximising information transfer without causing loss of data integrity (this may require alerting human experts to resolve incompatibilities).

4 Meeting the Revised Requirements: The e-2 Architecture

We can divide up our consideration of general characteristics of the new "e-2" architecture required to meet our requirements into three areas: ensuring sustainability and flexibility of the e-2 architecture; tools for scalable maintenance and expansion of the catalogue, and mechanisms for tracking change.

4.1 Ensuring Sustainability and Flexibility of the e-2 Architecture

The e-2 architecture must be maintainable over a considerable number of years, as computing technologies come and go. It needs to cope with heavy use in scenarios where it acts as a link to other information, and internally it must support a more flexible approach to the composition of the Catalogue from various kinds of information sources. It will also need to adapt to support uses of the Catalogue information unforeseen today. This leads to three main elements:

A maintainable Architecture. One key to maintainability is to ensure loose coupling and "thin interfaces" between system elements, allowing us to modify component implementation without affecting the behaviour provided to or expected by other linked components. As we intend to deploy these components over a distributed network of sites, we will adopt a Service-Oriented Architecture (SOA). An SOA provides a component-based model, at a granularity appropriate for deployment in a distributed system. However, our experience with CORBA in previous versions of SPICE demonstrated the extent to which a technology can pervade the software architecture, making it very difficult to remove dependencies on the technology later. To achieve appropriately loose coupling between components, we intend to use the Service Component Architecture (SCA). SCA should also provide the ability to redeploy services easily on new technologies as they emerge. In addition, where possible we shall delegate specific tasks to "off-the-shelf" tools. For example, we are considering using software such as `balance`[7] to do the load balancing for us in a multi-server environment.

A Scalable, Service-Based Architecture. A network of regional hubs will allow the CoL to replicate its data on all or most of these hubs. The existence of the full global data set and the regional databases on a hub will provide the same advantage mentioned above (searches do not need to address the original distributed sources), but also will permit the extensive comparisons and cross-mapping of regional databases needed to support a variety of services for Catalogue maintainers and users, without needing queries to be distributed across all regional hubs. Replicating data sets in this way improves robustness but requires a synchronisation mechanism. The Catalogue changes relatively slowly, so guaranteed full instantaneous consistency between sites is non-critical; rapid update propagation is sufficient. Incremental scalability on demand could be achieved by adding more sites, one option being to use a cloud computing facility.

Communications between hubs and services such as described in section 3 will need to be flexibly defined. The precise way services will be exposed may change over time due to technology changes. Thus, decoupling service implementation from the way they are exposed is important. A particular way the CoL could be used is in the Linked Data [2] "ecosystem", allowing CoL species information to be used to link other resources. In such scenarios CoL services will need to be published to a public registry, to allow them to be discovered.

[7] `http://www.inlab.de/balance.html`

One complexity in implementing a scalable system is that we plan to provide user accounts, to allow users to benefit from persistent customisations, such as annotations of data, and to control the visibility of such customisations, whether they are private, public, or only visible to a defined group. The maintenance of such dynamic data and its composition with the relatively static public Catalogue in a scalable manner is an on-going research issue within the project.

A Platform Architecture for Future Unforeseen Needs. Platform architectures [1] can be characterised by a small set of essential static core components, together with a larger more variable set of peripheral components. Interfaces between components, especially those of the core components, are tightly defined and controlled. Platform designers can publish the interface details as open specifications, permitting third-party development of novel peripheral components. New "products" based on the platform become easier to engineer, saving the need to "re-invent the wheel" each time. The CoL designers cannot necessarily foresee all possible future applications of the content nor can they in all circumstances justify expenditure of effort to develop and support all applications. Our approach will be influenced by platform architecture principles, permitting others with sufficient motivation to develop a wide range of niche applications.

4.2 Tools for Scalable Maintenance and Expansion of the Catalogue

Two strategies for assembling the Catalogue were described above. The first, using GSDs to assemble the Catalogue, was originally adopted and used as the mode in which SPICE gathers species lists. However, this strategy relies on the existence, enthusiasm and commitment of major taxonomic institutions, experts in the various groups, and even dedicated amateurs, to assemble these check-lists and databases. Such committed institutional projects and individuals do not exist for many less well studied groups of organisms, and funding for truly international projects is scarce.

The alternative strategy, of collecting together lists from regional and national projects and databases, has the advantage that many projects are funded from national or regional sources to list the species in their region. To make use of these regional and national resources, the CoL has established the beginnings of a system of "regional hubs". Assembling a world check-list from regional and national lists poses challenges in overcoming the duplication of species and potential for the use of alternative names.

A third strategy is to use name lists (nomenclators). Projects such as the International Plant Names Index (IPNI)[8] and the Online Registry for Zoological Nomenclature (ZooBank)[9] exist to catalogue all species names as they are published or encountered in older literature. In less well studied groups, where fewer changes in classification have taken place, these lists approximate to rough lists of species, which can be used to fill the gaps where GSDs do not yet exist.

[8] `http://www.ipni.org/`

[9] `http://zoobank.org/`

- A tool is needed to allow a user or data manager to view two taxonomic hierarchies side by side. When he or she performs a browsing operation on one tree (such as scrolling up or down, or opening or closing branches), the view of the other tree is adjusted correspondingly so that the parts of the two trees being viewed are as similar as possible. The differences between the trees, if any, may be highlighted. It must also be possible to create links between the nodes in the trees viewed, to support (i) connection of the Species 2000 hierarchy to the internal GSD hierarchies at the appropriate points and (ii) switching between different taxonomic hierarchies.
- Another tool is needed to provide a report on the differences between two or more checklists for the same organisms in different regional hubs . This might be used for various purposes, including creating a merged consensus database (for example, as a candidate for further development to form a new GSD) or establishing a "cross-map" (for example, to facilitate navigating between hubs, so that a user viewing a species or small set of species in one regional data set can easily find the corresponding species in another regional view).

Further tools are needed to support synchronization of hubs; propagation of alerts, e.g. for GSD updates; recovery from system failures such as synchronisation, server or service failure; report generation; and automated integrity/quality control checking when taxonomic components of the system are updated.

4.3 Mechanisms for Tracking Change

Scientific names of organisms are not sufficiently precise in their use to permit the accurate integration of data from different sources. A common feature of various solutions for this problem that have been proposed is the definition of a unique identifier for each species or other taxon in the classification. However, not all these solutions provide for a mechanism to track the changes in a species concept when a species is revised or split or merged with other species as knowledge improves. For example, ITIS defines Taxon Serial Numbers (TSNs) which denote species in the ITIS catalogue, but technically refer only to the accepted names of the species.[10] Having a unique identifier for each species name is important, and is key to the success of nomenclatural resources such as IPNI and ZooBank, but it is a different need from the one we are discussing here in the context of the CoL as a global authoritative reference source of taxonomic relationships. Here, each taxonomic *concept* needs to be uniquely and unchangingly labelled.

The solution currently being considered to document and report changes in biodiversity databases is to use persistent unique identifiers which are "resolvable", i.e. they can be used to obtain further information about the taxa, potentially including the changes which have occurred from one version to the next. GBIF recently commissioned a task group, the LSID-GUID Task group (LGTG) to research and write a report with recommendations on the possibilities [4]. Two approaches have been identified for providing a resolution mechanism, which are being adopted by some biodiversity organisations – Life Science Identifiers

[10] `http://www.itis.gov/itis_primer.html`

(LSIDs)[11] [3] and HTTP URIs following the "Linked Data" principles [2]. It is important to note that the essential common requirement is to track and maintain information about changes as they occur, in a standard format interpretable by software. LSIDs and Linked Data are merely two different ways to provide the same information; they are not mutually exclusive.

These solutions to the taxon-change tracking problem do not themselves actually obtain information about species concept changes directly. They rely on one or other of two possible methods for taxon changes to be discovered: either the GSD responsible for a sector of the Catalogue needs to report the changes which it has made in its list, or alternatively and less reliably, these changes have to be inferred by comparison of versions of their lists at different points in time. An example of the latter is the program used by the CoL to compare successive Annual Checklists to identify taxa which have been changed or newly introduced in an Annual Checklist edition.

These developments will support the unification of the Annual and Dynamic Checklist creation workflows, because it will become much easier to track the concepts selected for each Annual Checklist edition, and the Annual Checklist becomes simply a special edition among the large number of editions of the Dynamic Checklist which will be available as each year progresses.

5 Summary and Future Work

In this paper we have described how the user requirements have evolved substantially since the inception of Species 2000. We did anticipate that some performance requirements would eventually emerge, e.g. by implementing a cache before most GSDs would allow us to use it, in order to demonstrate its importance if adequate performance is to be achieved. Nevertheless, we discovered that committing ourselves to a complex system such as CORBA made it difficult to extend or revise the system. Architectural simplicity, loose component coupling and a platform approach are the key aims that we are now seeking to achieve.

We have explained how the CoL evolved from being primarily a system serving Web pages into a system which must interoperate with other systems that use species data, to support their operations with synonymic information, etc. The Catalogue has an important role to play in linking and allowing information to be discovered which might otherwise not be found due to naming differences.

We have described the general features of our new e-2 architecture. The next stage is to define more precisely how the mechanisms described in this paper – such as alerting – will function. This depends on more fundamental decisions such as what technology will be adopted for interacting between the components.

Acknowledgements

SPICE and LITCHI were originally funded by grants from the UK BBSRC/EPSRC Bioinformatics Committee. The second phase of these two projects was

[11] `http://lsids.sourceforge.net/`

funded as part of an EC Framework Programme 5 grant, and the current phase is funded as part of an EC Framework Programme 7 grant, as is the proposed i4Life project. The initial introduction of LSIDs into the CoL was funded by the TDWG infrastructure project.

Collaborators in these projects at the Universities of Cardiff and Reading and other institutions are gratefully acknowledged, as are a wide range of institutions whose data has been used in the projects.

References

1. Baldwin, C., Woodward, C.: The architecture of platforms: A unified view. In: Gawer, A. (ed.) Platforms, Markets and Innovation, pp. 19–44. Edward Elgar Publishing Ltd., Cheltenham (2010)
2. Bizer, C., Heath, T., Berners-Lee, T.: Linked data - the story so far. Int. J. Semantic Web Inf. Syst. 5(3), 1–22 (2009)
3. Clark, T., Martin, S., Liefeld, T.: Globally distributed object identification for biological knowledgebases. Briefings in Bioinformatics 5(1), 59–70 (2004)
4. Cryer, P., Hyam, R., Miller, C., Nicolson, N., Tuama, É.Ó., Page, R., Rees, J., Riccardi, G., Richards, K., White, R.: Adoption of persistent identifiers for biodiversity informatics (November 2009), `http://www2.gbif.org/Persistent-Identifiers.pdf`
5. Hammond, P.: Magnitude and distribution of biodiversity. In: Heywood, V. (ed.) Global Biodiversity Assessment, pp. 113–138. Cambridge University Press, Cambridge (1995)
6. Jones, A.C., Sutherland, I., Embury, S.M., Gray, W.A., White, R.J., Robinson, J.S., Bisby, F.A., Brandt, S.M.: Techniques for effective integration, maintenance and evolution of species databases. In: Günther, O., Lenz, H.J. (eds.) 12th International Conference on Scientific and Statistical Databases, pp. 3–13. IEEE Computer Society Press, Los Alamitos (2000)
7. Jones, A.C., Xu, X., Pittas, N., Gray, W.A., Fiddian, N.J., White, R.J., Robinson, J.S., Bisby, F.A., Brandt, S.M.: SPICE: a flexible architecture for integrating autonomous databases to comprise a distributed catalogue of life. In: Ibrahim, M., Küng, J., Revell, N. (eds.) DEXA 2000. LNCS, vol. 1873, pp. 981–992. Springer, Heidelberg (2000)
8. Paton, A.: Biodiversity informatics and the plant conservation baseline. Trends in Plant Science 14(11), 629–637 (2009)
9. Patterson, D., Remsen, D., Marino, W., Norton, C.: Taxonomic indexing - extending the role of taxonomy. Systematic Biology 55(3), 367–373 (2006)
10. Peterson, A.T., Vieglais, D.A.: Predicting species invasions using ecological niche modeling: New approaches from bioinformatics attack a pressing problem. BioScience 51(5), 363–371 (2001)
11. White, R.J.: Species 2000 Common Data Model, version 1.21 (2004), `http://biodiversity.cs.cf.ac.uk/sp2000/protocol/CommonDataModel1.21.rtf`
12. Xu, X., Jones, A.C., Pittas, N., Gray, W.A., Fiddian, N.J., White, R.J., Robinson, J.S., Bisby, F.A., Brandt, S.M.: Experiences with a hybrid implementation of a globally distributed federated database system. In: Wang, X.S., Yu, G., Lu, H. (eds.) WAIM 2001. LNCS, vol. 2118, pp. 212–222. Springer, Heidelberg (2001)

A Consensus Method for Checklist Integration

Kevin Richards[1], Aaron Wilton[1], Christina Flann[2], and Jerry Cooper[1]

[1] Landcare Research, Lincoln 7640, New Zealand
[2] Netherlands Centre for Biodiversity Naturalis (section NHN), Biosystematics Group, Wageningen University, Netherlands

Abstract. A project was initiated in 2006 to compile a draft global checklist of the Compositae plant family by integrating existing electronic checklists. A Checklist Integrator Tool (C-INT) was developed. C-INT provides an editorial interface alongside managed and automated workflows for the importation and integration of multiple checklists. A number of technical issues were encountered and addressed during the project, including the requirement to maintain provided data values, difficulties integrating data updates, issues with integrating relational data, and computational efficiency.

Keywords: Taxonomic Checklists, Taxon Name Matching, Biodiversity Data Integration.

1 Introduction

A project was initiated in 2006, supported by the Global Biodiversity Information Facility (GBIF) [1], to compile a working global checklist of the Compositae plant family by integrating existing regionally comprehensive electronic data sources (e.g., Euro+Med, SANBI, IPNI, Tropicos, Conabio, Flora of Japan).

The project had two primary informatics objectives: (1) to develop the informatics tools necessary to manage the integration of multiple existing checklists into a single consensus view, whilst maintaining the linkages to original records in the individual checklist datasets, and, (2) to create a baseline integration of key datasets from identified project partners.

To meet these objectives the Checklist Integrator Tool (C-INT) was developed. C-INT provides an editorial interface alongside managed and automated workflows for the importation and integration of multiple checklists. The main features of C-INT are: (a) importation of data structured as an XML document or a local database, (b) integration of taxon names, taxon concepts, and literature data using a configurable decision tree and matching process that may be optimised to reduce computation time, (c) calculation of consensus checklist records based on a simple majority of the linked original records, whilst allowing for editorial contributions to have priority, (d) management, integration and presentation of other data components that are often included in checklists (e.g. vernacular names and distribution data), (e) produce an associated website and web services.

C-INT is now actively being used to create the draft Global Compositae Checklist (GCC). To date C-INT integrates more than 60 datasets ranging in size from 220 records to nearly 160,000 records to maintain the GCC [2].

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 497–505, 2010.

A variety of requirements and issues were identified during the project. Some of these were addressed, whilst others were not. Importance considerations were:

- The technical ability of data providers varied, and therefore the structure, complexity and content of provider datasets also varied.
- Consensus data values for each checklist object (primarily taxon names), resulting from integrating provider values, needed to be calculated whilst maintaining a transparent audit trail of each provider contribution. This audit trail needed to be summarised as reports for end-users of the system.
- Changes to provider datasets potentially resulted in inconsistencies in the recalculated consensus data.
- Fields that are expected to be reasonably standardised, such as authors, actually vary substantially. The use of standard data dictionaries and thesauri during the integration process proved effective.
- The calculation of consensus values for each field for large numbers of records is a computationally intensive process. Improving the efficiency of this process for real-time editing and dynamically changing provider data is critical.

2 Provider Data - Consensus Record Calculation

Various methods for the integration of taxonomic checklists have been attempted. For example, LITCHI [3] and GPC [4] adopted an integration process that included matching algorithms within a managed workflow/decision tree process. Other projects have focussed on the matching algorithms for reconciling names, e.g. TaxaMatch [5] and Taxon Finder [6]. The general application of these algorithms has been on matching names from various sources and merging the result into a single list of taxonomic names. A decision tree approach was taken in developing C-INT. The key elements of automated process facilitated by C-INT are: 1) provider name records and taxon concept relationships are maintained unchanged, 2) the integration process consists of a decision tree and matching algorithm to identify potentially equivalent records from different providers, 3) equivalent provider records are merged using a set of rules to calculate a consensus record, 4) this consensus record is linked to the provider records from which it is generated. Thus there is a transparent audit trail of the integration process and how provider records differ from each other and the consensus record. These differences provide valuable feedback to the data providers. This approach also allows the consensus record to be recalculated at any time by running the integration process over the sets of provider records. The integration process is shown in Figure 1, together with a manual editing component which will be described shortly.

The integration of simple lists of taxonomic names does not provide any additional information other than the concordance of name strings amongst providers. Additional value is added to an integrated checklist when the provider recordsets contain a richer set of data, including synonyms and parent/child taxon concept relationships. Within the C-INT integration process the relationships between fields within a recordset were modelled according to Taxonomic Concept Schema (TCS) [7]. The use of a TCS-based model allowed the encoding of names distinct from the taxon concept relationships which link those names. This facilitates the integration process and sub

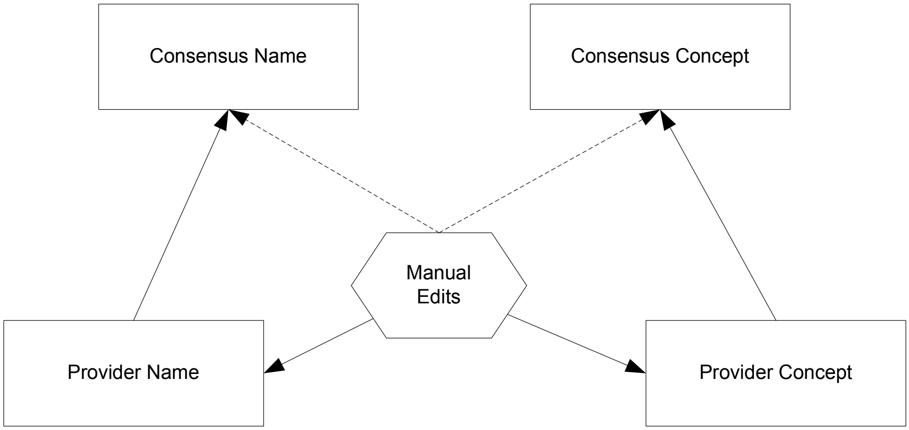


Fig. 1. The data structure used for provider integration – consensus calculation

sequent calculation of consensus records by using the provided parent-child concepts and synonymy concepts to provide a richer context to assist the matching process. The end user of data from the integrated checklist can see these calculated consensus records together with the linked provider records, and reports of equivalences and differences between providers.

It is unlikely that any automated process for integration will result in a correct set of calculated consensus records and there will remain many instances where manual intervention is necessary to edit and correct consensus checklist records. The role of an editor is therefore an essential part of the integration process and the compilation of a quality GCC. However, because C-INT was designed to operate against dynamically changing data from linked provider records, it remains important that the consensus records can be recalculated from the set of provider data at any time. This is facilitated by treating the editorial input as another data provider, the "Checklist Editor" (Fig. 1). Records produced by the Checklist Editor differ in status from other provider records because they have priority over other provided records in recalculating the consensus records. This approach has worked well.

3 Provider Data

The scope of the data that were provided in this project varied enormously, which is to be expected given the large number of providers. In some cases a simple list of taxon names and synonyms was provided, and in the most complete case, a full TCS XML document was provided. A requirement for importing the provider data into C-INT was that all data provided in a simple form were converted to the more structured TCS format. This data preparation step occupied a relatively large portion of the project time.

An important issue concerning the conversion of simple data (without an explicit taxonomic hierarchy) into the more structured TCS format is that parent concepts of records may have to be inferred from the context of the data and this may result in an incorrect attachment of the record into a consensus hierarchy, with the potential creation of unnecessary duplicate consensus records.

Even when a complete taxonomic hierarchy is provided there can be issues. For example two provider datasets may incorporate the same species but employ a different chain of parent taxa to encode the hierarchical classification. The hierarchy may differ in a minor way, for example by including an extra rank (e.g. tribe). A simple comparison of parent fields for each taxon would then result in the insertion of two identical species records with two different parent hierarchies, where in fact there is just one species record with a largely identical hierarchy. It is necessary to configure the decision tree and matching process to accommodate this broader context of a name and a chain of parent concepts.

Globally Unique Identifiers (GUIDs) play and important role in identifying and linking biodiversity information [8] based on unambiguous references to objects. For example, in our integration of provider recordsets they allow individual records within the integrated dataset to be identified and updated without uncertainty. If a data provider does not maintain and supply GUIDs for their data objects it is often impossible to determine which record should be updated during a data update. The only solutions for this issue that have been found, so far, are to: a) identify the appropriate provider record to update by using the name-matching rules (e.g. compare epithet, author and year values), and then recalculate a new consensus record; or b) by removing all data for that provider from the integrated dataset, refreshing the entire provider dataset, and then recalculating the entire set of consensus records.

A related issue is the inability of some data providers to provide deprecated record information. For example, if a record is deleted from the provider's dataset, and this is not recorded explicitly, it may not be possible to determine that a provider record should be removed and the consensus record recalculated. It is possible, therefore, to end up with consensus records that no longer reflect the current state of the provider records. Under these circumstances deprecating all the records from such providers is a pragmatic solution. The current provider data is recorded in a separate audit trail before it is replaced with the new provider data.

4 Integration Method and Issues

4.1 Integration Method

The integration mechanism consists of the following workflow:

1. Import new provider record(s)
2. Invoke a decision tree and matching process against all new provider records to determine potential matches with the existing set of consensus records. This process uses a configurable decision tree and matching rules, comparing field within records, such as parent name, rank, name epithets, authors, year and publication. Where necessary the decision tree also triggers a more detailed comparison of a new provider record against the set of existing provider records linked to each consensus record.

3. A new record is inserted into the checklist for each provider record where no matches were found, and an equivalent consensus record generated.
4. If multiple matches were found then the record is flagged as requiring editorial intervention to resolve potential matches.
5. If exactly one match was found then the provided record is linked to the existing consensus record for the group of equivalent linked provider records. The consensus record is then recalculated, so that it reflects the impact of the newly linked provider record.

The process for recalculating a consensus record is:

1. For each field of the consensus record (e.g., epithets, authors, year, basionym, publication year), retrieve the set of values from the linked provider records (example shown in Fig. 2)
2. If there is a simple majority value it is used for the consensus value of that field. (e.g., if the provider values for Publication Year are “1838”, “1838”, “1841”, the majority value applied to the consensus record is “1838”). However, if there is conflict with no majority, then a null value for that field is inserted into the consensus record (Fig. 2).

Provider records

Id	**Rank**	**Canonical (terminal epithet)**	**Authors**	**Year**	**Published In**
1234	genus	*Flotowia*	Endlicher	1838	Ench.
4433	genus	*Flotowia*			
5520	genus	*Flotowia*	Endl.	1841	

Consensus Record

Id	**Rank**	**Canonical (terminal epithet)**	**Authors**	**Year**	**Published In**
AC44-FE…	genus	*Flotowia*	Endlicher	NULL	Ench.

Fig. 2. Consensus record generation

4.2 Integration Issues

Several issues have arisen specific to this integration process.

A null value is used for a consensus field when there was no majority value for the contributing provider fields. This may result in false matches when subsequent provider records are matched against consensus records. For example, a set of provider records have conflicting data for the year of publication (provider records with years of 1935 and 1934) and application of the decision tree results in a consensus value

a) Provider records

Provider	Name	Author	Year	Rank	Parent Record
1	*Eupatorium cinereum*	Raf.	1836	Species	*Eupatorium* Tourn. ex L.
2	*Eupatorium cinereum*	Raf.	1836	Species	
3	*Eupatorium cinereum*	Raf.		Species	
1	*Eupatorium*			Genus	*Eupatorieae* Cass.
1	*Eupatorieae*	Cass.		Tribe	*Compositae*

b) Calculated consensus records

AA123…	*Eupatorium cinereum*	Raf.	1836	Species	*Eupatorium*
AA124…	*Eupatorium*	Tourn. ex L.	Null	Genus	*Eupatorieae*
AA125…	*Eupatorieae*	Cass.	Null	Tribe	*Compositae*

c) Resulting hierarchy

- *Compositae*
  - *Eupatorieae* Cass.
    - *Eupatorium* Tourn. ex L.
      - *cinereum* Raf. 1836

d) Updated provider records

Provider	Name	Author	Year	Rank	Parent Record
1	*Eupatorium cinereum*	Raf.	1836	Species	*Eupatorium*
2	*Eupatorium cinereum*	Raf.	1836	Species	
3	*Eupatorium cinereum*	Raf.		Species	

e) Recalculated consensus records

AA123…	*Eupatorium cinereum*	Raf.	1836	Species	Null
AA124…	*Eupatorium*	Null	Null	Genus	Null

f) Resulting hierarchy

??

- ??
  - *Eupatorium*
    - *cinereum* Raf. 1836

Fig. 3. Updated recordset and orphaned records

that is null. Subsequently a new provider record is added with the year 1922, but other fields match the consensus fields. The new record would also match the consensus record, but the recalculated consensus record would again contain a year of publication field that is null. This is clearly incorrect. In this case we have two publications that are probably the same because they differ by one year (usually the result of delayed distribution), whereas a publication date of 1922 probably refers to a different publication, and therefore constitutes a different record. The consensus values for the year of publication fields for these two records should not be null.

There are a number of possible refinements to the general decision tree rules where there is no majority value. For example, a selected provider value can be copied to the consensus field (for example the later year, where publication year differs by one). Alternatively, the decision tree could trigger a comparison of new provider record fields with all the provider record fields linked to the same concept. This fine tuning of the decision tree rules depends on the field being considered. For example there are a set of core fields, name, rank and parent, where conflict without majority should not default to a null consensus value and one provider value must be selected.

The effectiveness of the integration process can be significantly increased with richly structured and populated datasets from providers. However, dynamically changing richly structured data may also create issues. For example, consider the following situation: Providers 1, 2 and 3 provide details for the taxonomic name "*Eupatorium cinereum* Raf." (Fig. 3a) which results in calculated consensus records (Fig. 3b), and a consensus classification (Fig. 3c).

Provider 1 then updates a recordset (Fig. 3d) and, for some reason, removes the explicit parent concept hierarchy.

The recalculated consensus record (Fig. 3e), linking all similar provider records for the species *Eupatorium cinereum,* is now orphaned in the consensus classification (Fig. 3f). For this reason it is necessary to adjust the decision tree to ensure the appropriate calculation and propagation of a valid parent classification.

4.3 Integration Aids

The use of external standard thesauri (values and synonyms) for fields is useful for increasing the accuracy of integrating taxonomic checklists. Several data fields are particularly suited to this approach, including authors of names, taxon ranks and publication titles. Provider records often contain various representations of author names (full author names, author teams, abbreviations, etc). Integration is markedly improved by using a list of standard author values and known equivalents. This thesaurus may be derived from the data itself and incrementally constructed during the integration process, or it may be an existing standard from an external source. However, providers are often consistent in the way they use an abbreviation within their dataset, but that same abbreviation may not be used consistently across different providers. As a consequence it is preferable to map abbreviations to a standard list per provider rather than globally.

5 Efficiency of the Integration Process

The algorithms used to match provider records to consensus records, and for calculating a consensus record from the associated provider records, are computationally intensive and the time taken to process an entire set of provider records can quickly become unmanageable. Processing time is a significant practical consideration in C-INT because human intervention for editing the data is part of the workflow. It is therefore essential to provide a rapid process for data integration and consensus calculation.

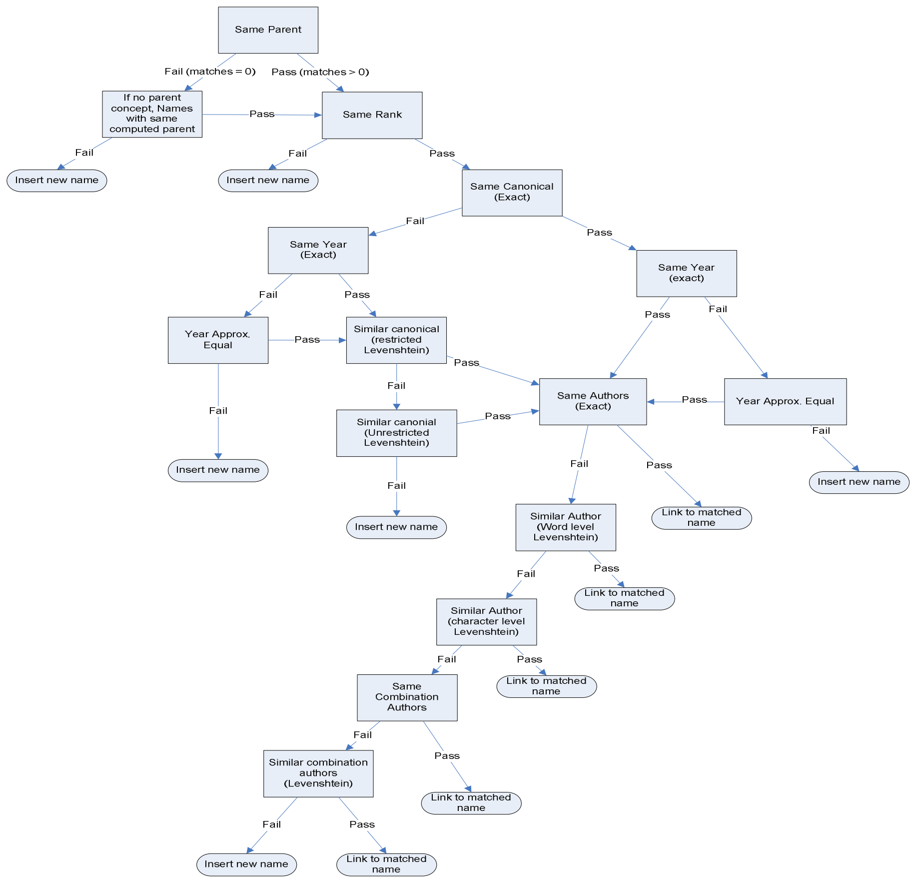


Fig. 4. Example decision tree & matching process

The decision tree and field matching process within the C-INT application is configurable and can be manipulated to improve the efficiency of the whole process. The processing of a provider record originally took anywhere from 12 to 60 seconds, depending on whether the calculation of a Levenshtein distance metric for string comparisons of text was required. The performance was improved by optimising the decision tree to prune the possible matches that require subsequent computationally intensive processing. An example of the configurable decision tree is shown in Figure 4. Optimisations include limiting the possible set of matching names/concept records by including only those records with the same rank and parent name. Such optimisation lowered the average processing time to approximately 2 seconds per record. However, even with this efficiency gain, processing an average dataset of 20,000 new provider records against 150,000 consensus records and 400,000 linked provider records takes approximately 11 hours on a current standard Windows-based server.

Similarly, the process for recalculating a consensus record from its associated provider records is computationally intensive, and was optimised to approximately 2 seconds. However, for the approximately 200,000 names in the GCC, the total computation time required to recalculate the set of consensus records on the same platform is 4.6 days.

It will be possible to improve the calculation times of these algorithms by introducing multi-threaded processing, and cluster computing.

6 Summary

The C-INT tool integrates multiple checklist data using a provider-consensus data model, where the provider records are maintained in their original form, and linked to automatically generated consensus records calculated using a configurable decision tree and matching process. Manual editorial changes to the calculated consensus checklist are treated as a special type of provider record.

The processes adopted in C-INT have proven to be accurate and efficient. The application of C-INT to over 60 provider datasets in the Global Compositae Checklist project has resulted in a draft working checklist containing, 152,000 consensus names and 253,000 consensus concepts integrated from 425,000 provider names and 313,000 provider concepts within 60 datasets from across the globe.

The main issue encountered in C-INT relates to optimising the process to deal with real-time updates and deletions of provider data.

References

1. Global Biodiversity Information Facility, `http://gbif.org/`
2. Global Compositae Checklist, `http://compositae.landcareresearch.co.nz/`
3. LITCHI project, `http://biodiversity.cs.cf.ac.uk/litchi/`
4. Global Plant Checklist project of IOPI (International Organization for Plant Information), `http://plantnet.rbgsyd.nsw.gov.au/iopi/iopigpc1.htm`
5. TaxaMatch, `http://code.google.com/p/taxon-name-processing/wiki/TaxamatchInfo`
6. uBio Taxon Finder, `http://code.google.com/p/taxon-finder/`
7. Taxon Concept Schema, `http://www.tdwg.org/standards/117/`
8. Life Science and Guid Task Group of GBIF, `http://imsgbif.gbif.org/File/retrieve.php?PATH=4&FILE=2efc20187e6ad3dd828bbeadaa1040e6&FILENAME=LGTGReportDraft.pdf&TYPE=application/pdf`

Tools for Semantic Annotation of Taxonomic Descriptions

Hong Cui, Partha Pratim Sanyal, and Chunshui Yu

University of Arizona, 1515 First Street, Tucson
Arizona, USA 85719
{hongcui,ppsanyal,chunshui}@email.arizona.edu

Abstract. A software application for automated semantic annotation of taxonomic, especially morphological, descriptions is reported in this paper. The tool is based on unsupervised machine learning methods. It is designed to annotate descriptions in a deviated syntax that is not normal English but often used in morphological descriptions. The unsupervised annotation system does not need any training examples to annotate text descriptions. It uses a relevant glossary available to it but aims to learn as much information as possible from the text itself. Tools such as this are needed to reformat free-text or OCRed taxonomic documents to a semantic-explicit format for easy and intelligent access, providing character data for phylogenetic research, climate impact on biodiversity, and traditional biosystematics research.

Keywords: Semantic Annotation, morphological descriptions, supervised machine learning, unsupervised machine learning, software applications.

1 Introduction

The majority of taxonomic descriptions of organisms are in an "unstructured" format, such as print, Word, PDF, HTML, or TXT. These formats are designed for human consumption and are difficult for computers to process. The knowledge locked in unstructured documents, on the other hand, is on high-demand by biological research from whole organism to genomic studies. Efforts such as those led by the Biodiversity Heritage Library (BHL) are converting print literature to an electronic format: As of March 15, 2010, BHL has OCRed 28,658,777 pages from 76,484 volumes. OCRed text, although in digital format, is still unstructured and difficult to use by computers. Automated semantic annotation tools are needed to complete the mission. Once the documents are semantically annotated, other sophisticated, even intelligent, applications can be built to utilize the knowledge explicitly annotated in the documents. A few foreseeable applications include: (1) A organism identification key generator; (2) A taxonomic information integrator that merges incomplete descriptions from multiple sources to form a more complete record; (3) A taxonomy author's aid that automatically annotates a manuscript to a preferred digital format and compares the manuscript with existing authorized documents for quality assurance; and (4) A character comparator that compares a selection of morphological descriptions quickly and intelligently for similarity and differences.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 506–516, 2010.

Although the development of semantic annotation tools for biodiversity literature started at least 20 years ago, there are still several open challenges. Two of these are: (1) The vocabulary diversity inevitably presented in biodiversity literature. Not having a machine readable lexicon for all biodiversity domains makes all NLP (Nature Language Processing) tools lose their power in processing domain text. (2) The deviated syntax that is commonly used in legacy documents (see the *description* element shown in Fig. 1b for an example) but is not typically dealt with by existing NLP tools. Most current semantic annotation tools work around these challenges by either limiting the target document collections, reengineering the tool on a collection by collection basis, or relying heavily on user input and intervention.

While workarounds are often necessary, we believe the design of the tool must aim to address the challenges, at least over time if not immediately. In this paper, we report on our design and implementation of a semantic annotation system for biodiversity descriptions. We argue that the system helps to address the challenges by using an unsupervised learning process that obtains needed information from the text itself, as opposed to hand-crafted training examples, grammar, or lexicons. We also note the limitations of the system and make suggestions for further research.

The paper is organized as follows. Section 2 reviews related research. Section 3 describes the unsupervised system and explains why unsupervised methods may be able to tackle the first two challenges noted above. Section 4 provides a description of the annotation system itself. Finally, the paper is concluded with directions for future research.

2 Review of Related Research

The earliest work related to semantic annotation of biodiversity literature we found is [1], in which a syntactic parser was made to extract organ names, characters, and character states from volumes of Flora of South Wales and Flora of Australia. The grammar needed for the parser was hand-made after studying the syntax of the descriptions. The lexicon was created using the accompanying glossaries as a base. While one can expect a fine-tuned grammar and lexicon to achieve close to perfect precision, the author did not scientifically evaluate the recall of the system, but gave an estimate at 60%-80%. The author noted that in moving the parser to a new volume or a different taxon group (e.g., ants), the grammar would have to be revised and the lexicon changed. Taylor suggested a number of ways to learn the lexicon automatically, but noted learning grammar rules would be harder. Some of Taylor's suggestions on learning the term were implemented in [2]. Abascal and Sanchenz [3] took a similar approach, but only parse a description to paragraph level.

Studies [4] and [5] performed a character level extraction with hand-crafted dictionaries, ontology, and/or checklists on a limited number of plant descriptions. The accuracy reported in [4] was 76%, while Wood et al. reported 66% recall and 74% precision. The descriptions used in [5] were from six English Floras and they reported that the redundancy existing across the floras helped to improve the recall.

Study [6] used a supervised machine learning algorithm to extract leaf shape, size, color, arrangement and fruit/nut shape character states from volumes of Flora of North America. The algorithm learned regular expression extraction patterns from hundreds of manually annotated examples, and then applied the extraction patterns to new examples. Depending on the character, the extraction precision ranged from 75% to 100% and recall from 33% to 80%.

The system reported in [7] is one of the two systems currently available for public download (http://plazi.org/?q=GoldenGATE). The system is called GoldenGATE for it was built on the architecture of GATE (http://gate.ac.uk/). Different from all above systems, GoldenGATE is an interactive editor that offers a number of functions a user can invoke to annotate taxonomic documents, for example, annotated scientific names appearing in a document. Although, character annotation is not a function currently provided, GoldenGATE has a very flexible and easy to use mechanism for an expert user to develop a plug-in for character annotation. We plan to integrate our unsupervised annotation algorithm as a plug-in to GoldenGATE after it has been tested on a good number of taxon groups.

The other publicly available system is our supervised learning based annotation system for plant descriptions, MARTT (MARkuper for Taxonomic Treatment, http://sites.google.com/site/biosemanticsproject/project-progress-wiki). MARTT learns a statistical annotation model from the training examples and applies it to annotate a morphological description in plain text format to sentence level (a sentence is a string of characters terminated by a period or a semicolon). The downloadable MARTT has been trained with five volumes of Flora of North America and seven volumes of Flora of China and can be used to annotate other plant morphological descriptions according to an XML (eXensible Markup Language) schema accessible at http://publish.uwo.ca/~hcui7/research/xmlschema.xsd. MARTT can be re-trained with other examples annotated according to this or another schema. A recorded demo of MARTT showing the use of it for creating training examples, correcting annotation errors, and annotating a newly composed description is accessible from the download page. Detailed information about MARTT and its learning algorithms can be found in [8]. The recall and precision performance of the system when evaluated on Flora of North American (FNA), Flora of China (FoC), and Flora of North Central Texas descriptions was 80%-95%.

The limitations of MARTT include: (1) MARTT processes morphological descriptions only. If morphological descriptions are embedded in their parent documents, they must be extracted first. (2) MARTT requires hundreds of training examples. Like all supervised learning algorithms, it is not capable of learning outside of the training examples. In other words, if "phyllaries" is not tagged in the training example, MARTT will not be able to annotate a statement about phyllaries. It will tag the statement as "unknown". (3) MARTT only annotates sentences (i.e., organ or structure names) and not characters.

The main reason we discontinued developing MARTT for character annotation is the amount of manual work it would require preparing an adequate number of training examples. These training examples would have little value beyond annotating descriptions of the same or related taxon groups.

3 Unsupervised Annotation Methods

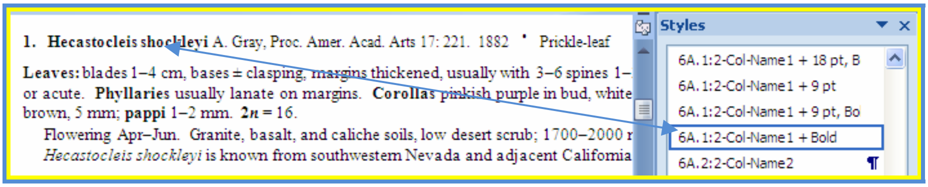
Several design decisions were made with respect to the three challenges discussed in Introduction section. (1)Do not assume the availability of a lexicon; instead, use the unsupervised machine learning algorithm reported in [9] to learn domain terms as they appear in document text. (2)Use modular design and minimize the interaction among the modules. Any intermediate results are saved in a file or a database to make it possible to adopt a new method in the annotation process when needed.

The unsupervised algorithm uses a bootstrapping procedure to discover organ names and character states in description text without needing any training examples, predefined schemas, or dictionaries. It exploits a basic syntax pattern of descriptive statements: a statement = organ/structure name + a set of descriptors. Starting with a few organ/structure names, the algorithm infers that the word following the organ/structure name is a descriptor (i.e., belonging to some character state). With the discovery of new descriptors, words before them are then learned as organ/structure names. These bootstrapping iterations continue until no new organ/structure name or character state is learned in an iteration. In a sense, the algorithm learns a lexicon in which a word is either an organ/structure name or a character state. It then uses the lexicon to annotate descriptive sentences. The reported annotation performance on a volume of FNA and a volume of Treatise of Invertebrate Paleontology was above 90% accuracy.

Study [1] suggested that using "or" and "to" may help discover character states from description text. For example, if either *green* or *brown* is known to be a character state, the statement "bark dark green *or* brown" can be used to infer the other is also a character state, because they are connected by *or*. We have implemented a simple algorithm bootstrapping ("that bootstraps" seems clearer than "bootstrapping," but I'm not sure this is the correct terminology) character states using this clue. The algorithm discovered hundreds of character states from volumes of FNA and FoC, many of which were not included by the Categorical Glossary for FNA or PATO (Phenotypic Quality Ontology) [2].

Taxonomic description text itself can provide other useful clues for solving other problems typically seen in semantic annotation. Examples include: (1) Normalizing hyphened words; [9] reported a dynamic programming algorithm that normalizes hyphened words by finding their normal form from description text. (2) Extracting morphological description paragraphs from heterogeneous source documents. The bootstrapping procedure we developed for this task made use of the fact that in morphological description paragraphs, sentence subjects are typically organ/structure names. Even though the algorithm starts knowing zero organ/structure names, it can easily find a few from the text itself by looking for paragraphs that are long but without any articles (a, an, the), help verbs (i.e., is, are, etc.), and pronouns. The few paragraphs with these traits are morphological descriptions in Linnaean style (e.g., bark brown), from which organ/structure names can be extracted using the unsupervised lexicon learning algorithm mentioned above. With more organ/structure names, more morphological descriptions can be identified, from which more organ/structure names can be learned. The precision of this procedure is estimated to be above 95% and the recall is currently under evaluation.

These examples suggest that there is much to be exploited from taxonomic descriptions themselves. Gathering needed information from text itself reduces an annotation system's reliance on manually created lexicons, grammar rules, or training examples; hence addressing the vocabulary diversity challenge identified above. The capability of learning a lexicon from domain text may also address the deviated syntax challenge. In a small scale exploratory experiment, we evaluated the Stanford Parser's



(a)


```xml
<?xml version="1.0" encoding="UTF-8"?>
<treatment>
  <number>1.</number>
  <species_name>Hecastocleis shockleyi</species_name>
  <authority>A. Gray</authority>
  <place_of_publication>Proc. Amer. Acad. Arts 17: 221. 1882</place_of_publication>
  <common_name>Prickle-leaf</common_name>
  <description>
    <leaf>Leaves:</leaf>
    <leaf_blade size="1-4 cm">blades 1-4 cm,</leaf_blade>
    <bases architecture="clasping">bases ± clasping,</bases>
    <margins count="3-6" size="1-3 mm">margins thickened, usually with 3-6 spines
    1-3 mm near bases,</margins>
    <apices apex="mucronate;acute">apices mucronate or acute.</apices>
    <phyllary pubescence="lanate">Phyllaries usually lanate on margins.</phyllary>
    <corolla color="pinkish purple, white" size="10 mm">Corollas pinkish purple in
    bud, white at flowering, ca. 10 mm.</corolla>
    <cypsela color="brown" size="5 mm">Cypselae brown, 5 mm;</cypsela>
    <pappus size="1-2 mm">pappi 1-2 mm.</pappus>
    <chromosome count="16">2n = 16.</chromosome>
  </description>
  <flowering_time>apr</flowering_time>
  <flowering_time>spring</flowering_time>
  <flowering_time>may</flowering_time>
  <flowering_time>jun</flowering_time>
  <flowering_time>summer</flowering_time>
  <habitat>granite, basalt, and caliche soils</habitat>
  <habitat>low desert scrub;</habitat>
  <elevation>1700-2000 m;</elevation>
  <us_distribution>Calif.</us_distribution>
  <us_distribution>Nev.</us_distribution>
  <discussion>Hecastocleis shockleyi is known from southwestern Nevada and adjacent
  California.</discussion>
</treatment>
```


(b)

Fig. 1. A taxon description from FNA. (a) In Word format before annotation. (b) In XML format after annotation.

syntax parsing performance on a set of 20 randomly selected sentences before and after a human-mediated assignment of Part of Speech (POS) tags to organ and character names in those sentences. In the first parsing, 75% of the sentences were parsed incorrectly. After assignment of POS tags, only 10% of the sentences were parsed incorrectly. This result suggests that we may not need to create special syntax parsers for biodiversity domains after all. We only need to feed general parsers like the Stanford Parser good POS tags.

4 Unsupervised Semantic Parser

The development of the unsupervised annotation system was started with a grant for the Flora of North America Project. While it meets the requirements of the FNA project, it is designed with non-flora descriptions in mind as well. The Semantic Parser is implemented in Java and Perl, but also needs MySQL and WordNet. It currently runs on Windows computers.

The Semantic Parser was designed to handle three types of taxonomic works. Type I includes works that have a proof in MS Word® format with styles meaningfully named. Examples include Flora of North America (FNA) and Flora of China (FoC). In FNA, for example, the styles applied to family names and synonyms of family names are named "#A#FamName#" and "#B#FamSyn#" ("#" represents a number). These named styles are useful for the Parser as they can be directly transformed to semantic annotations. Fig. 1a shows the named styles used in an FNA volume. Type II includes works that are a compilation of taxonomic descriptions like FNA but lack named styles. Examples include Treatises on Invertebrate Paleontology. Regardless of the original file format, these works are converted to plain text format, preserving only paragraph breaks. Type III includes works that have morphological descriptions embedded in them. Examples include journal articles from ZooTaxa and text OCRed by BHL. For Type I and II works, the Parser will attempt to annotate a complete taxonomic description, including sections on nomenclature, habitat, and distribution, in addition to morphological descriptions. In contrast, the Parser will only identify and annotate morphological description paragraphs in a Type III document, as the document may include a variety of other unrelated information, for example introduction, literature review, methods, and future research in a journal article.

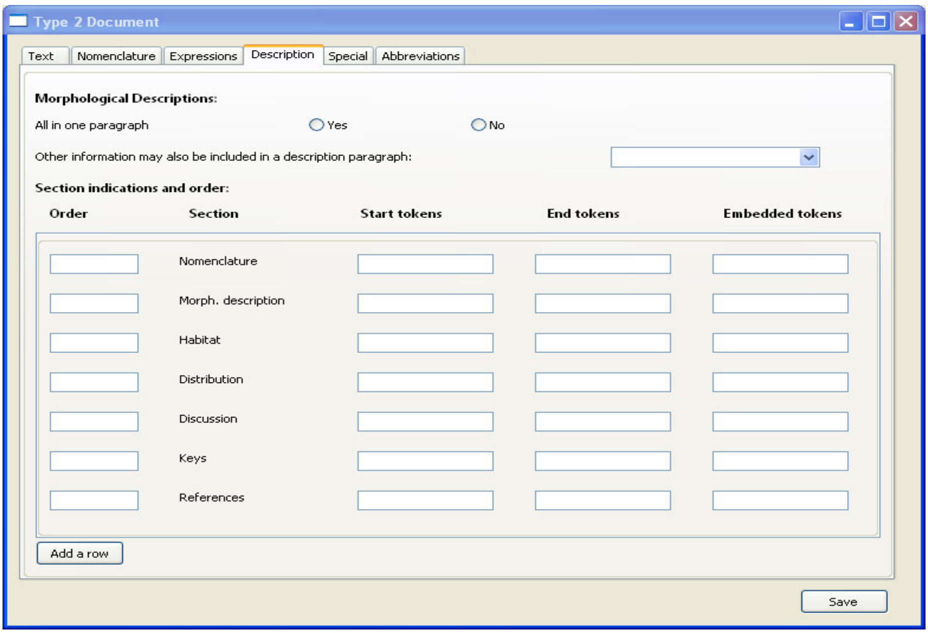
Fig 4. shows the main interface of the Semantic Parser for Type I works. The tabs show the steps one needs to take to complete the annotation process of a volume of a Type I work. In what follows, we shall explain the way the Parser annotates a Type I work in detail, using FNA as an example, and then note Type II and III documents can be handled with a few additional modules.

To use the Parser to annotate a volume of Type I work (i.e., a Word document with named styles), the user will need to set up a database and three folders in the file system: source, target, and conf. The source folder is where the to-be-annotated document resides, the target folder is where the output of the Parser is saved, and the conf folder holds configuration information the Parser uses. Next, the user will need to save the Word document as .docx format, then change the .docx file extension to .zip. When the file is unzipped, a file named **document.xml** can be found in the folder. This XML file is in Office Open XML(ISO/IEC 29500:2008) format. The named styles are used in the XML file as attribute names. Put document.xml in the source folder.

If there is a relevant glossary for the volume, the glossary file in CSV (Comma Separated Values) format may be saved in the conf folder. If a list of taxa included in the volume is available, the list should also be saved in the conf folder. If a taxon list is found in the conf folder, the Parser will use it to verify part of its parsing results. Otherwise, the Parser will generate such a list from its parsing results for the user to verify. Initially the target folder is empty.

Now the user can start the parser. When the Parser starts, the user is asked to identify the type for the task document. When the user chooses "Type I", the user is presented with a table to map the Word styles to the semantic tags to be used in the annotation. The mapping information can typically found in an Editor's Guide. Fig. 2a shows the mapping table. The result of the mapping is saved in the conf folder. Next the user is presented with the interface shown in Fig. 3.

(a)



(b)

Fig. 2. Style configuration of the Semantic Parser for Type I and II works. (a) Style configuration for Type I works. (b) Style configuration for Type II works.

The tabs list a series of steps to be taken to parse the entire volume to individual files annotated in a well-formed XML format.

On the "*General*" tab, the user enters the directories for configuration information (the conf folder), source file (document.xml), and target directory to save intermediate and final markup results. "*Save*" the information would allow the user to come back at a later time to resume a stopped process by simply "*load*"ing the information and

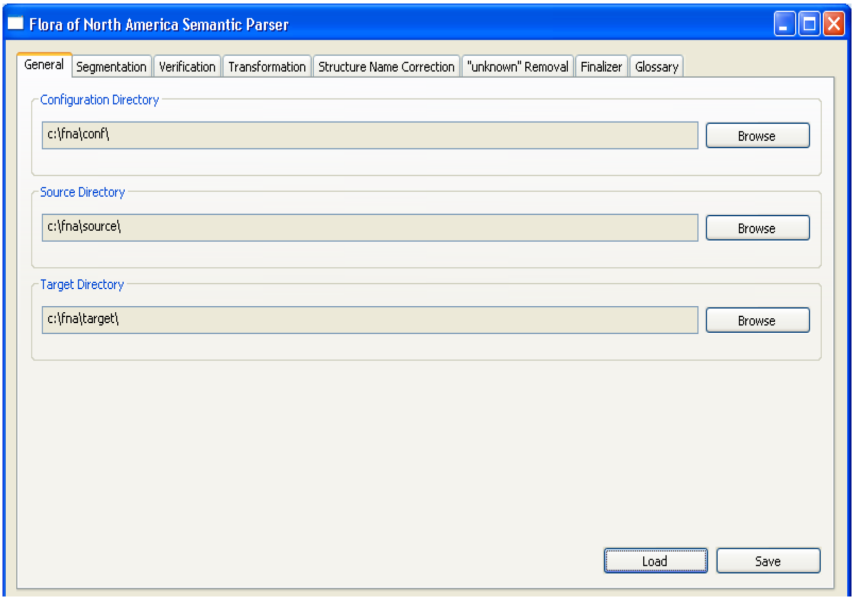


Fig. 3. Main interface of the Semantic Parser for Type I works. The tabs show the steps one needs to take to complete the annotation process.

going directly to the stop point. Then the user goes step by step through the annotation process:

1. *Segmentation* segments a volume to individual taxon files. Resulted files are saved in the *extracted* folder in the *target* directory. Fig. 4a shows the result of this step. Each file is a live link; when clicked, the taxon file is displayed to the user for review.

2.*Verification* checks individual taxon against the taxon list file in the conf folder, if there is one, and reports any discrepencies between the two. Any discrepencies discovered in this step must be resolved before proceeding to the next step. Fig.4b shows three extracted names do not match those in the taxon list.

3. *Transformation* transforms Word style tags to semantic tags using the style mapping generated earlier. This step also annotates nomenclature, flowering time, habitat, and distribution information, but not morphological descriptions. Certain transformations are performed on the annotation. For example, when flowering time is described as "summer" in a description, annotations of specific summer months (i.e., June, July, and August) are added, and vice versa. The Transformation tab lists transformed files with a column listing parsing errors (e.g., unparsed text: ") sect. Rumex" for file 1056.xml). The user can click open a file to correct the errors. Example annotations produced for nomenclature, flowering time, habitat and distribution in this step can be seen in Fig. 1b. It is in this step that morphological description paragraphs are saved to a separate folder for the next step to annotate.

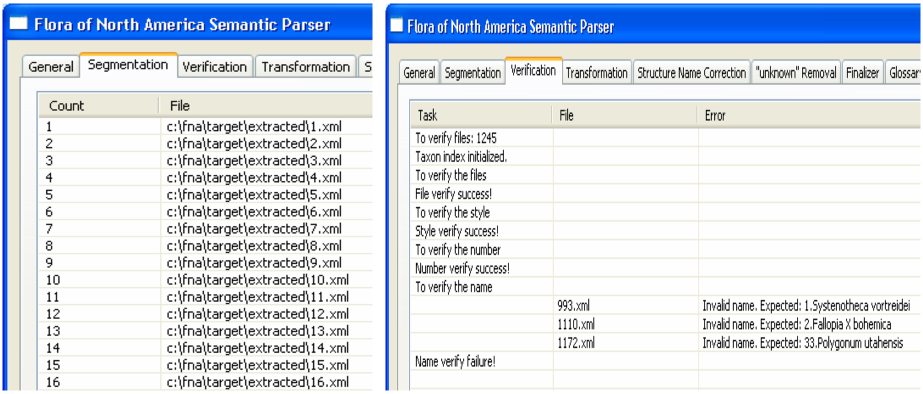


Fig. 4. Steps for annotating non-morphological description sections. (a) Result of Segmentation. (b) Result of Verification (a and b are not labeled in this figure).

4. *Structure Name Correction* annotates morphological description paragraphs using the bootstrapping-based unsupervised algorithm reported in [9]. The algorithm exploits the information redundency in descriptions to learn/discover organ names and character states.The table displayed on this tab lists the structure names learned by the algorithm. The user can select inappropriate ones and remove them from the list. Due to the removal of certain structure names, some sentences annotated with the names are re-annotated as “unknown” for the next step to handle.

5. *“unknown” Removal* interacts with the user to annotate sentences marked as “unknown”. Clicking on the ***Load*** button causes these “unknown” cases to be displayed for the user to assign tags. Selecting any sentence, the context in which the sentence appears is displayed in the context box. The user may use “modifier” and/or “tag” drop-downs to select an appropriate annotation or enter new values. *Save* saves all annotated sentences and refreshes the list. The user may terminate the program at any time. When the user comes back the next time, the Parser displays the “unknown” cases they left the last time. Behind the scenes, the updated annotations are saved in the MySQL database. Besides morphological descriptions, other information saved in the database includes taxon names, authority, place of publication, and glossary terms. Fig. 5 shows this tab.

6. *Finalizer* performs two functions at present: (1) annotates characters/character states in morphological descriptions, and (2) puts the morphological description paragraph back to their parent taxon description, outputting the complete XML file for an individual taxon. We plan to add a new tab before *Finalizer* and devote it solely to character level annotation, as it is a module still requiring much improvement. We will discuss this function more in the Future Research section.

7. *Glossary* compares the glossary, if one is found in the conf folder, with what is learned by the annotation algorithms automatically. It reports the terms (structure names or character states) that are not covered by the given glossary and how the annotation algorithms understand these new terms.

Each tab, except Configuration, displays a progress bar when needed to show the user the percentage of the step completed.

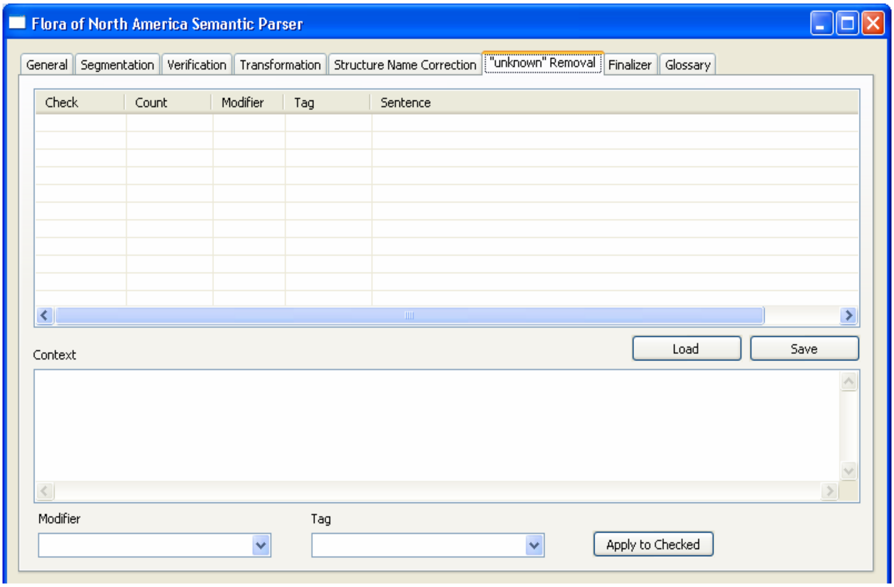


Fig. 5. The Unknown Removal tab of the Semantic Parser for Type I works

For Type II works, where no explicit styles are included in the text, a more detailed configuration template (Fig. 2b) is provided to collect information from the user. This information will be used by a module to segment and annotate non-morphological description paragraphs. Then the user can continue the annotation process with the Structure Name Correction, which annotates morphological description paragraphs.

As mentioned before, the parser only attempts to identify and annotate morphological description paragraphs in a Type III work. An unsupervised algorithm described in section 3 will be used to extract morphological description paragraphs from their heterogenous parent documents. We are developing and evaluating the performance of the Parser on Type II and III works.

5 Conclusion and Future Research

We argue in this paper that unsupervised learning methods can obtain much information needed by an automated semantic annotation system for biodiversity publications from the text itself. The rest of needed information is likely obtainable from the formalism of the publication process. We present an application built on this belief.

Improvements are needed in several respects. First, while the application is accumulating domain terms learned from its annotation experience, these terms need to be fed back to the annotation process for future tasks. The application reported here used the FNA Glossary as the source for annotating character states, but through systematic study [2], we know the algorithm discovered more terms than are included in the Glossary. We plan to add a terminology tab before the character annotation step in the application, on which a user can verify the learned terms and integrate them into the Glossary. Second, even with an enhanced glossary, the current character annotation algorithm has to be improved to capture the complex relations presented in morphological descriptions. For example, the statement “margins thickened, usually with 3–6 spines 1–3 mm near bases” shown in Fig.1b mistook spine count of 3-6 as margins count, which should not have a count semantically. The annotation did not capture the relation that “margins has spines”, nor the constraint “near bases” for the measure 1-3 mm. To annotate relations and constraints, the sentence structure needs to be parsed. We plan to experiment with the Stanford Parser, with a supply of POS tags for organ name and character states from our unsupervised annotation algorithm.

References

1. Taylor, A.: Extracting Knowledge from Biological Descriptions. In: Proceedings of 2nd International Conference on Building and Sharing Very Large-Scale Knowledge Bases, pp. 114–119 (1995)
2. Cui, H.: Competency Evaluation of Plant Character Ontologies against Domain Literature. Journal of American Society of Information Science and Technology (in press)
3. Abascal, R., Sanchenz, J.: X-tract: Structure Extraction from Botanical Textual Descriptions. In: Proceedings of the String Processing & Information Retrieval Symposium, pp.2–7 (1999)
4. Diederich, J., Fortuner, R., Milton, J.: Computer-Assisted Data Extraction from the Taxonomical Literature, `http://math.ucdavis.edu/~milton/genisys.html`
5. Wood, M.M., Lydon, S.J., Tablan, V., Maynard, D., Cunningham, H.: Using Parallel Texts to Improve Recall in IE. In: Proceedings of Recent Advances in Natural Language Processing, pp.70–77 (2003)
6. Tang, X., Heidorn, P.B.: Using Automatically Extracted Information in Species Page Retrieval, `http://www.tdwg.org/proceedings/article/view/195`
7. Sautter, G., Agosti, D., Bohm, K.: Semi-Automated XML Markup of Biosystematics Legacy Literature with the GoldenGATE Editor. In: Proceedings of Pacific Symposium on Biocomputing, pp.391–402 (2007)
8. Cui, H.: Converting Taxonomic Descriptions to New Digital Formats. Biodiversity Informatics 5, 20–40 (2008)
9. Cui, H., Boufford, D., Selden, P.: Semantic Annotation of Biosystematics Literature without Training Examples. Journal of American Society of Information Science and Technology (in press)

Automated Pre-processing Strategies for Species Occurrence Data Used in Biodiversity Modelling

Marshall J. Heap and Alastair Culham

Center for Plant Diversity and Systematics, School of Biological Sciences, University of Reading, Whiteknights, P.O. Box 217, Reading, Berks, RG6 6AS, U.K.

Abstract. To construct Biodiversity richness maps from Environmental Niche Models (ENMs) of thousands of species is time consuming. A separate species occurrence data pre-processing phase enables the experimenter to control test AUC score variance due to species dataset size. Besides, removing duplicate occurrences and points with missing environmental data, we discuss the need for coordinate precision, wide dispersion, temporal and synonymity filters. After species data filtering, the final task of a pre-processing phase should be the automatic generation of species occurrence datasets which can then be directly 'plugged-in' to the ENM. A software application capable of carrying out all these tasks will be a valuable time-saver particularly for large scale biodiversity studies.

Keywords: Biodiversity richness; environmental niche modelling; pre-processing species occurrence data; automated filtering and occurrence file generation.

1 Introduction

The term *biodiversity* is used in literature to describe the variety of biological organisms present within a specific geographic extent. High biodiversity is generally considered synonymous with a healthy ecosystem [1]. Most biodiversity studies are concerned with understanding and mitigating biodiversity loss. The most common causes of biodiversity loss amongst plant species, include; land-use change, climate change, atmospheric gas composition change, soil damage and the spread of invasive species [2]. Species richness, the number of species present in a specified geographic area, is the most commonly used measure of biodiversity [3]. Where species richness is estimated using Environmental Niche Models (ENM's), the typical approach is to model distribution probabilities for each species and then sum these probabilities for each grid cell in the chosen geographic extent [4]. ENMs combine species occurrence data with layers of environmental data in raster format. Species occurrence data consists of geo-referenced occurrence points and the species' scientific name. Environmental data consists of raster layers that define both geographic extent and the spatial resolution of the modelled species probability distributions. A machine learning algorithm (e.g. neural network, genetic algorithm, maximum entropy etc.) first finds the environmental grid cell values corresponding to each species occurrence point and

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 517–526, 2010.

then divides pattern space between environmentally suitable/unsuitable habitat. Computer-based ENM's are split between two funadamental approaches; namely, presence only (PO) and presence/absence (PA) models. They are distinguished by the inclusion of geo-referenced absence data in PA models. However, since absence data is often difficult to obtain, a pseudo presence/absence (PA) generative approach, is often used instead. MAXENT [5] is an example of an ENM that uses this approach - computing its own absence data by drawing random data samples from the models environmental layers excluding species occurrence locations.

Model accuracy is measured by omission and commission error. The extrinsic omission rate (also known as sensitivity) represents the fraction of test species samples located in an unsuitable environment i.e false positives. Commission error (also known as specificity) is the fraction of absences falling in a suitable environment i.e. false negatives [6]. Model performance is measured quantitatively by the area under the receiver operating characteristics curve (AUC of ROC curve, henceforth referred to as AUC) which is a threshold independent method. Creating the ROC requires division of species point data between training and test data sets. Essentially, AUC measures the probability that a presence location is ranked higher than a random background location [7]. Maximum AUC values are close to, but less than 1, with 0.5 representing a prediction no better than random. While the Test AUC score indicates good model performance, it does not indicate if the right model has been built. Here reliance is placed on expert knowledge in choosing the right combination of environmental layers combined with expert analysis of the resulting probability distribution to see if it is a fair representation of the species' fundamental niche.

An important source of species occurrence data for global and other spatially extensive ENM's is the Global Biodiversity Information Facility (GBIF) [8]. Some GBIF data statistics, as of 1st March, 2010, are:

- 198,721,699 Species occurrences for all kingdoms
- 51,572,239 Plantae occurrence records
- 39,184,950 Plantae occurrence records with a geo-reference
- 31,994,765 Plantae occurrence records with geo/temporal references
- 579,946 Plantae species with 1 or more occurrence records
- 186 Institutions contributing Plantae data to GBIF

In a recent European plant biodiversity study, although 1,350 species were modelled they represented only a fraction of the number of species to be found there [9]. To obtain a more accurate map of European plant diversity will therefore require modelling the highest possible number of species. GBIF offers various ways of downloading Plantae data including by country and by geographic extent. However, as there are limits on the maximum file size of downloads (up to 250,000 occurrence records, depending on file format selected), several such downloads maybe required which must then be consolidated into a single file. The consolidated csv file for a European plant diversity study will consist of several million records. This is too large to be opened by standard spreadsheet packages. In fact, due to the need to filter out several types of data error, it makes sense to develop custom programs in Java ® [10], for example, to do this. Another motive for automating the extraction of species occurrence datasets from the consolidated file is that preparing these datasets manually takes at least one hour per species based on our experience. Lets say that we wish to do a biodiversity study modelling 6,000 species. To prepare species occurrence datasets manually would take one person around three years whereas automating the process will take around 200 hours to develop the custom programs and, at most, another few hours of computing time to execute them. Earlier, we mentioned the need to filter out data error. In this paper, we will describe these errors and stipulate requirements for custom species occurrence data extraction tools that will, to the extent possible, pre-process the raw data filtering out these errors before preparing species occurrence point data files that can be 'plugged-in' directly to an ENM.

2 Taxonomic Disambiguation

The taxonomic structure for GBIF is the 2007 copy of the Catalogue of Life a database updated annually [11]. The catalogue has grown by quarter of a million species since 2007 but is not complete, nor can it ever be, due to the thousands of new names published each year. However the coverage is close to 66% at species level for all named life and much higher at genus level. The Global Strategy for Plant Conservation [12] has, as its first target: A widely accessible working list of known plant species, as a step towards a complete world flora to be established by 2010. This shows the importance placed on the completion of a catalogue of life. Construction of such a catalogue offers many challenges, not least the lack of funding to support taxonomic expertise to populate the catalogue. Even if the catalogue was complete there are complexities to the application of scientific names to living things. As well as the routine synonymy of names that form accepted species within a kingdom, there are separate codes of nomenclature for plants [13], animals [14] and microbes [15] and the consequence is that organisms in different kingdoms can have the same valid name providing they are covered by different codes. This can lead to problems of synonymy (the use of the same name more than once) in global databases of life because the component databases, if they are purely for plants, animals or microbes, may well not include information on Kingdom. Resolving these conflicts and ensuring names are applied to the right things is called disambiguation. Without this process the result is that erroneous records can be gathered in an automated search of a data portal. There are many examples of such synonymy causing problems. A study by Culham & Yesson (in press) [16] showed that automated retrieval of data for a family of tropical timber and fruit trees, Ebenaceae, yielded 21,000 data points of which 11,000 were in the Atlantic ocean! This was because the genus *Paralia* is used both in the Ebenaceae and for a genus of phytoplankton [16]. Other such examples are cited by Page [17, 18] and Chavan et al [19]. A second major issue is the inconsistent use of valid names as accepted names or as synonyms. The broad bean is widely referred to as *Vicia faba* L. but in some parts of the world is known as *Faba vulgaris* Moench., an equally valid name that is not widely accepted. The use of taxonomically intelligent network services [20] may help automate the process of identification of inconsistent use of names so directing taxonomic expertise to the solving of these problems. Currently both a single scientific name can validly apply to as many as three different organisms (Plant, Microbe, Animal) and multiple scientific names can apply to a single species (through the application of different taxonomic systems). Such problems of ambiguity may reduce when the Catalogue of Life becomes more complete by providing a single reference point. Manual scrutiny of data by the CoL editorial board is needed to prevent the same taxon being placed under different names if it is found in different source databases.

GBIF has a facility for filtering species occurrence data by Kingdom and for returning species synonyms [8]. However, this facility is only presently available for single species occurrence data downloads.

3 Excluding Extraneous Data Fields and Non Plantae Records

Species occurrence data files must be prepared in the precise format required by the chosen ENM. Typically, these are csv or txt files, that at a minimum, include the species name (including genus) followed by longitude and latitude in decimal degrees. Consequently, extraneous data fields must be removed from the consolidated GBIF csv file. We also recommend carrying out a simple check that each record belongs to the Kingdom Plantae and removal of any records that don't. After indexing records alphabetically, identically spelled species' names but with differences in lower/upper case lettering should be identified and harmonized to prevent them being treated by an ENM as separate species.

4 Temporal Error

Ideally, species occurrence observations should fall within the time period covered by the environmental layers. In practice, doing this will eliminate a large number of observations, possibly even the majority of them. In some cases, georeferenced GBIF data has no temporal reference. However, where the observation date is known, there are records ranging from the early 19^{th} Century to date. Filtering out records falling outside the environmental temporal framework is a straightforward task but there is a trade-off. If by excluding these records, we no longer have the minimum number needed to model the species (see Section 6), it may be preferable to maximize the number of species modelled even if doing so means introducing temporal error. Also, we may be excluding occurrences where in fact the species was still present during the time period of our model.

5 Data Duplication

Data duplication commonly occurs when the same occurrence records occur in different datasets, which could happen where two or more herbaria hold duplicates

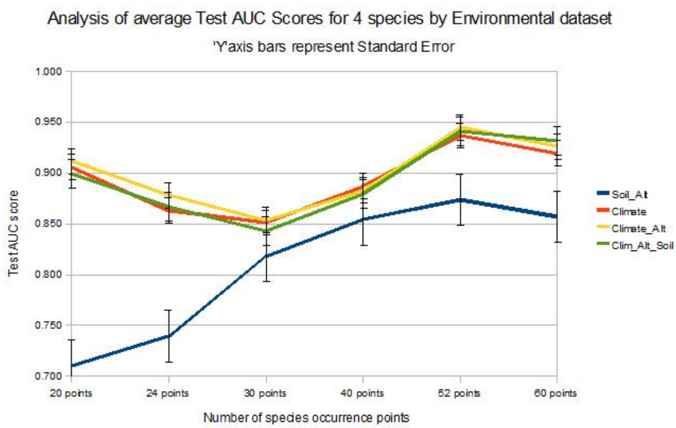


Fig. 1. Test AUC score variation by number of occurrences for 4 environmental datasets

of the same collection, or because the same occurrence was observed at different times. Although, ENM's typically remove these duplications (e.g [5, 21]), this may reduce the number of records below the minimum threshold for an accurate species ENM. Consequently, there is a need to remove these duplications during species occurrence data pre-processing to establish whether modelling a species is practicable. In a survey of 544 mainland European plant species listed under the Berne Convention, 169 (31%) had no georeferenced data in GBIF and only 69 (13%) had 52 or more records [22] before filtering for duplication.

6 Minimizing Test AUC Score Variation

In practice, Test AUC scores will vary by species. This is because it is unlikely that more than one species will have exactly the same occurrence points. While we cannot control AUC Test score variation, we can minimize variances due to differences in dataset size. In a study of 4 plant species occurring in Italy, we analyzed Test AUC scores by different sized occurrence datasets (Figure 1). The results show that 52 point occurrence datasets performed best overall (regardless of environmental dataset type). Models with more than 60 occurrence points were not possible due to the lack of GBIF data for the 4 selected species at the chosen geographic extent. Our results are somewhat consistent with results obtained by other researchers who have observed that Test AUC scores achieved with 50 occurrence points tend to plateau at this level when larger numbers of occurrence points are considered [23, 24, 25]. Interestingly though, a recent American study of spotted knapweed, with species datasets averaging several hundred points saw test AUC score values ranging from 0.65 to 0.75 [26], indicating that the predictive ability of ENM's can also deteriorate when presented with large occurrence point datasets.

While no correlation was observed between test AUC score and species occurrence dataset size for smaller datasets, we believe that the score for 20 point datasets was excessively high due to the small number of test points used (4 or 5). Elith observes that Test AUC scores exceeding 0.75 are useful [27] while Baldwin adds that scores over 0.9 are very good [24]. Although, Hernandez [25] and Pearson [28] suggest that useful scores can be observed with sample sizes below 10, we recommend a minimum of 20. Our reason for this is that the smaller the number of test points, the less evidence there is that the probability distribution can be relied upon and where these test points are clustered (see Section 9), potential reliability diminishes further. For a biodiversity study, there is an obvious need to maximize the number of species being modelled. While it makes sense to minimize Test AUC score variance, due to different sized point datasets, we also need to maximize test AUC scores. Ultimately though, the choice of dataset size will depend on data availability.

A final observation here concerns the most appropriate mix of environmental layers. Figure 1 suggests that the best Test AUC results were obtained for the Climate_Altitude and Climate_Altitude_Soil environmental datasets. The 19 Climate layers are BIOCLIM [29] layers derived from average WorldClim [30] climate data for the period 1950-2000. These layers together with the altitude layer are highly correlated. In contrast, the 4 soil layers [31] included in the Climate_Altitude_Soil environmental dataset were uncorrelated. At 1km spatial resolution, it is likely that soil is a factor in a species probability distribution. Therefore, we recommend providing the ENM with the widest possible range of environmental data believed to be acting at the chosen spatial resolution provided this results in decent Test AUC scores. So, in the case of the experiments summarized in Figure 1, our choice for the best (52 occurrence point) model is that built with Climate Altitude Soil even though its 0.93 average AUC Test score was not significantly different from that obtained with Climate layers alone.

7 Spatial Precision

The spatial precision of geo-referenced records is important because we need to ensure that this is not less precise than the spatial resolution of the environmental layers we use. Generally, the spatial precision of GBIF records varies in the range of 0 to 5 decimal places. For the sake of argument, lets say that we wish to conduct a biodiversity richness study at 30 Arc Seconds of spatial resolution (917 meters at the equator). Species occurrence point data stated to 2 decimal places equates to 1,110 metres at the equator. Therefore, only species occurrence records stated to 3 or more decimal places will be free of spatial precision error. At 2 decimal places, there is a small amount of spatial precision uncertainty but at 0 decimal places we are faced with not knowing which of 14,652 grid cells ($111{,}000^2/917^2$) in each of the environmental layers we should match the occurrence to? We use the qualifier 'possibly' because we may not be able to rely upon the positional accuracy of points whose coordinates are stated to a precision which is less than the spatial resolution of the environmental layers (some

GBIF data is very old and may not have been accurately recorded). Another source of spatial precision uncertainty occurs when occurrence coordinates were originally derived from a raster dataset using grid cell centroid values [32]. For example, some GBIF occurrence data is stated to 5 decimal places but came from 10km x 10km grid cell rasters! To resolve this problem, we must establish the spatial resolution of the original raster by consulting its metadata or, if necessary, the owner of the dataset. GBIF data also includes data where each coordinate is stated to a different number of decimal places. Consequently, for our study at approx. 1km resolution, we may wish to accept the minor spatial error of coordinates stated to 2 decimal places and we may also wish to include records where only one coordinate has been stated to 2 decimal places in the interests of preserving as many records as possible. Therefore, this filter should allow the user to specify the minimum number of decimal places of either both coordinates or one coordinate.

8 Points Missing Environmental Data

Typically, ENM's (e.g [5, 21]) automatically exclude from the model, occurrence points for which any environmental data is missing. If we wish to control the size of datasets and minimize Test AUC score variation due to this factor, then it would pay to filter out these points first rather than have the ENM do this. The first category of points to eliminate are those falling outside the geographic extent of the environmental layers. This is a simple operation. The second category concerns points falling within the geographic extent of environmental layers. Here, two common reasons why occurrence points maybe missing environmental data are; coastal data and recording error [33]. In the first instance, errors occur when the observation was recorded to an insufficient number of decimal places and, as a result, appears in the sea, just off the coastline. Frequent causes of recording error include transposition of longitude/latitude values, and missing coordinates but they also include unexplainable errors. This filter will, therefore, need to find the appropriate grid cells in the environmental layers for each occurrence point and remove records where any *null* or *no data* values are returned.

9 Species Dataset Preparation

In Figure 1, generally the highest Test AUC scores were realized with 52 point datasets. It is interesting, therefore, to contrast, probability distributions for the bilberry (*Vaccinium myrtillus* L.) in the Italian region with 20 and 52 point datasets (Figure 2).

Note, in Figure 2 on the left, how the Slovenian probability distribution is almost non existent and the Alpine distribution of limited extent compared to the right image. Adding more occurrence points (white dots) in the right image, obviously increases the range of environmental data values available to the ENM. However, Test AUC scores were 0.967 for the 20 point dataset and 0.95 for the 52 point dataset showing that these scores take on more importance when steps

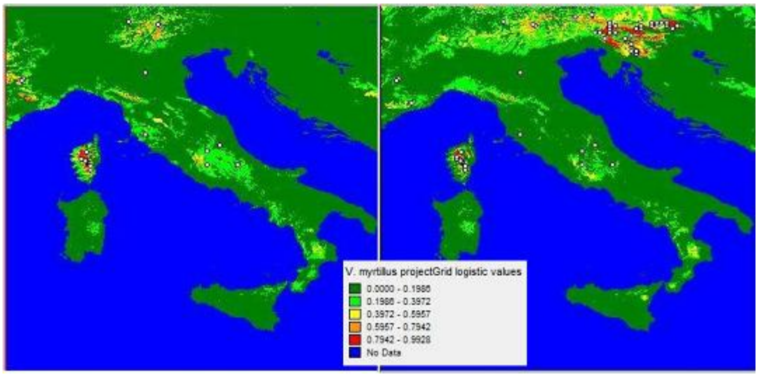


Fig. 2. *Vaccinium myrtillus* MAXENT logistic probability distribution Clim_Alt_Soil (20 points left, 52 points right)

have been taken to maximize the range of environmental data values provided to the ENM. GBIF Plantae species occurrence data is frequently clumped. This is almost certainly the result of the limited geographic extent of field trips to collect this data. We may be seeing this in Figure 2 and, of course, there are endemic plant species where we would expect to see clumped distributions. In his 2009 study of spatially autocorrelated sampling, Veloz concludes that the "AUC statistic is very sensitive to spatial autocorrelation between training and test points"[26]. ENM's attempt to overcome this issue by randomly allocating points between training and test data sets. In Section 6, we saw how generally a species occurrence data file size of 50 is probably the lowest file size choice if Test AUC scores are to be maximized. When more that 50 data points are available for a species from GBIF, it would make sense, therefore, to pick those points that provide the greatest geographic spread to avoid the spatial autocorrellation problem. Pythagoras's theorem [34] can help us here in calculating the distance d between two points (s_i,s_j):

$$d(s_i, s_j) = \sqrt{}(x_i - x_j)^2 + (y_i - y_j)^2 \quad (1)$$

If we calculate the distance between each species occurrence point and every other point, then we can choose the 50 points showing the widest dispersion. Finally, after filtering out data uncertainty, species occurrence file generation is easily automated with Java ® which includes library classes for the production of species occurrence point datasets in csv or txt format as used by MAXENT and openModeller, for example.

10 Conclusions

Divorcing the pre-processing of species occurrence data (e.g. removing duplicate occurrences and points with missing environmental data) from the ENM, gives the experimenter greater control over the size of species datasets and thus the ability to minimize test AUC score variance due to this factor. Including a widest dispersion filter for species with large numbers of occurrences similarly allows control over dataset size while providing the ENM with a wider range of environmental data values than that generated simply by random selection. The ability to filter out occurrences with coordinate precision lower than that required by the model's spatial resolution is an important pre-processing option. A separate species occurrence data pre-processing phase is also an opportunity to carryout temporal and synonymity filtering. After data filtering, the final task of the pre-processing phase is the automatic generation of species occurrence datasets which can then be directly 'plugged-in' to the ENM. A software application capable of carrying out all these tasks will prove to be a valuable time-saver particularly for biodiversity studies involving thousands of species. We are currently developing a software application implementing the species occurrence data pre-processing requirements described in this paper and we plan to present a tested version of this application at this special session of the KES 2010 conference.

References

[1] University of Idaho, college of Natural Resources, http://www.cnr.uidaho.edu/veg_measure/Modules/Lessons/Module%207/7_2_Biodiversity.htm
[2] Gurevitch, J., Scheiner, S.M., Fox, G.A.: The Ecology of Plants, 2nd edn. Sinauer Associates, Inc., Sunderland (2006)
[3] Gaston, K.J.: Species richness: Measure and measurement. In: Gaston, K.J. (ed.) Biodiversity, a Biology of Numbers and Difference, vol. 396, pp. 77–113. Blackwell Science, Cambridge (1996)
[4] Escalante, T., et al.: Ecological niche models and patterns of richness and endemism of the southern Andean genus Eurymetopum (Coleoptera, Cleridae). Rev. Bras. Entomol. 53(3), 379–385 (2009)
[5] The OpenNLP Maxent Homepage, http://maxent.sourceforge.net/
[6] Fielding, A.H., Bell, J.F.: A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24, 38–49 (1997)
[7] Phillips, S.J., Anderson, R.P., Schapire, R.E.: Maximum entropy modeling of species geographic distributions. Ecol. Model. 190, 231–259 (2006)
[8] GBIF Data Portal, 01/03/2010, http://data.gbif.org
[9] Thuiller, W., et al.: Climate change threats to plant diversity in Europe. Proc. Nat. Acad. Sci. USA 102, 8245–8250 (2005)
[10] Oracle and Java |Technologies, http://www.oracle.com/us/technologies/java/index.html
[11] Bisby, F.A., Roskov, Y.R., Orrell, T.M., Nicolson, D., Paglinawan, L.E., et al. (eds.): Species 2000 & ITIS Catalogue of Life: 2009 Annual Checklist Taxonomic Classification. Species 2000, Reading, UK (2009), CD-ROM
[12] Johnson, B.: Understanding the Global Strategy for Plant Conservation. In: Annual meeting of the North American Association For Environmental Education (2005)

[13] McNeill, J., Barrie, F.R., Burdet, H.M., Demoulin, V., Hawksworth, D.J., et al.: International Code of Botanical Nomenclature (Vienna Code) adopted by the Seventh International Botanical Congress Vienna, Austria, July 2005, p. 568. A.R.G. Gantner Velag, Ruggell (2006)
[14] International Commission on Zoological Nomenclature, International Code of Zoological Nomenclature, 4th edn. The Int. Trust for Zoological Nomenclature (1999)
[15] Sneath, P.H.A.: International Code of Nomenclature of Bacteria: Bacteriological Code, 1990 Revision. ASM Press, Washington (1992)
[16] Culham, A., Yesson, C.: Biodiversity informatics for climate change studies. In: Hodkinson, et al. (eds.) Climate Change and Systematics, Systematics Association, UK, pp. pp. XXX–YYY (2010)
[17] Page, R.D.M.: A taxonomic search engine: Federating taxonomic databases using web services. BMC Bioinformatics 6, 48 (2005)
[18] Page, R.D.M.: Taxonomic Names, Metadata, And The Semantic Web. Biodiversity Informatics 3, 1–15 (2006)
[19] Chavan, V., Rane, N., Watve, A.: Resolving Taxonomic Discrepancies: Role Of Electronic Catalogues Of Known Organisms. Biodiversity Informatics 2, 70–78 (2005)
[20] Patterson, D.J., Remsen, D., Marino, W.A., Norton, C.: Taxonomic Indexing Extending the Role of Taxonomy Systematic. Biology 55(3), 367–373 (2006)
[21] openModeller, `http://openmodeller.sourceforge.net/`
[22] Heywood, V.H., Culham, A.: The Impacts Of Climate Change On Plant Specie. In: Europe. Report to The Convention On The Conservation Of European Wildlife And Natural Habitats, Standing Committee 29th Meeting Bern (November 09, 2010)
[23] Stockwell, D.R.B., Peterson, A.T.: Effects of sample size on accuracy of species distribution models. Ecol. Model. 148, 1–13 (2002)
[24] Baldwin, R.A.: Use of Maximum Entropy Modeling in Wildlife Research. Entropy 11, 854–866 (2009)
[25] Hernandez, P.A., Graham, C.H., Master, L.L., Albert, D.L.: The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29, 773–785 (2006)
[26] Veloz, S.D.: Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models. J. Biogeogr. 36, 2290–2299 (2009)
[27] Elith, J.: Quantitative methods for modeling species habitat: comparative performance and an application to Australian plants. In: Ferson, S., Burgman, M. (eds.) Quantitative Methods for Conservation Biology, pp. 39–58. Springer, Heidelberg (2002)
[28] Pearson, R.G., Raxworthy, C.J., Nakamura, M., Townsend-Peterson, A.: Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. Journal of Biogeopraphy 34, 102–117 (2007)
[29] Bioclim — WorldClim - Global Climate Data, `http://www.worldclim.org/bioclim`
[30] WorldClim - Global Climate Data, `http://www.worldclim.org/`
[31] European Soil Data Center (ESDAC), `http://eusoils.jrc.ec.europa.eu/library/esdac/index.html`
[32] Biodivertido.blogspot.com, `http://biodivertido.blogspot.com/2009/02/grid-data-shared-as-point-data-errors.html`
[33] Yesson, C., Brewer, P.W., Sutton, T., Caithness, N., Pahwa, J.S., et al.: How Global Is the Global Biodiversity Information Facility? PLoS ONE 2(11), e1124 (2007), doi:10.1371/journal.pone.0001124
[34] O'Sullivan, D., Unwin, D.J.: Geographic Information Analysis. Wiley, USA (2003)

A Hybrid Approach for Indexing and Retrieval of Archaeological Textual Information

Ammar Halabi[1], Ahmed-Derar Islim[2], and Mohamed-Zakaria Kurdi[1,3]

[1] School of Informatics and Computing, Indiana University, Bloomington
ammar.halabi@gmail.com
[2] Financial Mathematics Department, Florida State University, Tallahassee
derarief@gmail.com
[3] Department of Computer Science, Mamoun University of Science and Technology, Aleppo, Syria
mzkurdi@yahoo.com

Abstract. This paper focuses on the problem of archaeological textual information retrieval, covering various field-related topics, and investigating different issues related to special characteristics of Arabic.

The suggested hybrid retrieval approach employs various clustering and classification methods that enhances both retrieval and presentation, and infers further information from the results returned by a primary retrieval engine, which, in turn, uses Latent Semantic Analysis (LSA) as a primary retrieval method. In addition, a stemmer for Arabic words was designed and implemented to facilitate the indexing process and to enhance the quality of retrieval.

The performance of our module was measured by carrying out experiments using standard datasets, where the system showed promising results with many possibilities for future research and further development.

Keywords: Information retrieval, Arabic Information Retrieval, Arabic Stemming, Arabic Lexical Analysis, Latent Semantic Analysis, Automatic Document Categorization.

1 Introduction

Today, with new archaeological sites being discovered and interesting findings being unearthed in already established ones, growing archaeological data places a significant amount of information at the archaeological community's disposal by every mission at the end of a successful season.

In principle, archaeologists and researchers in associated disciplines should be able to access this information in a convenient and consistent manner to effectively retrieve material for the support of their own research, and to conduct collaborative research, via information exchange, with other researchers in the community or in other research communities.

The needs of information recording, organization, acquisition and dissemination in the archaeological community suggest interesting possibilities for the adoption of

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 527–535, 2010.

computer-based information systems. We started our research with the goal of designing and implementing a robust Archaeological Information Retrieval system, which is capable of indexing and searching cross-language corpora as well as cross-media corpora. This information retrieval system is a basic need and a critical component in a complete Archaeological Information System, and can be considered as the starting point for developing such system [10].

The main differences between such an archaeological system and other common information retrieval engines are:

- Limitation of the domain: in domain oriented applications, the lexicon may be big but it is usually limited. This leads to a significant reduction of the textual ambiguity.
- Multilingualism: the archaeological data may be in several languages (especially old ones such as Acadian, Sumerian, Aramean, Assyrian, etc.)

In this paper, we tackle the problem of the retrieval of textual information archaeology. We introduce a background on textual information retrieval systems, followed by proposing architecture for the Archaeological Textual Information Retrieval System, and we end by presenting our results and conclusions. The current application covers Arabic language only but the adopted design makes it relatively easy to add new languages: only a lightweight stemmer need to be added per language.

2 State of the Art

2.1 Latent Semantic Analysis (LSA)

Latent Semantic Analysis (LSA) [5] is a technique used in statistics and natural language processing to find hidden – or latent – relations between a set of observations and a set of associated features. This is done by mapping features and observations onto an intermediate concept space which only preserves the most significant characteristics. By this new representation, relations that were unobvious between features and observations can be revealed. In the context of natural language, features are represented by terms, and observations are represented by documents. LSA can be applied for document retrieval by projecting user queries and indexed documents onto the concept space to uncover the relation between the user needs and documents in the corpus. The application of LSA in textual information retrieval is known as Latent Semantic Indexing (LSI).

In theory, to find a lower rank approximation of the 2D term-document matrix, LSA makes use of the reduced Singular Value Decomposition (SVD); a matrix factorization tool used in signal processing and statistics [1].

Using SVD, the term-document matrix X (where row vectors represent terms, and column vectors represent documents) is written as the product of three orthonormal matrices, U, Σ, and V, as shown in Fig. 2

As in Fig. 2, U holds the eigenvectors of the matrix $X.X^T$, V holds the eigenvectors of the matrix $X^T.X$, and Σ is a diagonal matrix having it's diagonal formed by the square roots of the eigenvalues of the matrix $X.X^T$ (or equally by the square roots of

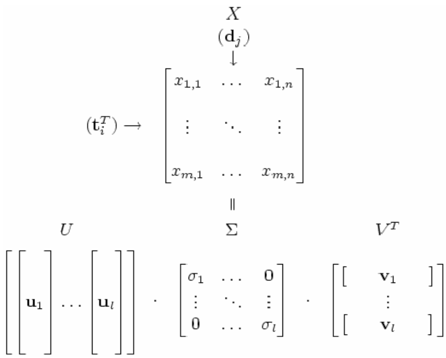


Fig. 1. SVD of the term-document matrix (X)

the eigenvalues of the matrix $X^T.X$). The values $\sigma_1, \ldots, \sigma_l$ are the singular values, where $u_1, \ldots, u_l$ and $v_1, \ldots, v_l$ the left and right singular vectors.

The LSA concept space is extracted by keeping only the *k* largest singular values and the corresponding *k* left and right singular vectors. These *k* singular values represent the new reduced concept space, where the left and right singular vectors provide the means to transform to and from this space.

After computing the SVD of the term-document matrix, relations between terms, documents, or terms and documents can be revealed. Moreover, user query vectors can be projected on the concept space to retrieve the closest set of documents that meet the user's needs.

LSI can be applied for document indexing and retrieval including cross-lingual corpora, and for modeling the process of human learning and text comprehension [13], [5]. LSI is also reported to outperform the *Vector Space* retrieval model, where comparisons between query vectors and document vectors are done in the original space [14].

2.2 Automatic Document Clustering and Classification

Document clustering can be defined as the partitioning of a dataset of documents into subsets (clusters), where all documents in each of these subsets share some common traits expressed using some certain distance measure. This technique is an unsupervised learning method that no prior information related to potential similarities between documents is used in the learning process [2].

Document classification is a supervised machine learning method which assigns documents to pre-defined labels (categories). This technique is used in pattern recognition and data analysis. It works by constructing a classifier which learns a model from a training dataset composed of documents along with their corresponding categories. This dataset has to contain enough information for the classifier model to be effective at predicting the classes of new documents.

3 The Archaeological Text Retrieval System

In Fig. 3, we illustrate our proposed hybrid architecture for The Archaeological Text Retrieval System.

This architecture was primarily based on the general architecture and conventions used in text retrieval systems, with an additional layer of functionality inspired by the work of Sahami [22].

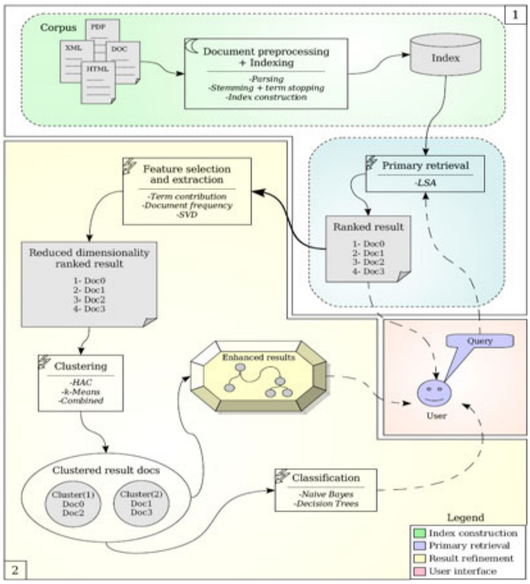


Fig. 2. Generic architecture of IR systems; the crescent refers to offline operations; lighting bolts refer to online operations

For illustrative purposes, the architecture can be divided, into two main blocks. The first block (1) is modeled after the general architecture of text retrieval systems which provides primary retrieval of documents. This is achieved by:

- Document pre-processing, which includes parsing files of different formats, term stopping, and stemming.
- Construction of inverted index.
- Utilizing a primary retrieval engine based on the well defined LSI retrieval model.

The second block (2) adds extra levels of functionality in order to refine the results of primary retrieval, i.e. set of retrieved documents, obtained from the first block and also to improve the quality of result demonstration, where documents are automatically clustered and classified into different categories giving the user a better perception of the retrieval result than in the case of simply presenting a list of retrieved documents. This is achieved by:

- Feature selection for primarily retrieved documents.
- Clustering of primarily retrieved documents.
- Classification of new documents into one of the learned categories.

By this architecture, we seek to improve the quality of retrieval and result demonstration of the overall system. In another words, the features provided by the Archaeological Text Retrieval subsystem are:

- Primary, semi-semantic text retrieval.
- Enhanced retrieval results automatically refined using clustering and classification techniques.

The problem of archaeological text retrieval is similar to the problem of general text retrieval where only minimal modifications of the general text retrieval system are required to adapt for retrieval in a specific field, i.e. special stop lists, and additional stemming rules. Therefore, we focused on the problem from a general point of view assuring that selected methods can be extended to build a cross-media retrieval system.

3.1 Document Pre-processing

Concerning Arabic in particular, we came out with our own stop-word list which includes functional vocabulary like prepositions, adverbs, pronouns, and others.

We also implemented a Light Stemming algorithm which strips frequently used suffices and postfixes of Arabic words. Light stemming of Arabic terms is reported to contribute to the effectiveness of retrieval better than Root Normalization, which adopts a more aggressive stemming approach by reducing words to their roots [15], [16]. Our algorithm is similar in concept to the stemming algorithm described in [16] and [4]. However, we suggested and implemented a different set of stemming rules, making more use of the knowledge of Arabic morphology.

3.2 Indexing

In this operation, the index is constructed and the term-document matrix is built, which serves as an abstract representation of documents for the retrieval model to act upon.

3.3 Primary Text Retrieval

Latent Semantic Indexing (LSI) was chosen as a primary retrieval model for the following reasons:

- LSI generally outperforms the vector space model and provides solutions to the problems of synonymy [3], [5].
- An LSI-based retrieval engine can be extended to achieve cross language information retrieval [19], [20].
- LSI uses the reduced SVD decomposition to project document and query vectors on a new dimensionally-reduced space. This new representation of documents can also be used for document clustering to refine primary search results [17], [24].

This LSA-based retrieval engine provides primary retrieval of relevant documents to be further refined by further processing of the retrieval result.

3.4 Document Clustering

Both HAC and k-Means clustering algorithms were used in document clustering. The output of HAC was used to seed the K-Means algorithm. By this approach, after documents are grouped under different levels of hierarchy using HAC, one level of the output hierarchy is used to seed K-Means which refines clustering at the given level by making useful re-assignments of documents into clusters. In addition, using HAC to seed K-Means can yield improvements in performance, where K-Means will potentially converge faster than in the case of randomized document seeding [22]. Upon clustering, cluster descriptors are extracted, which are the most representative terms of the documents contained in the corresponding cluster. They effectively assist users to understand the categories of clustered retrieval results. These descriptors are extracted using the Probabilistic Odds method [22].

3.5 Document Classification

A Naïve Bayesian classifier [18] was employed for the classification of user queries and new documents. Naïve Bayesian classifiers have yielded good results in text classification [6], [22], and they have been applied successfully [23].

After primary retrieval results are clustered, resultant clusters serve as a training set to train the classifier, so as new documents or queries presented to the classifier when the system is online will be classified as belonging to one of the learned classes.

This allows users to accurately determine the most a cluster which is most relative to their search queries or example documents.

4 Results

By using our light Arabic stemmer and refining primarily retrieved documents, we obtained promising results where these techniques proved efficient and effective in Arabic textual information retrieval.

Fig. 4 compares the effects of our stemming technique to the light10 stemmer [16] on the performance of document clustering using the k-means algorithm with document features projected on the reduced space of SVD. This experiment was performed using the Sulaiti dataset [25], which was assembled from newspapers, magazines, radio, TV and webpages, summing to 411 texts that are manually classified under 8 different categories. The quality of document clustering was measured using the F-measure:

$$F = \frac{2 \times precision \times recall}{precision + recall}$$

These results show that our stemmer matches the light10 stemmer for effectiveness regarding the document clustering. Moreover, by looking at Table I, It is shown by stemming the documents of three datasets that our proposed stemmer does a better job

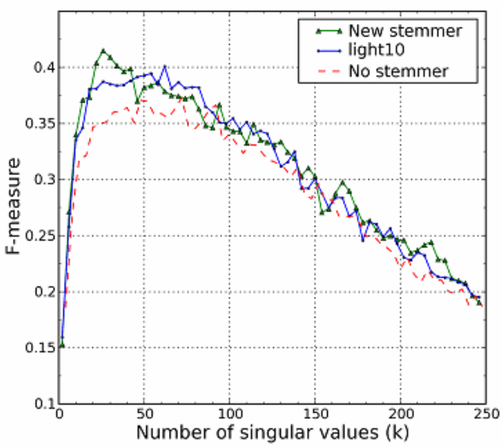


Fig. 3. Document clustering performance as a function of (*k*)

Table 1. Number of Terms Extracted from Arabic Datasets Using Different Stemmers

Dataset	Number of Terms		
	No Stemmer	Light10	New Stemmer
Sulaiti	108209	44329	31198
Hammadeh	22921	10197	7980
Mashkat	656807	222478	135369

in reducing the number of features (terms) to be used in indexing and retrieval. This considerably improves the performance of computationally-demanding retrieval systems when stemming is applied. The datasets used include the Sulaiti dataset [25] in addition to two datasets we assembled from researchers in archaeology and an online resource for ancient Arabic texts.

Regarding the performance of our proposed retrieval model, we were short of free Arabic datasets that are tailored for evaluating retrieval engines. Therefore, we were not able to conduct numerical experiments to measure the quality of our Arabic retrieval engine. However, empirical results and user reactions indicated that retrieval results of the primary LSI retrieval engine on indexed Arabic corpora were good. In addition, further clustering and classification operations successfully improved the quality of presentation for the primarily retrieved group of documents, and successfully assisted users reaching required information more rapidly.

5 Conclusions

In this paper, we have demonstrated our work in Arabic information retrieval and viewed the architecture of our Archaeological Text Retrieval System. This architecture was designed after the generic retrieval systems' architecture with an additional layer of functionality that improves presentation of retrieval results. Our results showed that our stemming algorithm is highly effective, and that statistical and probabilistic methods for retrieval and language modeling such as the LSI, automatic document clustering, and classification are effective for Arabic textual information.

References

1. Akritas, G., Malaschonok, G.I.: Applications of Singular-Value Decomposition. Mathematics and Computers in Simulation 67(1-2), 15–31 (2004)
2. Berkhin, P.: Survey of clustering data mining techniques. Tech. Rep., Accrue Software, San Jose, CA (2002)
3. Berry, M.W., Dumais, S.T., O'Brien, G.W.: Using Linear Algebra for Intelligent Information Retrieval. SIAM Review 37(4), 573–595 (1995)
4. Chen, F.G.: Building an Arabic Stemmer for Information Retrieval. In: Proc. Eleventh Text Retrieval Conference TREC 2002, Gaithersburg, Maryland, USA, pp. 19–22 (2002)
5. Deerwester, S., Dumais, S., Furnas, G.W., Landauer, T.K., Harshman, R.: Indexing by Latent Semantic Analysis. Journal of the Society for Information Science 41(6), 391–407 (1990)
6. Dumais, S., Platt, J., Heckerman, D., Sahami, M.: Inductive learning algorithms and representations for text categorization. In: 7th ACM International Conference on Information and Knowledge Management ACM-CIKM 1998, Bethesda, USA, pp. 148–155 (1998)
7. Fox: Lexical Analysis and Stoplists. In: Frakes, W.B., Baeza-Yates, R. (eds.) Information Retrieval: Data Structures and Algorithms. Prentice Hall, Englewood Cliffs (1992)
8. Frakes, W.B.: Stemming Algorithms. In: Frakes, W.B., Baeza-Yates, R. (eds.) Information Retrieval: Data Structures. Prentice Hall, Englewood Cliffs (1992)
9. Frakes, B., Baeza-Yates, R.: Information Retrieval: Data Structures and Algorithms. Prentice-Hall, Englewood Cliffs (1992)
10. Halabi, A.D.I., Keshishian, R., Rehawi, O.: The Archaeological Text Retrieval System. BSc. thesis, Dept. Artificial Intelligence, Faculty of Informatics, University of Aleppo (2007)
11. Hearst, M.A., Pedersen, J.O.: Reexamining the Cluster Hypothesis: Scatter/Gather on Retrieval Results. In: Proc. 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 1996), Zurich, Switzerland, June 1996, pp. 76–84 (1996)
12. Hull: Stemming algorithms – A case study for detailed evaluation. Journal of the American Society for Information Science 47(1), 70–84 (1996)
13. Landauer, T.K., Foltz, P.W., Laham, D.: Introduction to Latent Semantic Analysis. Discourse Processes 25, 259–284 (1998)
14. Landauer, T.K., Littman, M.L.: A statistical method for language-independent representation of the topical content of text segments. In: Proc. Eleventh International Conference: Expert Systems and Their Applications, Avignon, France, vol. 8, pp. 77–85 (May 1991)
15. Larkey, L., Ballesteros, L., Connell, M.E.: Improving stemming for Arabic information retrieval: light stemming and co-occurrence analysis. In: Proc. 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Tampere, Finland, pp. 275–282 (2002)
16. Larkey, L., Ballesteros, L., Connell, M.: Light Stemming for Arabic Information Retrieval. In: Soudi, A., van den Bosch, A., Neumann, G. (eds.) Arabic Computational Morphology: Knowledge-based and Empirical Methods. Series on Text, Speech, and Language Technology. Kluwer/Springer's (2005)

17. Lerman, K.: Document Clustering in Reduced Dimension Vector Space (1999) (unpublished), http://www.isi.edu/~lerman/papers/papers.html (retrieved on 13-08-2007)
18. Lewis, D.D.: Naive Bayes at forty: The independence assumption in information retrieval. In: Nédellec, C., Rouveirol, C. (eds.) ECML 1998. LNCS, vol. 1398, pp. 4–15. Springer, Heidelberg (1998)
19. Littman, M.L., Dumais, S.T., Landauer, T.K.: Automatic cross-language information retrieval using latent semantic indexing. In: Grefenstette, G. (ed.) Cross-Language Information Retrieval, pp. 51–62. Kluwer Academic Publishers, Dordrecht (1998)
20. Littman, M.L., Jiang, F.: A Comparison of Two Corpus-Based Methods for Translingual Information Retrieval. Tech. Rep. CS-98-11, Duke University, Department of Computer Science, Durham, NC (June 1998)
21. Manning, C.D., Raghavan, P., Schütze, H.: Introduction to Information Retrieval, August 13. Cambridge University Press, Cambridge (2007), http://www-csli.stanford.edu/~schuetze/information-retrieval-book.html
22. Sahami, M.: Using Machine Learning to Improve Information Access. Ph.d. thesis, Dept. Computer Science, Stanford University (1999)
23. [Sahami et al 1998] Sahami, M., Dumais, S., Heckerman, D., Horvitz, E.: A Bayesian Approach to Filtering Junk E-mail. In: Proc. AAAI 1998 Workshop on Learning for Text Categorization, Madison, Wisconsin, USA, pp. 55–62 (1998)
24. Schutze, H., Silverstein, C.: Projections for efficient document clustering. In: Proc. 20th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, Philadelphia, Pennsylvania, USA, pp. 74–81 (1997)
25. Al-Sulaiti, L., Atwell, E.: Designing and Developing a Corpus of Contemporary Arabic. In: Proc. Sixth TALC Conference, Granada, Spain, p. 92 (2004)

Prognosis of Breast Cancer Using Genetic Programming

Simone A. Ludwig and Stefanie Roos

Department of Computer Science, University of Saskatchewan, Canada
ludwig@cs.usask.ca

Abstract. Worldwide, breast cancer is the second most common type of cancer after lung cancer and the fifth most common cause of cancer death. In 2004, breast cancer caused 519,000 deaths worldwide. In order to reduce the cancer deaths and thereby increasing the survival rates an automatic approach is necessary to aid physicians in the prognosis of breast cancer. This paper investigates the prognosis of breast cancer using a machine learning approach, in particular genetic programming, whereas earlier work has approached the prognosis using linear programming. The genetic programming method takes a digitized image of a patient and automatically generates the prediction of the time to recur as well as the disease-free survival time. The breast cancer dataset from the University of California Irvine Machine Learning Repository was used for this study. The evaluation shows that the genetic programming approach outperforms the linear programming approach by 33 %.

Keywords: genetic programming, selection, crossover, mutation, breast cancer, recurrent surface approximation.

1 Introduction

Worldwide, breast cancer is the second most common type of cancer after lung cancer (10.4 % of all cancer incidence) and the fifth most common cause of cancer death. In 2004, breast cancer caused 519,000 deaths worldwide (7 % of cancer deaths; almost 1 % of all deaths) [1]. Breast cancer is the most common malignancy in women, except for non-melanoma skin cancers. It continues to be a major health care problem worldwide. Cancer occurs when cells in a part of the body begin to grow out of control. Normal cells divide and grow in an orderly fashion, but cancer cells do not. They continue to grow and crowd out normal cells. Although there are many kinds of cancer, they all have in common this out-of-control growth of cells [2].

Different kinds of cancer can behave very differently. For example, lung cancer and breast cancer are very different diseases. They grow at different rates and respond to different treatments. That is why people with cancer need treatment that is aimed at their kind of cancer. Therefore, it is important to identify the type of cancer accurately, so that the correct treatment can be started.

Breast cancer is a cancer that starts in the tissues of the breast. There are two main types of breast cancer [3]: (1) Ductal carcinoma starts in the tubes (ducts) that move

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 536–545, 2010.

milk from the breast to the nipple. Most breast cancers are of this type. (2) Lobular carcinoma starts in parts of the breast, called lobules that produce milk.

The good news is that early detection and new treatments have improved survival rates of breast cancer. The 5-year survival rate for women diagnosed with cancer is 80 %. About 88 % of women diagnosed with breast cancer will survive at least 10 years. Unfortunately, women in lower social and economic groups still have significantly lower survival rates than women in higher groups. The good news is that women are living longer with breast cancer. Survivors must live with the uncertainties of possible recurrent cancer and some risk for complications from the treatment itself [4].

Recurrences of cancer usually develop within 5 years of treatment. However, 25 % of recurrences and half of new cancers in the opposite breast occur after 5 years.

In order to aid the physicians in the prognosis of whether the breast cancer is likely to recur in patients, the linear programming approach has be introduced in 1994 [7]. It has achieved good results with an expected error of 13.0 to 18.3 months, which was better than the prognosis correctness achieved by other available techniques at that time. However, the proposed approach outlined in this paper, takes another step towards the improvement of the prognosis correctness using a genetic programming approach.

This paper is structured as follows. First, some background information regarding the previous work using the linear programming technique, the data collection and analysis is outlined in Section 2. In Section 3, the proposed approach using genetic programming is introduced outlining the different parameters involved. The experiments and results are given in Section 4. Section 5 concludes this paper with a summary and analysis of the results obtained.

2 Previous Work

The dataset used for the prognosis of breast cancer is publicly available at the UCI Machine Learning Repository [5]. The dataset was collected by researchers from the University of Wisconsin as follows. First, a sample of fluid was taken from the patient's breast. This fluid was then placed on a glass slide and stained to highlight the nuclei of the constituent cells. An image from the FNA is transferred to a workstation by a video camera mounted on a microscope. A program, called Xcyt [6], was developed which uses a curve-fitting program to determine the exact boundaries of the nuclei. The boundaries were initialized by an operator using a mouse pointer. Ten features were computed for each nucleus: area, radius, perimeter, symmetry, number and size of concavities, fractal dimension (of the boundary), compactness, smoothness (local variation of radial segments), and texture (variance of gray levels inside the boundary). The mean value, extreme value (i.e. largest or worst value: biggest size, most irregular shape) and standard error of each of these cellular features were computed for each image. Furthermore, tumor size and lymph node status make up a total of 32 real-valued features. The dataset contains 198 observations; with 47 recur cases and 151 non-recur cases.

Using this dataset, Street et al. [7] applied linear programming to predict the Time To Recur (TTR), which is a mapping of an n-dimensional input of cytological and other features to a one-dimensional time output. One complicating factor for the prediction is that TTR is known for only a subset of patients and not for others for which we know only the time of their last check-up, or Disease-Free Survival time (DFS). However, all available cases were used for this investigation. The solution to this estimation program is termed the Recurrence Surface Approximation (RSA) technique [7,8,9]. RSA basically uses linear programming to determine a linear combination of the input features that accurately predicts TTR. The linear program to be solved for a given training set is given in [8]. The motivation for the RSA approach was the following: (1) Recurrences actually take place at some point in time prior to their detection. However, the difference between the time a recurrence is detectable (actual TTR) and the time it is actually detected (observed TTR) is assumed to be small. (2) Observed DFS is a lower bound on the recurrence time of that patient.

Therefore, three types of absolute errors need to be considered. Let err_1 be the average overestimation using only the set of patients with a recurrence (i.e., the set of recurrent cases). The average underestimation of DFS using only the set of patients without an observed recurrence is denoted with err_2. Since DFS is a lower bound for the TTR, overestimation is not considered an error. Thus, the third type of error, err_3, is the average underestimation of TTR (on the set of patients having a recurrence). Underestimation of TTR is not considered as severe as overestimation, hence the RSA approach tries to minimize:

$$err_1 + err_2 + \delta \cdot err_3 \tag{1}$$

where $0 < \delta \leq 1$ defines the influence of underestimating TTR in relation to overestimating TTR. For the genetic programming approach, Equation (1) is used as the fitness function, thus allowing a direct comparison. Note that it is not sensible for machine learning algorithms to set $\delta = 0$, even considering that the recurrence appeared at some point before it was detected. Allowing $\delta = 0$, the algorithm could simply perform a classification into recurrent and non-recurrent cases. Then 0 could be predicted for all recurrent cases (underestimating TTR in all cases), while some large value (meaning greater than the maximal DFS in the set) is predicted for the non-recurrent cases (always overestimating DFS). However, the goal of this research is to perform regression analysis for medical prognostics, and therefore, symbolic regression and not classification was used.

In [7], δ was chosen such, that of those values for which $err_1 + err_2$ are minimal, the one that minimizes err_3 is chosen. Based on the perturbation theorem [10], such that a δ, $0 < \delta \leq \bar{\delta}$ for some $\bar{\delta}$, exists. For the proposed genetic programming approach, the influence of various δ on the fitness and the expected error is investigated, as well as on err_1, err_2 and err_3 each.

In order to obtain the best generalization, it is important to choose the right subset of features. The appropriate feature set was chosen for the linear programming approach in the following automatic fashion. A tuning test (one tenth of the training cases) was first set aside. The RSA linear program was then solved using all input features, and the resulting surface was tested on the tuning set. Features were then removed one by one. Each new problem was solved and the result tested on the tuning set, until only one feature remained. Using the features that showed the best performance on the tuning set, all the training data was then re-optimized. Feature selection is done differently for the proposed genetic algorithm approach and will be outlined in the following section.

3 Genetic Programming Approach

The origins of evolutionary computation reach back to the 50's of the last century. Genetic programming, in itself, was not considered until the middle of the 80's. The term first appeared in [11], the main development took place in the early and middle 90's, particularly through work by Koza [12].

Genetic programming uses the concepts of genetics and Darwinian natural selection to generate and evolve entire computer programs. Genetic programming largely resembles genetic algorithms in terms of its basic algorithm. The notions of mutation, reproduction (crossover) and fitness are essentially the same, however, genetic programming requires special attention when using those operations. While genetic algorithms are concerned with modifying fixed-length strings, usually associated with parameters to a function, genetic programming is concerned with actually creating and manipulating the (non-fixed length) structure of the program (or function). Therefore, genetic programming is more complex than genetic algorithms [13] and works as follows. In genetic programming the aim is to find solutions to some problem in the form of a computer program. It is a stochastic search strategy that is particularly powerful in the circumstances where one cannot make any assumptions about the characteristics of the solution. The solution is developed by first creating a number of initial programs, which are then recombined and changed in each evolution step [14].

Genetic programming performs the following steps:

```
  Step 1: Assign the maximum number of generations to be run and
probabilities for cloning, crossover and mutation.
  Step 2: Generate an initial population of computer programs of
size N by combining randomly selected functions and terminals.
  Step 3: Execute each computer program in the population and
calculate its fitness with an appropriate fitness function.
Designate the best-so-far individual as the result of the run.
  Step 4: With the assigned probabilities, select a genetic
operator to perform cloning, crossover or mutation.
  Step 5: If cloning operator is chosen, select one computer
program from the current population of programs and copy it into a
new population. If crossover operator is chosen, select a pair of
computer programs from the current population, create a pair of
offspring programs and place them into the new population. If
mutation operator is chosen, select one computer program from the
current population, perform mutation and place the mutant into the
new population.
  Step 6: Repeat Step 4 until the size of the new population of
computer programs becomes equal to the size of the initial
population N.
  Step 7: Replace the current (parent) population with the new
(offspring) population.
  Step 8: Go to Step 3 and repeat the process until the termination
criterion is satisfied.
```

The Java Genetic Algorithms Package (JGAP) [15] was chosen as the programming platform. JGAP is a Genetic Algorithms and Genetic Programming package written in Java. It is designed to require minimum effort to use, but is also designed to be highly modular. It provides basic genetic mechanisms that can be used to apply evolutionary principles to solve problems.

4 Experiments and Results

The experiments were done as follows. First of all, the dataset was pre-processed normalizing the values by subtracting the mean and dividing by the standard deviation. There were 4 missing values for the lymph node status, which were replaced by the mean. For the division of training and test data, the leave-one-out method was used as this sampling method was also used for the linear programming approach. Then, the genetic programming was fine-tuned with the selection of parameters such as function set, feature selection, maximal crossover depth, crossover and mutation rate, population size, and number of generations. Afterwards, the varying δ values were investigated.

In general for the fine-tuning, the initial choice of parameters was: tournament selection of size 4, population size 1,000, minimal initial depth 5, maximal crossover depth 12, crossover rate 0.9, mutation rate 0.3, and a δ value of 0.01. Basic mathematical operations (addition, subtraction, multiplication, division), comparison operators (<,>), if-statement, logarithm and exponential made up the function set. For those parameters that were analysed, the best performing value was used for all following tests. If not otherwise stated, 30 tests were performed for each choice of parameter. Note that the results were compared using the average expected error, not the fitness value. However, the lowest expected error always corresponds to the lowest fitness value. The expected error only considers known errors, meaning the underestimation of DFS and the overestimation of TTR.

Feature selection: 180 test cases over 300 generations were run, and it was counted how often each feature showed up in the best solution of the run. Afterwards test cases each with the 4 and 8 most frequent features (all others were contained in less than 10 %) were run. The best result was achieved when 8 features were used: 13.02 in comparison to 13.62 (4 features). Using all features led to an average error of 13.80. The 8 features are (in the order of the number of times they are used): Tumor size, lymph node status, mean symmetry, extreme area, standard deviation of radius, standard deviation of area, extreme compactness, and extreme concave points. This is quite different to the 5 features found by Street et al. [5]: mean area, mean perimeter, mean fractal dimension, extreme value for area and perimeter. The explanation for this lies in the resulting programs of the genetic programming method which produces, given the function set used, also if conditions. An example of part of the resulting programs frequently contained the condition: `if tumor size > mean symmetry and lymph node status > 0.0` … . Such relationships between the different features can influence the outcome more than the actual values of the features.

Function set: 300 generations were used and every function set tested contained at least the four basic mathematical operations (addition, subtraction, multiplication and division). Additionally, function sets using comparison operators, if-statements, logarithm, exponential, min, max, cosine, sine, random numbers or sigmoid functions were considered. The best performance was achieved when using nothing but basic mathematical operations, comparison operators and if-statements. Function sets without the later two performed considerably worse (above 15), while function sets consisting at least of the above functions had on average an expected error of 12.60 to 14.44.

For reasons of time constraints only 200 generations were used for finding the next 3 parameters (maximal crossover depth, crossover and mutation rates).

Maximal crossover depth: Maximal crossover depths of 9, 12 and 17 were tested. Using 9 resulted in a higher expected error (13.61) than the 12 or 17.

Crossover and mutation: The crossover rates 0.5, 0.7, 0.8, 0.9 were used, together with mutation rates between 0.1 and 0.5. The crossover rate of 0.9 and the mutation rate of 0.1 produced the highest expected error of 13.09.

Population size: Taking 2,000 individuals and 300 generations improved the average error to 12.24. For 5,000 and 10,000 individuals only 10 test cases were run. In case of 5,000 individuals the expected error decreases significantly to 10.88. Surprisingly, the test runs with a population size of 10,000 do not achieve a similar increased performance. With an average expected error of 12.17, the performance is hardly better than for a population size of 2,000. This might be due to the low number of tests, so that one or two tests with a bad result influence the average result significantly. More tests would be needed to verify if 5,000 is indeed the best choice for the population size. However, using higher population sizes increases the time cost. The execution times measured, taking the average of 10 runs, are shown in Figure 1. As estimated, a linear increase in the execution time can be observed while increasing the population size.

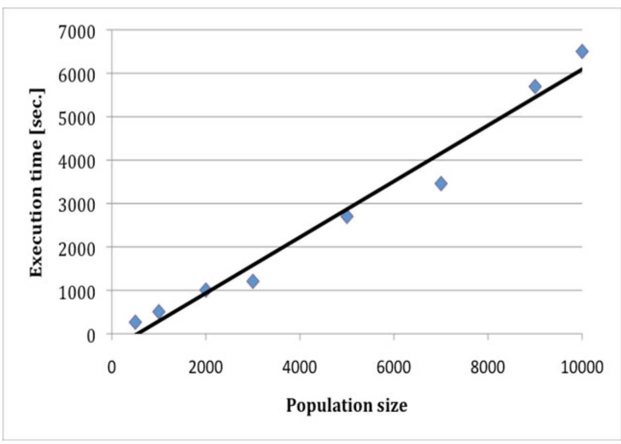


Fig. 1. Performance vs. population size

Number of generations: Running 10 tests with 1,000 generations resulted in an expected error of 10.9. On average, the best result was found after 734 generations.
The second part of the evaluation was to investigate the average error in comparison to varying δ values. The best parameter settings found earlier were used: population size: 2,000, number of generations: 300, mutation rate: 0.1, crossover rate: 0.9, maximal init depth: 5, maximal crossover depth: 9, new individuals per generation:

0.1, function set: addition, subtraction, multiplication, division, logical and, logical or, >, <, if, feature set: tumor size, lymph node status, mean symmetry, extreme radius, standard deviation of radius, extreme perimeter, mean fractal dimension and extreme area of cell nuclei.

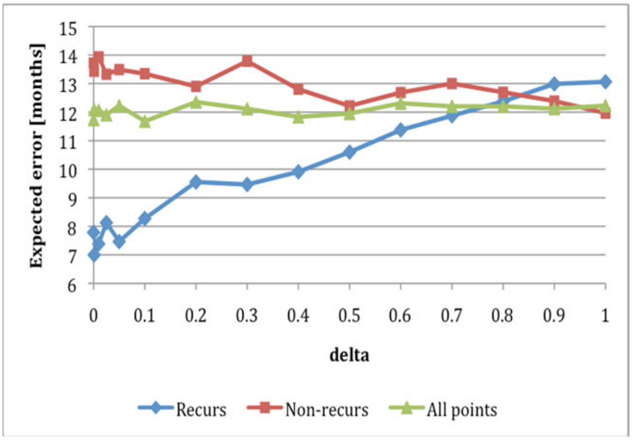


Fig. 2. Expected error vs. varying δ values

Figure 2 shows the expected error for varying δ values. As can be seen, the error for the recurs is clearly increasing with the value of δ, corresponding to the higher emphasis on underestimation of TTR in contrast to overestimation. The increase, considering all data points, is less noticeable since the slight decrease of the expected error on the non-recurs balances the increase on the recurs to some extent. The performance on recurs is better for low values of δ, while for higher values the performance on the non-recurs exceeds the one on recurs. From this one might guess that the prognosis of TTR works better than the one of DFS. After all, for δ=0 the underestimation of DFS and the overestimation of TTR are given the same weight in the fitness function, thus one might expect to get a similar performance. For δ=1, on the other hand, underestimation contributes to the fitness value for all 198 data points, but overestimation only for the 47 recur cases.

Figure 3 shows the overestimation and underestimation of the recur values. As mentioned before, δ determines the relation between the overestimation and the underestimation of TTR. It can be clearly seen that the programs are not able to focus very well on preventing overestimation when δ is higher, since they have to control underestimation of recurs as well as to minimize the fitness value. This leads to a higher emphasis on controlling underestimation on the whole dataset, explaining the slight decrease of the expected error on non-recurs with increasing δ. Note that preventing underestimation of TTR seems to work better than preventing overestimation. For high values of δ, the average underestimation is only 3.1 months, while the best result for the overestimation is 7.4 months, achieved with δ=0.01.

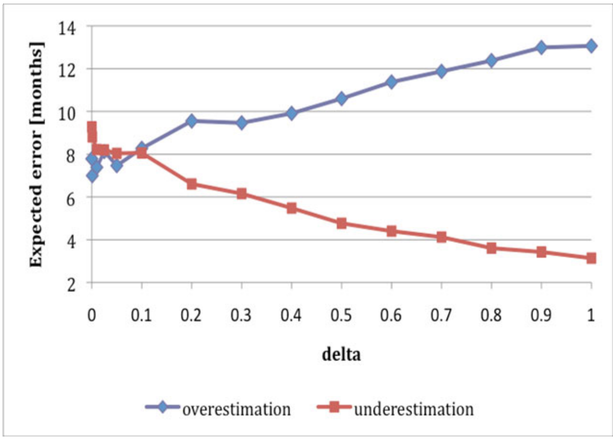


Fig. 3. Expected error of recur values for overestimation and underestimation

In Figure 4, the fitness value for Equation (1) including the δ term and without the δ term is shown. As can be seen, the larger the δ value gets, the larger the difference between the two curves gets, indicating the increase in importance of the third term. As this term becomes more important for the fitness, the programs focus more on the third term than on the remaining two. Therefore, we can see a higher fitness value without the delta term.

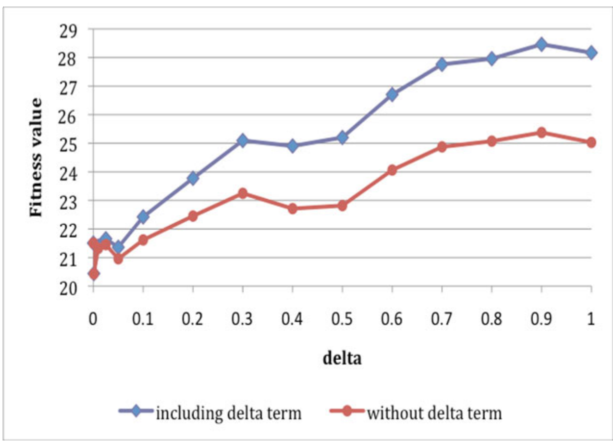


Fig. 4. Investigation of δ term in fitness function

Taking Figure 2 and 3 into account, the higher fitness value can be explained by the increased expected error for the recur cases. In contrast to measuring the performance on all data points, the fitness measure places an equal weight on the two classes. This is why the fitness value is influenced more significantly by a higher error on recurs than the overall performance, which puts an equal weight on each data point, thus the error on the non-recurs (151 cases) is more important than the error on the 47 recur cases.

The best values achieved with the genetic programming method compared to the linear programming technique for $\delta = 0.1$, shows that genetic programming achieves a higher accuracy on all prognostic formulations, as shown in Table 1; the improvement is above 33 %.

Table 1. Comparison of Linear Programming (LP) with Genetic Programming (GP) Results

	All points	**Non-recur**	**Recur**
LP	18.3 months	19.9 months	13.0 months
GP	11.7 months	13.3 months	8.3 months
Improvement	36.1 %	33.2 %	36.2 %

4 Conclusion

Breast cancer victims' chances for long-term survival are improved by early detection of the disease. Early detection in turn is enhanced by an accurate diagnosis. The choice of appropriate treatments immediately following surgery is largely influenced by prognosis, which provides the expected long-term behaviour of the disease. Therefore, an automatic method to aid physicians in their diagnosis and prognosis is of essence.

Previous work regarding the prognosis of breast cancer used a linear programming approach, which arrived at an expected error of 13.0 to 18.3 months, which at that time, attained a better prognosis correctness than other available techniques.

The proposed approach used genetic programming to enhance the prognosis accuracy even further. The fine-tuning of feature selection, selection of the function set, maximal initial depth and crossover depth, crossover and mutation, population size and the number of generations achieved an expected error between 11.7 and 12.4 months, which accounts for an above 36 % higher prognosis accuracy. The main contributor for the higher accuracy was the selection of functions that increased the accuracy. Given that the feature selection of both techniques resulted in a different feature set implies that feature selection is specific to the learning technique. In particular, the possibility of genetic programming to include non-linear functions and if-statements seems to make the difference. The performance for a function set without if-statements is only slightly better than the results achieved by Street et al., which resulted in an expected error of 15.63.

Future work involves running the genetic programming approach with a larger population size, as further improvement can be achieved. In addition, the number of generations also increases the accuracy, however, again at the cost of the execution time. As a larger population size and a larger number of generations imply a longer execution time, the parallelization of the genetic programming could achieve a better performance.

Acknowledgments. The authors would like to thank the Natural Sciences and Engineering Research Council (NSERC) for the partial funding of this research.

References

1. World Health Organization, Fact sheet No. 297: Cancer, http://www.who.int/mediacentre/factsheets/fs297/en/index.html (last retrieved October 2009)
2. Nouri, K.: Skin Cancer. McGraw-Hill, New York (2007)
3. Calgary Breast Health Program, http://www.calgaryhealthregion.ca/breasthealth/breast_cancer/types_of_breast_cancer.htm (last retrieved May 2010)
4. Breast Cancer Prognosis, http://www.healthcentral.com/breast-cancer/ (last retrieved May 2010)
5. Prognostic Wisconsin Breast Cancer Database, UCI Machine Learning Repository, http://www.ics.uci.edu/~mlearn/MLRepository.html (last retrieved May 2010)
6. Street, W.N.: Cancer Diagnosis and Prognosis via Linear-Programming-Based Machine Learning, PhD thesis, University of Wisconsin-Madison (August 1994)
7. Street, W.N., Mangasarian, O.L., Wolberg, W.H.: An Inductive Learning Approach to Prognostic Prediction. In: Proceedings of the Twelfth International Conference on Machine Learning, pp. 522–530. Morgan Kaufmann, San Francisco (1995)
8. Mangasarian, O.L., Street, W.N., Wolberg, W.H.: Breast cancer diagnosis and prognosis via linear programming. Journal of Operations Research 43, 570–577 (1995)
9. Wolberg, W.H., Street, W.N., Heisey, D.H., Mangasarian, O.L.: Computer-derived nuclear grade and breast cancer prognosis. Journal of Analytical and Quantitative Cytology and Histology 17(4), 257–264 (1995)
10. Mangasarian, O.L., Meyer, R.R.: Nonlinear perturbation of linear programs. SIAM Journal on Control and Optimization 17, 745–752 (1979)
11. Cramer, N.L.: A representation for the Adaptive Generation of Simple Sequential Programs. In: Proceedings of an International Conference on Genetic Algorithms and the Applications, Grefenstette (1985)
12. Koza, J.R.: Genetic Programming: On the Programming of Computers by Means of Natural Selection. MIT Press, Cambridge (1992)
13. Banzhaf, W., Nordin, P., Keller, R.E., Francone, F.D.: Genetic programming: an introduction on the automatic evolution of computer programs and its applications. Morgan Kaufmann Publishers, San Francisco (1998)
14. Poli, R., Langdon, W.B., McPhee, N.F.: A Field Guide to Genetic Programming, http://www.gp-eld-guide.org.uk (last retrieved May 2010)
15. Java Genetic Algorithms Package (JGAP), source code, http://jgap.sourceforge.net (last May 2010)

Classification of Software Artifacts Based on Structural Information

Yuhanis Yusof[1] and Omer F. Rana[2]

[1] College of Arts and Sciences, Information Technology Building, Universiti Utara Malaysia, 06010 UUM Sintok, Kedah, Malaysia
yuhanis@uum.edu.my
[2] School of Computer Science, Cardiff University, Cardiff CF24 3AA, Wales, UK
o.f.rana@cs.cardiff.ac.uk

Abstract. Classification of software artifacts, in particularly the source code files, are currently performed by administrator of a repository. Even though there exist automated classification on these repositories, nevertheless existing approach focuses on semantic analysis of keywords found in the artifact. This paper presents the use of structural information, that is the software metrics, in determining the appropriate application domain for a particular artifact. Results obtained from the study show that there is a difference in the metrics' trend between files of different application domain. It is also learned that results obtained using k-nearest neighborhood outperformed C4.5 decision tree and the one generated based on Discriminant Analysis in classifying files of database and graphics domain.

Keywords: classification, software metrics, decision tree, k-nearest neighborhood, discriminant analysis.

1 Introduction

As software functionality, size and application domains increase, the difficulty of categorizing and classifying software artifacts for reuse and maintenance increases. Automatic software classification has become an important topic in software engineering [11]. There are various reasons for this, for instance `SourceForge.net` had over seventy thousand registered software [12] in 2002. As this repository receives input (e.g software files) from various developers, categorizing software relies on the text provided by developers and comments contained within the software source code. One issue which arises from such situations is to find ways to support text-based classification and hence enhance the search process. There is therefore a need for an alternative method in software classification [11]. Manual classification requires more time and a high level of software understanding and classification [11]. Consequently, repository administrators need to put extra effort in identifying the specific application domain that a software belongs to [9]. The study tries to overcome such problem by introducing the use of structure information (i.e software metrics) extracted from source

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 546–555, 2010.

code. Software metric can be categorized into two types; internal metric such as the lines of code, characters and functions, and external metric responsible for measuring and identifying software behavior [18].

Our objectives are two folds in this work: 1) to mine structure information contained in source code and, 2) to use this for automating software classification. We extend the approach used in `SourceForge.net` [3] and `Freshmeat.net` [2] by using software metrics as the feature set in determining the appropriate application domain for software artifacts. Our work focuses on categorizing database and graphics software artifacts due to.

We use the freeware C and C++ Code Counter (CCCC) [1] to extract software metrics from source code. CCCC is a source code analyzer tool that analyses C++ and Java source code files and generates a report on various metrics of the code. At the time of our research, CCCC extracted a total of 19 software metrics, nevertheless, only twelve metrics were used in the experiment undertaken. The selection is made based on the strength of relationship that exists between the metrics and the application domain categorized in `Sourceforge.net` [3]. Regardless of the direction of the relationship (positive or negative), Cohen [6] suggests that the value of Pearson correlation, r, is considered to be large (i.e strong) if it is between 0.5 and 1.0. If r is in the range of 0.3 and 0.49, then it is considered as medium, and, if r is equal or less than 0.29, then the correlation is considered to be small. Based on this suggestion, only software metrics that depict a correlation as low as 0.5 are chosen to be used as the independent variables in determining the dependent variable (i.e application domain). The selected metrics are listed below:

1. McCabe's Cyclomatic Complexity (Mvg) - every decision point in a method (e.g `if`, `for`, `while`, or `case` statement) is counted; additionally, one is added for the method's entry point, resulting in an integer-based measurement denoting a method's complexity.
2. Coupling Between Objects (Cbo) - the use of another object's methods or instance variables. Cbo is defined for classes and interfaces, constructors and methods. It counts the number of reference types that are used in field declarations, formal parameters & return types, `throws` declarations and local variables.
3. Weighted Method per Class (Wmc) - the sum of a weighting function over the functions of the module. The *Wmc* uses the nominal weight of 1 for each function and hence measures the number of functions.
4. Fivis - a count in a file that is accepting information from another file and the count is restricted to the part of the interface that is externally visible.
5. Ficon - a count in a file that is accepting information from another file and the count implies changes to a client must be recompiled if the supplier's definition changes.
6. Fovis - a count in a file that sending information to another file and the count is restricted to the part of the interface that is externally visible.
7. Focon - a count in a file that is sending information to another file and the count implies changes to the client must be recompiled if the supplier's definition changes.

We used source code files obtained from `SourceForge.net`, which are then automatically classified into application domains using the predictions made by Discriminant Analysis (DA) [13], C4.5 decision tree [20] and k-nearest neighbour (KNN) [7]. By using DA [13] which is based on statistical analysis, we can determine how software metrics may be combined to predict the most likely application domain of a file. On the other hand, the C4.5 [20] is included in the experiment as it has always been used as a comparison base line in machine learning [21]. In addition, the work undertaken by Lim et al. [15] and Ganti et al. [10] shows that the C4.5 algorithm generates good classification accuracy and is the fastest among the compared algorithms (i.e neural networks and KNN). KNN [7] is chosen as it classifies an object (i.e source code file) based on the object's neighbourhood. With this, we can identify if files from the same domain contain similar metric values and are of the same pattern. Furthermore, as our repository contains artifacts from different developers, with varying styles of writing source code, the use of KNN (an effective classifier for large data sets [14]) would be suitable.

2 Related Work

Classifications are widely used in Internet directories, such as Google and Yahoo, and information portal. Classification systems provide means of organizing resources and quickly retrieving them when required. Existing work on software classification involves various artifacts including workbenches [22], requirements and source codes [23,16]. Most of the applications stored in the open-source repository systems such as Sourceforge.net [3] and Freshmeat [2] are classified into various categories (e.g application domain and programming language). In order to reuse source code files, a user may need to manually analyse each of the applications (that may contain more than one file) retrieved. This is because the applications in these repositories are classified into appropriate domains based on the overall description provided by the developers.

A developer attempting to organize a collection of software artifacts, would most likely categorize the artifacts based on information in the source code itself (e.g identifier names), some design specifications and the documentation provided with the source code file. Often, a developer would use comments, manuals and *ReadMe* files for categorizing source files.

2.1 Classifiers

Shafia et. al. [22] propose six different workbenches for classifying software – built by identifying, subjectively, relevant features to represent each workbench. These features are given different weighing to indicate their importance. Due to its subjective nature, the approach may differ across individuals.

Marcus et. al [16] employ the term occurrences approach to indicate domain knowledge and concepts embedded in a program source code. Identifier names and comments are extracted from the program before the use of Latent Semantic Indexing (LSI) [8]. In addition to recording which keywords a program contains, the LSI examines the program collection as a whole, to see which other programs contain some of those same words. LSI considers programs that have many words in common to be semantically close, and ones with few words in common to be semantically distant. Such an approach is similar to the one undertaken by Walters and Rajashekar [24]. They extract relevant keywords from software documents prior to the classification process.

Ugurel et al. [23] classify programs into appropriate application domains and also programming languages using three components: feature extractors, vectorizers and Support Vector Machine (SVM) [17] classifiers. They train the classifier with automatically extracted features from the code, comments, and the ReadMe files (i.e. tokens in the code, words, and lexical phrases in the comments and *ReadMe* files). The results imply that large archive collections of mixed software components such as software documentation and source code files can effectively be automatically classified and categorized. Nevertheless, such an approach is based on terms extracted from documentation. Therefore the approach is only applicable to software that are well-documented. To the knowledge of the researcher, there is no work undertaken in software artifact classification that is based solely on structure information contained in source code files. In addition, other than SVM, there is no other work that uses machine learning techniques in classifying source code files into application domains.

Statistical analysis for retrieval has also shown promising results [5]. Such approaches include the use of discriminant function analysis and linear regression. Discriminant function analysis (DA) is used to determine which attributes in an object can be used to separate two or more naturally occurring categories. The model is built based on a set of objects (training set) for which the categories are known.

3 Mining Structure Information

In this work, CCCC is used to extract software metrics contained in source code files. A total of 584 Java source code files (371 database and 213 graphics) were included in the mining process. The extracted metrics are later analyzed to identify similarities that exist between files of the same domain. Figure 1 shows the average value of software metrics, and identifies that artifacts of database and graphics application domains illustrate different trends in their metric values (Mvg, Wmc and Cbo). In this analysis, comparison is undertaken as the increment or reduction of values when two metrics are compared. For example, if metric X is 20 and metric Y is 25, then it is suggested that there is an increment of 25% in metric values when X is compared against Y, and there is a reduction of 20% in metric values if X is 25 and Y is 20.

In Figure 1 there is an increment in Wmc when compared to Mvg for the database category (i.e metric Wmc is 2250% bigger than metric Mvg). However, the graphic category shows a different trend; there is a reduction in the metric values when Wmc is compared to Mvg (i.e metric Wmc is 17.78% smaller than metric Mvg). In addition, the graph also reveals that the trend between metrics

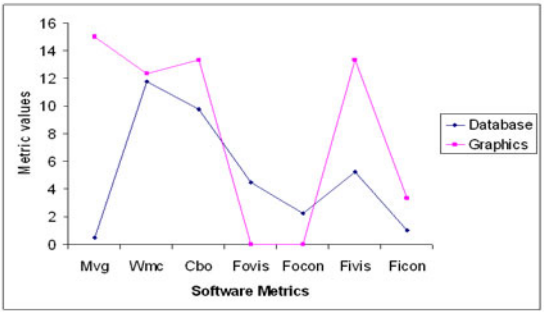


Fig. 1. Software Metrics of Database and Graphics Source Code Files

Wmc and Cbo differ for source code files from different application domains (database and graphics). The graphic file contains Cbo that is 8.11% bigger than Wmc, while the database file contains Cbo that is 17% smaller than Wmc.

4 Structure-Based Classification

To obtain the classifier models, we performed experiments on two data sets; training and testing. The training data set includes the ones used in the mining process - a total of 584 source code files, while the testing data set contains 136 Java source code files. The classifier models obtained during training process were later verified using the testing data set. We used classifiers C4.5 and KNN which are provided by *WEKA* [19] and DA was obtained using *SPSS* [4]. In *WEKA*, the C4.5 decision tree is known as J48 while the KNN is known as an instance based neighbour (IbK). Therefore, in this paper, the results are reported using these names (J48 and IBk).

In measuring the effectiveness of the three classifiers, derivatives of measurements from the information retrieval domain is used; falsePositives, Precision and Recall. Tables 1(a) and 1(b) reveal the falsePositives, precision and recall scores which were calculated after the source code file classification experiment was completed. Data in Table 1(a) shows that J48 and IbK-3 have similar capabilities in identifying database files. This is shown by the recall scores depicted in table 1(a) which shows that there is only a difference of 0.1 between the two classifiers. However, J48 outperforms both DA and IbK-3 by 10% in the precision scores. 92% of source code files classified into the database domain, using the J48 model, are actually relevant, hence generating the lowest falsePositives score of 0.122.

Data in Table 1(b) shows that there is a reduction in the precision and recall scores for J48 and IbK-3 when compared to the scores obtained for database source code files. On the other hand, the DA model generates a better classification then the one made for database files. An increment of approximately

Table 1. Classification Analysis on Training Data Set

(a) Database

Measurement	J48	IbK-3	DA
Precision	0.929	0.882	0.837
Recall	0.916	0.927	0.593
falsePositives	0.122	0.216	0.202

(b) Graphics

Measurement	J48	IbK-3	DA
Precision	0.858	0.861	0.53
Recall	0.878	0.784	0.732
falsePositives	0.084	0.073	0.407

15% in recall was achieved by DA in classifying graphics files compared to the database files.

The classifier models from the previous experiment were then used to classify a new set of source code files (testing data set) containing 136 Java files. This data set contains 64 database and 72 graphics files which were also obtained from `Sourceforge.net` [3]. From the data depicted in table 2, it is noted that the classification accuracy obtained using the three classifiers are in the range of 57% to 73%. Out of 136 source code files, the IbK-3 has correctly classified 98 files while 94 files classified by the J48 model was relevant. On the other hand, the DA has correctly classified 57% of the files (i.e 78 out of 136).

Table 2. Classification Accuracy: Testing Data Set

Data Set	J48	IbK-3	DA
Testing	70.149	73.134	57.46

In Tables 3(a) and 3(b), the precision, recall and falsePositives scores obtained after completing the classification experiment are provided. As expected (based on results obtained in the previous experiment), the IbK-3 model has outperformed the DA and J48 in classifying source code files into the application domains. The average precision score for IbK-3 was 0.739 while J48 generated an average of 0.714. On the other hand DA generated an average of 0.581 and 0.567 for precision and recall, respectively. The average recall scores for IbK-3 and J48 were 0.725 and 0.694, respectively.

The interpolated precision recall curve for the classification of database and graphics source code files are illustrated in Figures 2(a) and 2(b). Based on Figure 2(a), we noted that IbK-3 outperformed other classifiers in classifying database files. This is achieved by correctly classifying 51 out of 64 database

Table 3. Classification Analysis on Testing Data Set

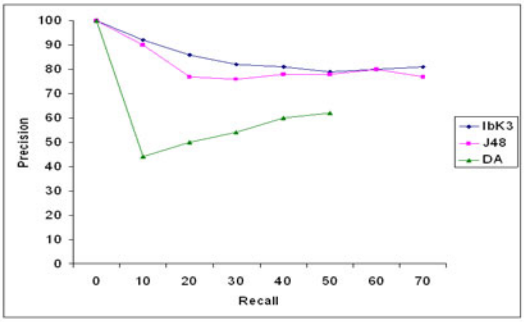
(a) Database

Measurement	J48	IbK-3	DA
Precision	0.762	0.797	0.607
Recall	0.75	0.797	0.578
falsePositives	0.239	0.203	0.393

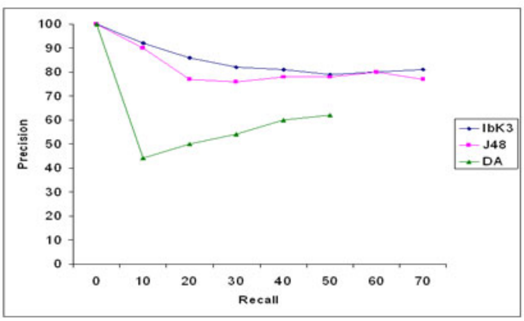
(b) Graphics

Measurement	J48	IbK-3	DA
Precision	0.667	0.681	0.556
Recall	0.639	0.653	0.557
falsePositives	0.406	0.347	0.444

files, hence generating a recall of approximately 80% (rounded). It is therefore supports the result obtained earlier (refer to Table 3), which showed that the IbK-3 is a better classifier than the DA or J48 in classifying database files. The interpolated precision recall curve that is depicted in Figure 2(b), illustrates that IbK-3 has also out performed DA and J48 in classifying graphics source code files. The IbK-3 has correctly classified 65% of the graphics files compared to J48 which correctly classified 64% of files from the same domain.



(a) Database



(b) Graphics

Fig. 2. Interpolated Precision Recall Curve for Source Code Files Classification using Testing Data Set

The precision and recall scores that are depicted in Table 3 suggest that the IbK-3 is a better classifier than the J48 or DA. We investigate if the result is supported by statistical analysis. A normality test was performed using *SPSS* version 11.5 and the results is depicted in Table 4(b). It is noted that there is a significant difference at $\alpha = 0.05$ in the precision and recall scores across the three classifiers. Both of the `Asymp.Sig.` values for the precision and recall scores are 0.0 which are less than the α value. This indicates that there is a difference in precision and recall scores across the J48, DA and IbK-3.

The mean ranks depicted in Table 4(a) reveals that the IbK-3, with 63.52 and 68.37 had the highest precision and recall scores, with the DA reporting the lowest. Such a result indicates that the IbK-3 is a better classifier compared to DA and J48.

Table 4. Kruskal-Wallis Test Result Relative to Source Code File Classification into Database Domain

(a) Mean Rank

Scores	Classifier	Mean Rank
Precision	J48	50.22
	IbK-3	63.52
	DA	22.77
Recall	J48	45.50
	IbK-3	68.37
	DA	22.63

(b) Test Statistics

	Precision	Recall
Chi-Square	38.146	46.006
df	2	2
Asymp.Sig.	0.000	0.000

In order to identify whether there is a significant difference between classifiers in precision and recall scores for graphics source code files, we conducted a one-way between-groups analysis of variance (one way ANOVA). The statistical results depicted in Table 5 indicate that the difference in precision and recall scores, obtained using DA, J48 and IbK-3 are less significant at $\alpha = 0.05$. This is depicted by `Sig.` values of 0.879 for precision and 1.00 for recall which were greater than the α value used in the test, that is 0.05.

Table 5. ANOVA Test Results of Precision and Recall Scores of Graphics Files

		Sum of Squares	df	Mean Square	F	Sig.
Precision	Between Groups	0.005	2	0.002	0.129	0.879
	Within Groups	1.646	87	0.019		
	Total	1.651	89			
Recall	Between Groups	0.000	2	0.000	0.000	1.000
	Within Groups	1.301	87	0.015		
	Total	1.301	89			

5 Conclusion

In this paper, we presented an alternative approach to classifying software artifacts into application domains. As existing work only focuses on semantic analysis, we see the need for utilizing structure information hidden in the artifacts. Such an approach would be helpful especially to artifacts that are not well documented by their developers. Furthermore, as open source repositories are rapidly growing, relying on humans to classify the artifacts may not efficient. By mining software metrics contained in source code files, it is learned that there is a difference of trends (shown by the different shapes of graphs) in software metrics for database and graphics source code files. Based on the experiments undertaken, it is noted that database files contain greater value of Wmc when compared to metrics Mvg and Cbo. On the other hand, metric Wmc for graphics files are smaller than metrics Mvg and Cbo. Classification performed using k-nearest neighborhood has shown to better compared to the ones obtained using C4.5 and discriminant analysis. As k-nearest neighbor measures similarity, such a result indicates that the software artifacts (i.e source code files) of the same domain illustrate similar trends of software metrics.

The work described in this paper can further be extended by including more structural information which later can be reduced using feature set algorithms. In addition, multi-classification can also be experimented on artifacts that can be categorized into more than one application domain.

References

1. C and c++ code counter, `http://sourceforge.net/projects/cccc/` (last accessed on April 15, 2010)
2. Freshmeat, `http://freshmeat.net/` (last accessed on May 10, 2010)
3. Sourceforge, `http://sourceforge.net/` (last accessed on May 10, 2010)
4. Spss, `http://www.spss.com` (last accessed on April 15, 2010)
5. Chung, K.-P., Fun, C.C.: A hierarchical nonparametric discriminant analysis approach for a content-based image retrieval system. In: ICEBE 2005: Proceedings of the IEEE International Conference on e-Business Engineering, Washington, DC, USA, pp. 346–351. IEEE Computer Society, Los Alamitos (2005)
6. Cohen, S., Mamou, J., Kanza, Y., Sagiv, Y.: Xsearch: A semantic search engine for xml. In: Proceedings of the 29th VLDB Conference, Berlin, Germany (2003)
7. Cover, T.M., Hart, P.E.: Nearest neighbor pattern classification. IEEE Transactions on Information Theory IT-13(1), 21–27 (1967)
8. Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K.: Indexing by latent semantic analysis. Journal of the American Society for Information Science, 391–407 (1990)
9. Fuchs, N.E.: Specifications are (preferably) executable. Software Engineering Journal 7(5), 323–334 (1992)
10. Ganti, V., Gehrke, J., Ramakrishnan, R.: Mining very large databases. Computer 32(8), 38–45 (1999)
11. Kawaguchi, S., Garg, P.K., Makoto, M., Inoue, K.: Automatic categorization algorithm for evolvable software archive. In: Proceedings of the Six International Workshop on Principles of Software Evolution, pp. 195–200 (2002)
12. Kawaguchi, S., Garg, P.K., Makoto, M., Inoue, K.: Mudablue: An automatic categorization system for open source repositories. In: Proceedings of the 11th Asia-Pacific Software Engineering Conference, pp. 184–193 (2004)
13. Klecka, W.R.: Discriminant Analysis, 1st edn. Sage Publications, Thousand Oaks (1980)
14. Kwon, O.-W., Lee, J.-H.: Text categorization based on k-nearest neighbor approach for web site classification. Information Processing Management 39(1), 25–44 (2003)
15. Lim, T.-S., Loh, W.-Y., Shih, Y.-S.: A comparison of prediction accuracy, complexity, and training time of thirty-three old and new classification algorithms. Machine Learning 40(3), 203–228 (2000)

16. Marcus, A., Sergeyev, A., Rajlich, V., Maletic, J.I.: An information retrieval approach to concept location in source code. In: WCRE 2004: Proceedings of the 11th Working Conference on Reverse Engineering (WCRE 2004), Washington, DC, USA, pp. 214–223. IEEE Computer Society, Los Alamitos (2004)
17. DSFP Modeling and Forecasting. Svm - support vector machines, `http://www.dtreg.com/svm.htm` (last accessed on April 15, 2010)
18. Nagappan, N.: Toward a software testing and reliability early warning metric suite. In: Proceedings of International Conference on Software Engineering, pp. 60–62 (2004)
19. U. of Waikato. Weka, `http://www.cs.waikato.ac.nz/ml/weka` (last accessed on April 15, 2010)
20. Quinlan, J.R.: C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers Inc., San Francisco (1993)
21. Ruggieri, S.: Efficient c4.5. IEEE Transactions on Knowledge and Data Engineering 14(2), 438–444 (2002)
22. Shafia, Mustafa, T., Raza, A., Jamil, U., Shahzad, F.: A classification model for software workbenches. European Journal of Scientific Research 41(1), 109–121 (2010)
23. Ugurel, S., Krovetz, R., Giles, C.L.: What's the code?: Automatic classification of source code archives. In: Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 632–638. ACM Press, New York (2002)
24. Walters, S., Rajashekhar, T.B.: Mapping of two schemes of classification for software classification. Cataloging and Classification Quarterly 41(1), 163–182 (2005)

Clustering Client Honeypot Data to Support Malware Analysis

Yaser Alosefer and Omer Rana

School of Computer Science & Informatics,
Cardiff University, UK
{Y.Alosefer,O.F.Rana}@cs.cf.ac.uk

Abstract. Client honeypots visit and interact with suspect web sites in order to detect and collect information about malware. Malicious websites may cause a number of activities to be performed on a victim's system; each activity is performed in different stages. We use a state machine to represent the activities performed by the malicious web page into pre-defined states. These states can be used to summarise interactions with malicious web pages using the same state machine structure. The states are then passed to a clustering algorithm to group similar malicious web page exploits in order to better understand how software can be developed to better respond to such attacks. The outputs of the clustering algorithm are categorized to build up groups of similar states that represent the malicious activities performed on the victim's system. The benefit of using this process is to build families of malicious web pages with similar behaviours (behaviour families) leading to the development of common approaches to deal with such exploits.

Keywords: web-based malware, client honeypot, automated state machine, clustering.

1 Introduction

A honeypot is defined as a "security resource whose value lies in being probed, attacked or compromised" [1]. The aim of a honeypot is to collect high-value data on attacks and attackers by monitoring the state of a real operating system or services. The automated state machine, when applied to a client honeypot as described in this paper, gives honeypot technology a number of benefits in the analysis of malware and tracing the details of activities performed by malware on a machine.

An automated state machine is applied to a client honeypot to convert the log file of a high interaction client honeypot [4] to a state machine. The benefits of this are the following:

1. The state machine encodes each action performed by a malware and saves the action along with its parameters.
2. Each state machine contains all the system states during the interacting between the malware and the system, which helps to review the system states subsequently.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 556–565, 2010.

3. The state machine for each interaction helps to analyse the malware and identify its processes in order to develop a patch for it or analyse it's attack behaviour.
4. As the state of the host system is known prior to the interaction, it is possible to accurately record changes in the system as a consequence of the interaction with the malware.

The active client honeypot interacts with with the site referred to in the URL and any modifications made to the host system are recorded. In our work, we use a client honeypot application called Capture-HPC [4] to scan the URLs provided and then to produce log files, which are then converted to the state machine file structure by our system. We apply clustering algorithms on the generated state transitions to find similarity between different attacks and to group attacks by different distance measures. In order to do this, we identify a framework that can help the analyser understand relationships between web-based malware associated with a particular web site, as well as mechanisms to discover inconsistent activity that can be performed by a malware with a different name.

2 The Automated State Machine

The finite state machine (FSM) is used to monitor changes to finite and limited aspects of the system, such as the file system, the registry and processes generated based on interactions with an external server. These are the three main state types that we want to monitor. The state machine diagram is shown in Fig. 1. This FSM can also be illustrated with the 3-tuple as A ={S, Σ, δ}, where:

- The Σ alphabet represents the finite number of events Σ = {e1, e2, e3 …, e18}.
- S is the non-empty set of finite states: S = {s0, s1, s2, s3, s4}.
- δ is the transition function between states and represented as: $\delta: S \times \Sigma \rightarrow S$.

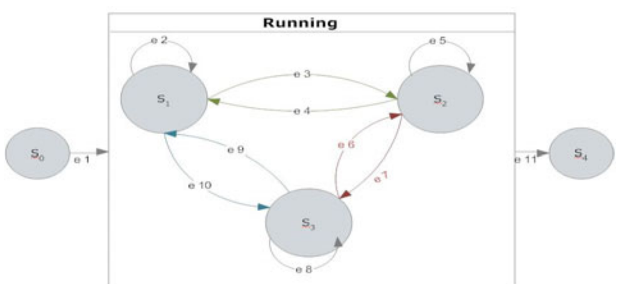


Fig. 1. State machine diagram (the boxes indicate states and the arrows represent transitions)

The state machine includes four main states that are monitored through the 11 possible transitions between the states. We have built the state machine using Windows Workflow Foundation [3] and monitor three states: file system, processes and the registry.

3 State Machine Structure

The client honeypot state machine (CHSM) tool is used to convert Capture-HPC log files into the state machine representation that produces more information about each URL that is scanned. The state machine can be used to extract useful information from the client honeypot; the client honeypot tool scans specific URLs and then produces log files only. This makes it difficult for an analyser to interpret the output or to determine the relationship between one URL and another, as shown in Fig. 2. The state machine is therefore used to expose malicious activities in order to improve behaviour detection.

```
"file","Nov 11, 2007 2:50:50 PM","C:\Program Files\Internet Explorer\IEXPLORE.EXE","Write","C:\WINDOWS\TEMPmbroit.exe"
"process","Nov 11, 2007 2:50:50 PM","C:\Program Files\Internet Explorer\IEXPLORE.EXE","terminated","C:\WINDOWS\TEMPmbroit.exe"
```

Fig. 2. Snapshot of the Capture-HPC log file

The state machine tool converts the Capture-HPC log file to the state machine version; a real example of a *state machine chain* is shown in Fig. 3 and Table 1 shows the state name and action type descriptions:

```
S0---> S1W[C:\Program Files\Internet Explorer\IEXPLORE.EXE]=[C:\WINDOWS\TEMPmbroit.exe]--->S2N[C:\Program
Files\Internet Explorer\IEXPLORE.EXE]=[C:\WINDOWS\TEMPmbroit.exe]--->S2K[C:\Program Files\Internet
Explorer\IEXPLORE.EXE]=[C:\WINDOWS\TEMPmbroit.exe]---> S4
```

Fig. 3. Example of state machine file

The *state machine chain* as shown in Fig. 3 uses a state name and action description. A detailed analysis of the example above is given below:

1. **S0:** is the initial state of the state machine chain, which shows the starting point of the URL activity.
2. **--->**: shows the next activity.
3. **S1W:** is the second state in our example, which shows that there was activity in the file system and the action taken was (W), which means that there is a new file: a write operation in the file system.
4. **[C:\Program Files\Internet Explorer\IEXPLORE.EXE]:** is the source of the change in the system.
5. **[C:\WINDOWS\TEMPmbroit.exe]:** the result of the previous action, which led to the file “TEMPmbroit.exe” being written to this destination path.

6. **S4:** the final state; this indicates that the end of the activity of the URL has been reached.

Table 1. State name and action type descriptions

State	Character/Description	
Initial State (S0)	-	-
File system (S1)	W	Write
	D	Delete
	C	Create
Process (S2)	N	New
	K	Kill
Registry (S3)	S	Set
	R	Remove
Completed State (S4)	-	-

3.1 State Machine Stamp

The ordering of activities within the state machine chain is important as we group activities according to their ordering. The order is important within a web-based exploit; an example of a state machine chain is shown below:

S0,S1W1=2,S1W1=3,S2N1=3,S2K1=3,S4

An exploit can change the order of some of its activities (without changing the impact on the victim's machine), which can cause our clustering algorithm to categorize the same exploit into two different groups. A solution is to have a unique ID for each state machine chain and then compare this ID to others to compare similar exploits.

S0,S1W1=2,S1W1=3,S2N1=3,S2K1=3,S3S3=4,S4
S0,S1W1=3,S1W1=2,S2N1=3,S2K1=3,S3S3=4,S4

Fig. 4. Example of a state machine file

The state machine chain in Fig. 4 shows the same web-based exploit but with different ordering of activities. However, both of the exploit functions and impact on the victim's machine are the same. We have made a state machine stamp, a unique ID, which can recognize this change. An example of a state machine and its stamp is shown in Fig. 5.

State machine chain:
S1W1=2,S1W1=3,S2N1=3,S2K1=3,S3S3=4

State machine stamp:
2S1W(1=2,1=3),1S2N(1=3),1S2K(1=3),1S3S(3=4)

Fig. 5. Example of a state machine chain and stamp

The stamp represents the state machine chain in a unique way by grouping similar activities together. Therefore, the state machine stamp can determine some of the activity of the same exploit without changing its impact.

Table 2. Three examples of state machine chains and their stamps

State machine chain	State machine stamp
S1W1=2,S1W1=3,S2N1=3,S2K1=3,S3S3=4	2S1W(1=2,1=3),1S2N(1=3),1S2K(1=3),1S3S(3=4)
S1W1=3,S1W1=2, S2N1=3,S2K1=3,S3S3=4	2S1W(1=2,1=3),1S2N(1=3),1S2K(1=3),1S3S(3=4)
S1W1=3,S2N1=3,S3S3=6,S3S3=7,S3S3=8,S3S3=11	1S1W(1=3),1S2N(1=3),4S3S(3=6,3=7,3=8,3=11)

An example of the above idea is below, where the first string represents the state machine chain, which encodes the malicious activities and the outcome, and the impact of this malicious web page is called the state machine stamp, which represents its impact on the victim's machine:

```
[S1W1=2,S1W1=3,S2N1=3,S2K1=3,S3S3=4] => 2S1W(1=2,1=3),1S2N(1=3),1S2K(1=3), 1S3S(3=4)
```

4 State Machine Clustering

The state machine chains that contain the activities of web-based malware do not enable gathering of information about the relationships between one web-based malware and another as shown in the figure below:

S1W1=13,S1W1=14,S2N1=14,S1W1=15,S3S14=16,S1W14=17
S1W1=18,S1W1=14,S2N1=14,S1W1=15,S1W14=17,S3S14=16
S1W1=2,S1W1=3,S2N1=3,S2K1=3,S3S3=4
S1W1=13,S1W1=14,S2N1=14,S1W1=15,S3S14=16,S1W14=17

Therefore, one way to solve this is to group the state machine chains using data mining techniques, especially clustering algorithms. The term clustering covers a number of different techniques used to group similar items together and can be defined as "the process of organizing objects into groups whose members are similar in some way" [5]. Clustering can help us with our state machine files by:

1) Grouping state machine files that are the same, which can be used to group different URLs that use the same web-based malware.
2) Grouping similar state machine files to build behaviour families of web-based malware, containing malware that uses similar techniques or attacks, which could lead to a better analysis of the relationships between the state machine activity files.
3) Reducing the analysis time, providing the analyser with groups of web-based malware to analyse, whereas before each log file would have to be analysed separately to understand the behaviours of the URL scanned.

There are different types of clustering algorithms, such as hierarchical, partitioned or density-based algorithms. The one that we use is a hierarchical algorithm because it can help us to process and group our state machine chains easily. The hierarchical agglomerative algorithm can group clusters (each state machine chain is one cluster) into one cluster by combining the clusters together in different stages. The technique used in the hierarchical agglomerative algorithm to group similar clusters uses various “distance measures”, which in our approach can be of three types:

- **Single linkage clustering:** the minimum distance between items of each cluster.
- **Complete linkage clustering:** the maximum distance between items of each cluster.
- **Average linkage clustering:** the mean distance between items of each cluster.

Using the above three distance measures, the hierarchical agglomerative algorithm can calculate the distance between clusters to merge similar clusters together, which could help to identify similar state machine chains and to combine them to build an automated analysis engine, as well as to build families of similar web-based malware (behaviour families).

5 Clustering Applied in State Machine Files

The clustering algorithm is used to group similar state machine chains together automatically to identify the existence of web-based malware *families*. The clustering algorithm groups the state machine chains according to the ordering of the activities and states as well as using the state machine stamp to determine how ordering can be modified to mislead the resulting cluster formation. The activities undertaken within our clustering framework are shown in Fig. 6.

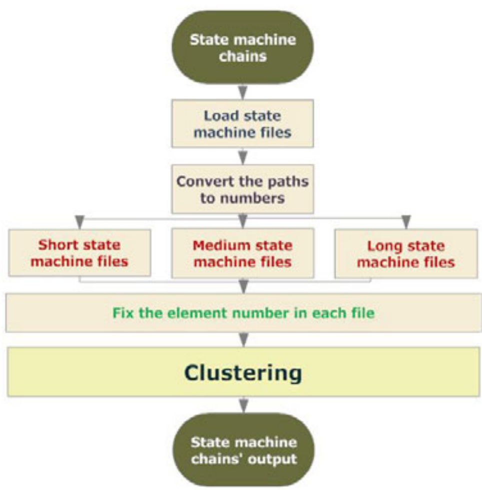


Fig. 6. State machine clustering process

The clustering process as shown in Fig. 6 is divided into six steps, which prepare the state machine files for clustering and then generates the clustered groups. The six steps are discussed below.

1) **Load the state machine files**.
2) **Convert the path to numbers:** this step converts the URL in the source and destination to unique numbers to simplify the state machine chains.
3) **Divide the state machine files:** in this step, the state machine files are divided into three different groups.
   - *Short*: this contains the short state machine file activity. We have found the optimum short length to be less than six activities.
   - *Medium*: medium files contain between six and 15 activities.
   - *Long*: this group contains files with more than 15 activities.

   The benefit of this division is to compare each group with ones of similar lengths, which gives more accurate results.
4) **Fix the element numbers in each file:** this step is carried out to ensure that all files are of the same length (as required by the clustering algorithm), as the state machine files may have different numbers of activities and lengths. Table 3 shows some state machine files which are passed to the clustering algorithm.

Table 3. State machine files before clustering

ID	Activity number	State machine file
1	5	S1W1=2,S1W1=3,S2N1=3,S2K1=3,S3S3=4
2	6	S1W1=5,S1W1=3,S2N1=3,S3S3=6,S3S3=7,S3S3=8
3	4	S1W1=9,S1W1=3,S2N1=3,S3S3=4

To equalize the length of state machine files, additional states can added to ensure that all are of the same length; however these additional states do not count when clustering. This is state “S4” in Table 4. The longest state machine file within Table 3 is number 2, which has six activities, hence “S4” is added to 1 and 3. The new version is as shown in Table 4 below.

Table 4. Missing element in the state machine files

ID	Activity number	State machine file
1	6	S1W1=2,S1W1=3,S2N1=3,S2K1=3,S3S3=4,S4
2	6	S1W1=5,S1W1=3,S2N1=3,S3S3=6,S3S3=7,S3S3=8
3	6	S1W1=9,S1W1=3,S2N1=3,S3S3=4,S4,S4

5) **The clustering:** The data is ready to be passing through the clustering algorithm as shown in Table 4, which compares each activity with others in the same column. The clustering on the state machine is shown below.

The clustering table

Column 1	Column 2	Column 3	Column 4	Column 5	Column 6
S1W1=2	S1W1=3	S2N1=3	S2K1=3	S3S3=4	S4
S1W1=5	S1W1=3	S2N1=3	S3S3=6	S3S3=7	S3S3=8
S1W1=9	S1W1=3	S2N1=3	S3S3=4	S4	S4
S1W1=2	S1W1=3	S2N1=3	S2K1=3	S3S3=4	S4

The clustering algorithm compares all entries in each column using one of three distance measures within the hierarchical agglomerative algorithm, which can be calculated using the following.

- Single linkage clustering:

$$\min\{\, d(x,y) : x \in \mathcal{A},\ y \in \mathcal{B}\,\}. \tag{1}$$

- Complete linkage clustering:

$$\max\{\, d(x,y) : x \in \mathcal{A}, y \in \mathcal{B}\,\}. \tag{2}$$

- Average clustering:

$$\frac{1}{|\mathcal{A}| \cdot |\mathcal{B}|} \sum_{x\in\mathcal{A}} \sum_{y\in\mathcal{B}} d(x,y). \tag{3}$$

The same two entries produce “0” in each distance measure, and the height difference produces “1”. The clustering process calculates the distance between each two entries in each column and then merges the lowest two clusters together, and then moves to the second stage. The second stage also calculates the distance between each entry in each column, and the lowest two entries are merged. This process continues until all the clusters are merged into one large cluster.

6) **State machine file output:** the State machine file output is shown in Fig. 7.

Stage	Combined Clusters: Cluster1	Combined Clusters: Cluster2	Distance
1	1	2	0.0
2	4	5	0.16666667
3	1	4	0.5
4	3	1	0.6666667
5	6	1	1.0

Fig. 7. Table view of our clustering output

6 The State Machine Clustering Experiment

To validate the state machine clustering, we have carried an experiment using 116 Capture-HPC log files generated by scanning 116 unique websites. Firstly, we have converted all the log files to a state machine version as described previously in section 3. Then we loaded all 116 of the state machine chains into our implementation and clustered them all at once to receive the groups that can be passed to the analyser. Fig. 8 shows the result from the clustering framework.

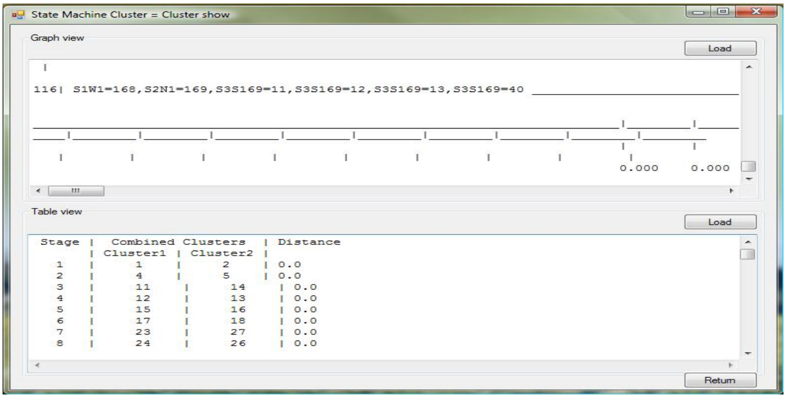


Fig. 8. Result of the state machine clustering experiment

From the result shown in Fig. 8, there are 77 main groups in this experiment shown in table 4. As an illustrative example, two state machine chains from the group number five shown below:

S1W1=29,S1W1=19,S2N1=19,S1W19=21,S3S19=20,S1W1=5
S1W1=18,S1W1=19,S2N1=19,S1W19=21,S3S19=20,S1W1=5

These show that the two state machine activities on the victim's machine are exactly the same except that the second exploit writes a different file in the first activity compared to the first. This can help the analyser to explore the differences by looking for the code that performed this activity, and determining if this change is an example of a real difference or whether the two exploits are the same.

Table 5. The web-based exploit experiment groups

Group	Log files	Group	Log files	Group	Log files
1	**3**	**4**	**8**	**7**	**6**
2	**3**	**5**	**7**	**8**	**5**
3	**8**	**6**	**3**	**9**	**3**
10	**3**	**11-77**	**1 in each group**		

7 Conclusion

The analysis tools discussed in this work can convert Capture-HPC log files into state machine files. Moreover, we have used the state machine as a better and easier way to present malicious website activity found using a client honeypot. A clustering algorithm is then applied to the outcome of this, in order to group similar files together to reduce the time taken to analyse malicious website activity, which can occur on a user's system without his/her knowledge. A graphical user interface is provided to enable a user to more easily interact with the generated clustering outcomes, and trace this activity back to the log files.

The current automated state machine supports three states: file system, processes and registry as Capture-HPC does not yet support network or port monitoring. However, Capture-HPC is currently being developed to support network monitoring, and when this is complete it will be added to our system.

Our future works involves investigating the following: (1) The behaviour groupings identified by the clustering algorithm may be stored in a database. This contains all the state machine chains, the URL, the state machine stamps and the behaviour family groups obtained from the clustering algorithm. (2) The behaviour families of malicious web pages could also be used to predict which family a particular exploit falls under. The algorithm will check the current web page and try to predict the subsequent activities of the malicious web page. If the algorithm can predict the next activity likely to take place when interacting with a web site, and determine that it is malicious, then it can be used to block the web page and protect the victim's machine. The aim of the prediction algorithm is to act as an intrusion detection system (IDS) by using the behaviours and activities collected from state machine chains.

References

1. Spitzner, L.: Honeypots: Tracking Hackers, 1st edn. Addison-Wesley Professional, Reading (2002)
2. Witten, J., Frank, F.: Data Mining: Practical Machine Learning Tools and Techniques, 2nd edn. Morgan Kaufmann, San Francisco (2005)
3. Windows Workflow Foundation, `http://msdn.microsoft.com/en-us/netframework/aa663328.aspx` (Accessed August 5, 2009)
4. Capture-HPC (2008), `https://projects.honeynet.org/capture-hpc` (Accessed February 1, 2009)
5. Song, I.-Y., Eder, J., Nguyen, T.M. (eds.): DaWaK 2007. LNCS, vol. 4654. Springer, Heidelberg (2007)
6. Vijaya, P., Murty, M., Subramaniana, D.: Efficient bottom-up hybrid hierarchical clusteringnext term techniques for protein sequence classification. Pattern Recognition Bioinformatics 39(12), 2344–2355 (2006)
7. Manning, C., Raghavan, P., Schütze, H.: Introduction to Information Retrieval. Cambridge University Press, Cambridge (2008)
8. Provos, N., Holz, T.: Virtual Honeypots: from Botnet Tracking to Intrusion Detection, 1st edn. Addison-Wesley Professional, Reading (2007)
9. Seifert, C.: Know Your Enemy: Behind the Scenes of Malicious Web Servers (2007), `http://honeynet.org/book/export/html/181` (Accessed March 1, 2009)

LocPriS: A Security and Privacy Preserving Location Based Services Development Framework

Gareth Ayres and Rashid Mehmood

School of Engineering, Swansea University, Swansea, SA2 8PP, UK
{g.j.ayres,r.mehmood}@Swansea.ac.uk

Abstract. With the ever increasing pervasiveness of devices with functionality to provide location based services comes the increased importance and reliance upon those services to provide user privacy and security. Many techniques to facilitate privacy and security in mobile and fixed networks have been developed, but surveys of user's show that this area still has a lot of work left to do to satisfy privacy fears and help developers of such services to choose the best techniques to use. In this paper we propose a security and privacy preserving location based services development framework. The framework will allow for future development, visualisation, comparison and analysis of location based services that preserve security and privacy in order to improve user confidence in such technologies.

Keywords: Location Based Service; Privacy; Security; Mobility; Visualisation; Wireless Networks.

1 Introduction

The recent explosion in popularity of mobile devices such as the Apple iPhone and Google Andriod phones has accelerated the use and deployment of Location Based Services (LBS). Such services are keen to help users make use of their location and mobility data in order to enhance or provide functionality. However, this functionality can sometimes come at the cost of privacy.

While LBS's are usually developed with security and privacy measures in mind, these are not always sufficient. Many LBS's come with the ability to turn location tracking on/off, while some come with only assurances of privacy with little or no technical explanation or justification of how privacy is achieved, or to what degree. The level of granularity of location data relative to the level of granularity required for a service to be functional is often never expressed or made transparent enough to users. Previous surveys of the value of location privacy to users have revealed the importance of privacy with regard to location, which is discussed further in this paper along with the result of a local study of user opinions.

Keeping the importance of location privacy and the security of techniques used in mind we have identified the need for a framework to help develop, visualise, test, compare and analyse existing and new ways of preserving location privacy and security in LBS's.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 566–575, 2010.

We propose a modular framework that will attempt to address these issues and provide an open source solution to help future researchers as well as developers and users of LBS's.

The remainder of this paper is broken down as such: Section 2 covers background material on LBS's and location privacy techniques. Section 3 provides an overview of related works in the area of network and mobility simulators and frameworks, and. discuses some previous studies of user's opinions on location privacy. Section 4 shows the results of a survey of 502 users carried out at Swansea University and the results discovered. Section 5 introduces our proposed LocPriS framework and its modular architecture while Section 6 will conclude the paper and highlight the future direction of our work.

2 Background Material

2.1 Origins of LBS

One of, if not the first, LBS's to be developed was the E911 system developed by Telecom operators in the early 1970's in collaboration with the US Governments Federal Communications Commission. These telephone systems allowed emergency calls made in some states in the United States to be routed to the appropriate emergency services call room. This is a simple example of a LBS with low granularity of location data, but later this service was enhanced to comply with additional regulations and new technologies such as mobile phones. This resulted in improved granularity of the location data as well as the better functionality of the service by displaying the data on maps [1].

Some of the first LBS's to be developed separately from the E911 system were developed under the vision of context-aware computing by Olivetti Research Ltd. ORL developed a LBS in 1992 that made use of the Active Badge system to inform receptionists where to forward phone calls too. This allowed receptionists to forward calls to the nearest phone to the recipient [2].

The development of the Active Badge system in Cambridge led to an increase in research in LBS's that made use of indoor localisation systems and users wearing a small localisation device. The next major development in indoor localisation did not appear until 1999 when AT&T developed the Active Bat system. Soon after in 2000 the cricket system [3] was also developed with much success.

With significant research in localisation techniques and other issues relevant to LBS's focusing on indoor systems throughout the 1990's until early 2000 the next wave of research was kick started by Microsoft with a paper they published called RADAR [4], which details a number of methods of performing localisation through wireless LAN's. This paper claimed the use of WLANS can provide localisation with an accuracy of 2-3 meters using existing 802.11 equipment.

The direction of research began to shift during the early 2000's from using indoor location tags to existing infrastructure to provide LBS's. The increase in popularity of GPS during this time as a result of the increase in its accuracy to 20meters from 100meters in 2000 also helped change the direction of research.

In 2005 the combination of the increase in availability of mobile phones with built in GPS, 3G networks and the arrival of Web 2.0 technologies resulted in a revival of LBS research and development. The combination of mobile phones that were location aware through WiFi signatures and AGPS, social networking sites booming and online GIS mapping systems has resulted in a surge of activity in the area of LBS's.

2.2 Location Privacy and Security

Most users of networks in large institutions such as Universities are happy to use computers to browse the internet and communicate with friends and colleagues without considering how private that activity is. They are likely unaware that their internet browsing activity maybe being logged and the chat communications is being sent unencrypted and open to interception by network administrators or other agencies or hackers.

A side effect of the functionality of some LBS's is that users become consciously aware of the fact that their location data is what drives the service and that data is being controlled by a computer somewhere.

Privacy is considered a fundamental human right by the Universal Declaration of Human Rights and most democracies around the world [5] and the security of location data and users privacy must be taken seriously.

One of the first considerations is the granularity of the location data. The granularity of meters could provide more information about a user than the granularity of kilometres [6], depending upon the context of the service. Granularity alone does not provide any real privacy, and is vulnerable to correlation attacks as well as inference and assumptions attacks based on historical data.

The fundamental problems of location data storage and visualisation can be addressed by annonymising the data using pseudonyms. Pseudonymity provides anonymity to location data while maintaining a relationship between the data that is used to help the LBS function. Recording a pseudonym and location as a location data record allows for the movement of a node to be tracked while removing any identifiable data from the record [7]. This adds a level of security to the system that would protect a user if the data was stolen or misused. However it does not offer complete privacy as a user's identity could still be inferred from the history of a nodes movement in some cases.

One solution to this problem is the addition of dummy nodes. Dummy nodes add a level of 'noise' to the LBS that does not affect the quality of the service but helps remove the ability of a possible attacker to infer the identity of a node based on the history of a nodes movements [8, 9]. Other possible solutions to this problem are temporal and special cloaking along with silent periods [9].

Another possible technique to add privacy is the use of mix zones. Mix zones provide a trusted middleware that provides anonymised location information to third-party applications by defining spatiotemporal zones where all users in that zone have their pseudonyms changed upon entering and leaving, therefore providing a new set of anonymity [10].

Configurable privacy preferences have also become common techniques used to provide privacy in LBS's, giving users the ability to control the volume, granularity or accuracy of location data they reveal.

3 Related Works

There exist a number of simulation and visualisation tools, with different objectives, which come close to meeting some of the objectives of our framework. The tools, simulators and frameworks surveyed can be broken down into two distinctive areas, network simulation and mobility visualisation tools.

3.1 Network Simulations

There are a number of established simulation tools available which address the problem of network/ wireless network simulation. Many of the simulation tools provide the ability to develop other tools built upon them and could be used to fulfil part of the needs of the simulation module of our framework.

There are commercial as well as open source simulators, all with different qualities. One popular open source simulator is NS-2 which is a discrete event based simulator that is very popular in academia for network simulation. NS-2 does not cater well for mobility modelling on its own, and has little in the way of visualisation functionality. There have been developments which attempt to add these features to NS-2, such as iNSpect.

GloMoSim (Global Mobile Information System Simulator) is a network protocol simulation tool that simulates wireless and wired network systems. It is designed using the parallel discrete event simulation capability provided by Parsec. QualNet is a commercial network simulator with many libraries and components. It supports visualization of simulations, and has support for some mobility patterns.

SWANS is a scalable wireless network simulator built on the parallel discrete event based java JiST platform. It makes use of virtual machines to improve speed of simulation. The Georgia Tech Network Simulator (GTNets) also provides limited support for mobility and provides particular attention to protocol simulation and analysis.

3.2 Mobility Visualisation Tools

While most network simulators have added functionality to allow wireless network simulation, some have also added mobility simulation. Mobility simulation is useful for wireless networks of many types, including WLAN, adhoc and sensor networks. A detailed survey paper on the area of mobile area network simulation is [11].

There are a number of mobility simulation tools such as Mobitools, MobiREAL and MoViTo which all aim to simulate mobility in mobile area networks through the use of mobility patterns. Many have functionality for Vehicular mobility simulation, such as Mobitools, while some aim more at human mobility using probalistic rule-based models such a MobiREAL. Some simulators make use of virtual environments while some make use of GIS systems. Most tools make use of 2D while some also make use of 3D such as ViTaN.

3.3 The Value of Location Privacy

While it is mostly accepted amongst researchers that privacy needs to be built into future technologies regardless of some user's value of it, a number of studies of user opinions on the value of location privacy have been performed. Cambridge University run a survey of their computer science undergraduate students to try to measure the monetary value students place on their location information. The survey was in the form of questions and an auction to determine the value students place on their location data. 74 students filled in the questionnaire.

The results show that students valued their privacy at a median bid of £10. They then doubled that when commercial interest was mentioned. It also showed that students who travelled outside of Cambridge valued their privacy more than those who did not travel far [12].

One of the authors of the Cambridge survey went on to question a sample of over 1200 people from five EU Countries. This also followed the form of an auction to determine the value placed on a month's location data. The survey produced a median value of £20 for a month's location data, but did not find the same correlation between users who travel more and location data value [13].

A survey on location privacy and social networks has been carried out by Intel Research which provided PDA's to 16 non-technical participants in order to retrieve information on how the participants value location privacy when disclosing information to friends, family and colleagues [14].

The Westin/Harris Privacy Segmentation Model was used to classify participant's privacy. The paper found that people were fairly specific about their location 77% of the time. Who was requesting the information had the strongest influence on the participants willingness to disclose. Why a request for information was made was also an important influence. Participants also provided more granular location information to people who were relatively close to them.

Another survey used two distinct types of location data in order to access which is a greater privacy concern in relation to location based services. Location-tracking and position aware location data. The authors claim that users are more concerned by location tracking data than position aware data. They also claim that the sampled users were positive towards location based services as long as they perceive them to be useful [15].

One survey studies the results from a questionnaire used to determine the importance of two factors – inquirers identity and the users situation at the time of inquiry. It is found that these two factors directly determine the accuracy of disclosed information [16].

4 LBS Privacy Questionnaire

An online questionnaire was sent via email to all users of the Swansea University Wireless Network on 18 March 2010 which ran for 2 weeks. The users of the wireless network consist of members of staff, students and some volunteers. The survey consisted of questions regarding the wireless network, but had a section on location based services and privacy. The questionnaire was anonymous and voluntary, with no incentive offered to participants. 502 people completed the questionnaire, but as answering each question was optional, there are different numbers of responses to some questions.

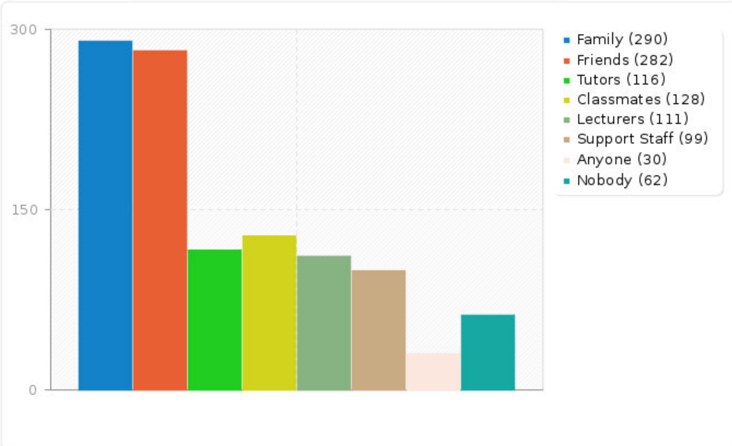


Fig. 1. Who would you share location information with?

4.1 Questionnaire Results

There was a fairly even split between participants of the questionnaire on who had used a LBS and who has not. 43% (190) had not, while 41% (181) said they had. Of the participants who had used a LBS, GPS directional assistance was the most popular service 30% (131) with mobile phone applications following with 27% (119).
When asked '*Is privacy an issue you would consider when using a Location Based Service*' the response was that 47% (203) said '*yes*' it was, with 35% (151) saying '*no*' it was not. When asked who they would be happy to share thier location data with (allowing for multiple choices), the response was unsuprising that family (66%) and friends (64%) was selcted most commonly with the extreme options of sharning with nobody (14%) and anybody (14%) being less popular.

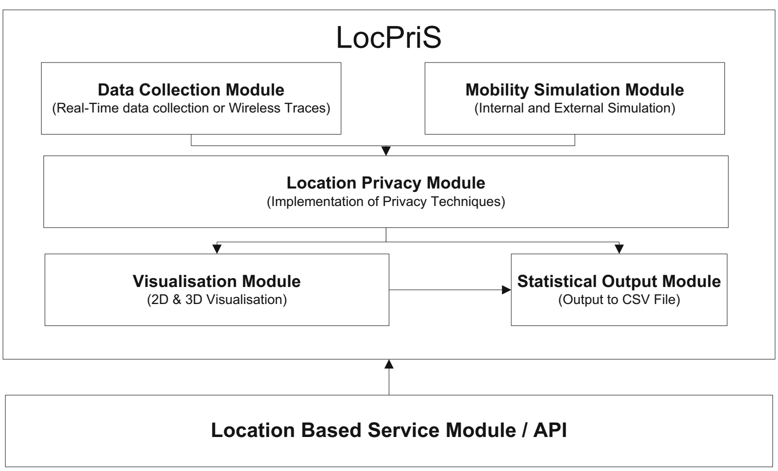


Fig. 2. The modular design of the LocPriS Framework

Interestingly, when asked '*If technology could guarantee the privacy of your location, would this encourage you to use a Location Based Service?*' the response was that 63% (277) of participants felt it would encourage them with only 9% (40) saying it would not.

5 LocPriS Framework

We propose a modular extensible framework that will provide tools for the development, analysis, comparison and visualisation of LBS's that preserve privacy and security. The framework will also assist in the development and testing of Location Based Services through an exposed API.

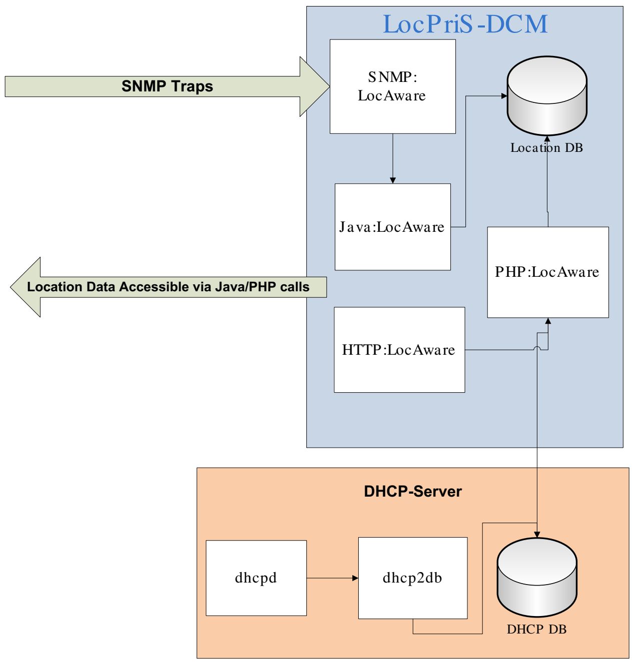


Fig. 3. LocPriS Data Collection Module

5.1 LocPriS: Data Collection Module

The Data Collection Module will provide for two main inputs of data:

- **Wireless traces** – Wireless traces are available from a number of archives and institutions, including Dartmouth (Crawdad) and USC. These archives are of different sizes, and do not all contain location linking data, but through linking techniques can used.

- **Real-time data –** Data from the Swansea University wireless network is being recorded in real time through the use of a SNMP traps and custom java/PHP programs. See figure 3 for a system design diagram.

5.2 LocPriS: Mobility Simulation Module

Although the use of real data will be beneficial, the use of simulated mobility data will allow for additional testing and comparison. There are a number of mobility simulators that simulate models such as the random waypoint and walk models which could be used as a single source of data or in conjunction with the implementation or other mobility models in the framework.

5.3 LocPriS: Visualisation Module

Through the use of Java, OpenGL and the Java Monkey Engine it is possible to visualise mobility in 2D or 3D. Visualisation of mobility and then the visualisation of the application of privacy and security techniques will allow for contrasting views of some of the strengths, weaknesses and characteristics of differing techniques. This will aid designers of LBS's as well as users who wish to quickly see the implications of using LBS's.

Through the use of open mapping data it will also be possible to overlay graphical maps on to 2d visualisations as well as allowing for the importing of user designed floor maps and layouts. See figure 4 for a screen shot of the visualisation module.

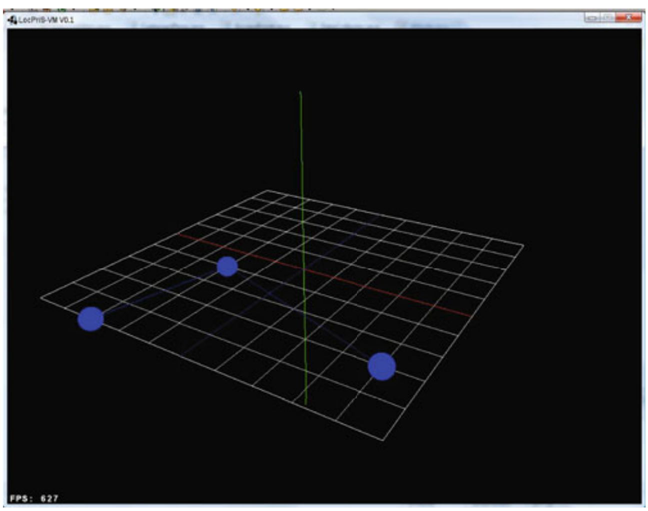


Fig. 4. LocPriS Visualisation Module

5.4 LocPriS: Location Privacy Module

As described in the previous sections, there have been a number of privacy and security techniques developed to improve privacy and provide security in LBS's.

These techniques will be implemented in this module so they can be applied to sets of location/mobility data. This module will also allow for the development of new novel techniques.

5.5 LocPriS: Statistical Output Module

The result of visualisation and application of privacy and security techniques will produce statistical output. This will mainly consist of numerical representation of node locations and movements. This data will be outputted in a CSV file for easy analysis.

5.6 LocPriS: Location Based Services

Through the use of an exposed API it will be possible to develop, test and analyse privacy techniques as well as new services through the use of the LocPriS framework. As has been shown [17], security can be considered a LBS and can be built as a service on the LocPriS framework.

6 Conclusion

Location based services have come a long way since the early inceptions by researchers and service providers, along with significant developments in privacy and security techniques to complement them. While these techniques are bringing some confidence and assurances to users, results from our and other user surveys and questionnaires have shown that there is still work to be done to provide greater privacy and security, and to build the trust of users of LBS's through continued transparency of the application and functionality of such techniques.

User surveys along with the continued growth of the number of devices with LBS functionality have shown that demand for such services is substantial and growing. It is important that the growth of privacy techniques follow this trend to keep up with such growth in order to ensure the continued trust of users in new and exciting technologies.

In this paper we have outlined a framework we believe will help to enhance the development of LBS's that preserve privacy and security. Development of the framework is still in the early stages, with work left to do on each of the modules. The continued development of the framework will result in the production of results from analysis and comparison of privacy and security techniques along with the possibility of development of new techniques and LBS's.

References

1. Woody, L.A., Acker, L., Curran, W., Susi, S., Velazquez, M.: The Impact of the FCC's Position on Wireless E911
2. Hightower, J., Borriello, G.: A Survey and Taxonomy of Location Systems for Ubiquitous Computing, pp. 57–66 (2001)
3. Priyantha, N.B., Chakraborty, A., Balakrishnan, H.: The Cricket location-support system. In: Mobile Computing and Networking, ACM, Boston (2000)

4. Bahl, P., Padmanabhan, V.N.: RADAR: an in-building RF-based user location and tracking system (2000)
5. Nations, U.: Universal Declaration of Human Rights, General Assembly Resolution 217 A (III) (1948)
6. Grlach, A., Heinemann, A., Terpstra, W.W.: Survey on location privacy in pervasive computing. Privacy, Security and Trust within the Context of Pervasive Computing (2004)
7. Chaum, D.: Untraceable electronic mail, return addresses, and digital pseudonyms. Communications of the ACM, 84–88 (1981)
8. Report, F.: Surviving the privacy revolution.: Deering, S., ICMP Router Discovery Messages, RFC 1256 (2002)
9. Gruteser, M., Grunwald, D.: Anonymous Usage of Location-Based Services Through Spatial and Temporal Cloaking. In: Mobile Systems, Applications and Services, ACM, New York (2003)
10. Beresford, A.R., Stajano, F.: U.o.C.: Location Privacy in Pervasive Computing. In: Pervasive Computing, p. 10. IEEE, Los Alamitos (2003)
11. Kurkowski, S., Camp, T., Colagrosso, M.: MANET simulation studies: the incredibles. SIGMOBILE Mob. Comput. Commun. 9, 50–61 (2005)
12. Danezis, G., Lewis, S., Anderson, R.: How much is location privacy worth (2005)
13. Dan, C., George, D.: A Study on The Value of Location Privacy. In: Privacy in the Electronic Society, pp. 109–118. ACM, New York (2006)
14. Consolvo, S., Smith, I.E., Matthews, T., LaMarca, A., Tabert, J., Powledge, P.: Location disclosure to social relations: why, when, & what people want to share. In: Proceedings of the SIGCHI, pp. 81–90. ACM, New York (2005)
15. Barkhuus, L., Dey, A.: Location-Based Services for Mobile Telephony: a study of users' privacy concerns. In: Proceedings of the 9th IFIP TC13, p. 709, p. 712 (2003)
16. Lederer, S., Mankoff, J., Dey, A.K.: Who wants to know what when? privacy preference determinants in ubiquitous computing. In: CHI 2003, pp. 724–725. ACM, New York (2003)
17. Localization to Enhance Security and Services in Wi-Fi Networks under Privacy Constraints. In: Communications Infrastructure. Systems and Applications in Europe, LNICST, vol. 16, pp175–188 (December 2009)

Controlling Real World Pervasive Environments with Knowledge Bases

Atif Alvi[1], Zubair Nabi[1], David Greaves[2], and Rashid Mehmood[3]

[1] Computer Science Department, Lahore University of Management Sciences, Pakistan
{atif.alvi,zubair.nabi}@lums.edu.pk
[2] Computer Laboratory, University of Cambridge, UK
david.greaves@cl.cam.ac.uk
[3] School of Engineering, Swansea University, UK
r.mehmood@swansea.ac.uk

Abstract. With the ever increasing popularity of mobile devices and the widespread deployment of different types of networks, pervasive environments have truly arrived. Although such environments are already being utilised to provide services based on the available data, the modelled domain knowledge is generally not formally verified for consistency and applicability within and across domains. This paper presents our knowledge-based verification and control approach and shows its need and effectiveness. In our system model, policies in the form of rules are authored at different levels of abstraction at different layers in the policy hierarchy. Top level rules are written in first-order predicate logic based on a formal ontology of the domain. Detection and resolution of feature interactions (or conflicts) is an important task in the policy hierarchy. Our running example is the pervasive mobile learning environment of a university campus. Our approach is easily and equally applicable to other pervasive environments.

Keywords: Pervasive computing, ontology, mobile learning, rule-based control, feature interaction, knowledge base, location-based services.

1 Introduction

The vision of pervasive computing is to have a wealth of embedded processors and applications serving user needs in an imperceptible way [1]. Computing power is now abundantly available but a large gap still exists between user needs and software fulfilling those needs.

Knowledge bases using ontologies are an attractive choice to model the pervasive environment: an ontology's flexible structure can best represent the dynamic environment that has changing data and relationships. In addition, the formal basis of an ontology makes it amenable to automated reasoning. Such robust modelling also expresses the desired functionality in a device-independent way.

In general, an ontology is used as a domain model to share, reuse, analyse and better engineer knowledge [2]. A particularly useful benefit is the separation of

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 576–585, 2010.

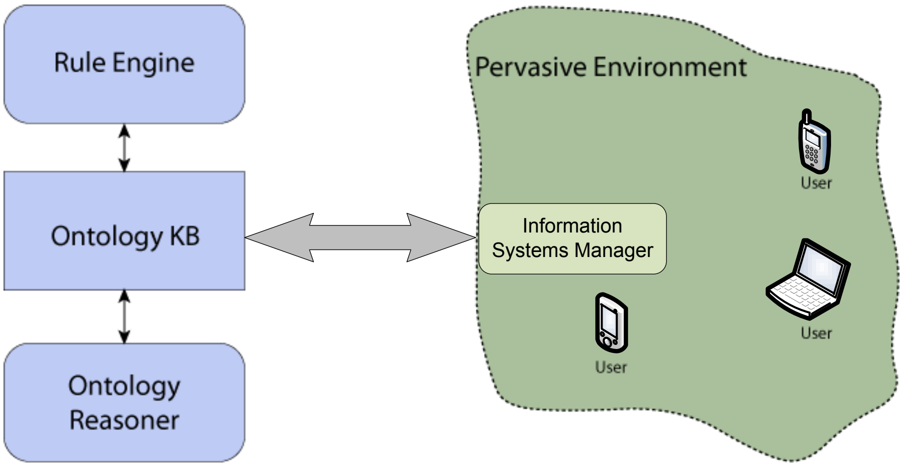


Fig. 1. Architectural overview

domain knowledge from operational knowledge, which enables the development of applications without requiring specialized knowledge of a particular domain. This also promotes reuse of ontologies. Ontologies model data at a higher level of abstraction than specific database designs [3] and can be used to export, translate, query, and unify data across independently developed systems and services [4]. Hence, an ontological model serves as a convenient data reflection API for programmers while avoiding the rigid schema of databases, thus allowing systems to evolve better. Pervasive domains are diverse by definition and their requirements of *interoperability* and *evolvability* make them suitable for modelling by ontologies.

Ontologies are made machine-processable when expressed in knowledge representation languages with well-defined formal semantics. These include OWL (Web Ontology Language) [5], KIF (Knowledge Interchange Format) [6] and RIF (Rule Interchange Format) [7], among others. Adding rule axioms to the ontological substrate enables a control plane: policies can then be defined by writing rules to shape behaviour in the domain, e.g. by using SWRL (Semantic Web Rule Language) [8]. Sets of rules can be written by experts in various subclasses (subdomains) of the ontology of the larger domain and any potential conflicts among these sets can be caught by supervisory rules operating at a higher level in the ontology hierarchy. We discuss this in Section 3.2. Figure 1 depicts the framework of our approach in which an ontological KB of the domain or pervasive environment is verified by a reasoner and the rules in the KB are verified by a rule engine. An information systems manager then asserts these rules in the pervasive domain comprising heterogenous devices and subdomains of varying smartness.

Web-enabled portable devices are a component of pervasive infrastructure as they are owned by the majority of population these days. Mobile learning or m-learning [9,10] is an emerging paradigm—a type of e-learning or distance learning—that uses mobile technologies to deliver educational content to users who are mobile. In this paper, we take the motivating example of m-learning in a pervasive campus environment. We focus on wireless connectivity as the means of delivering educational material to students on campus and discuss the various associated issues and their resolution. Moreover, when users are mobile, knowing where they are is important for providing context-based and location-based services, including those that enhance their learning experience. Doing so in an imperceptible and conflict-free way is a challenge, as is the ability to write a rule that can be implemented on a variety of disparate devices. We tackle all of these challenges through our knowledge-based formal verification approach.

In Section 2 we describe a mobile learning enabled, wirelessly connected campus environment. In Section 3, we present our technique of defining and verifying a rule-based policy and resolving any conflicts and illustrate through several examples. We also discuss there the implementation of policies at different levels. We conclude in Section 4.

2 Pervasive Campus Environment

As a real world deployment example of our approach, we choose the campus environment of the Lahore University of Management Sciences[1] (LUMS), Lahore, Pakistan. In particular, we consider the mobile learning approach enabled through a wireless local area network (WLAN) on campus and study how having an ontological formal approach can help enable a pervasive environment. Some of the policies and techniques that we suggest are already implemented in high-end WLAN controllers[2], but we discuss in Section 3.3 how to implement this in a more device-independent way. We do not discuss the particular technique to determine location using WiFi. Tracking could be performed based on a signal strength map or, less accurately, on the proximity to an access point (AP), or triangulation (if several APs are in range) [11,12].

LUMS is a medium-sized university with an urban campus. The campus is served by wired and wireless local area networks, providing intranet and Internet services to students and staff. LUMS has two learning management systems (LMS) or virtual learning environments (VLE), viz. Sakai and Zambeel. Sakai (`http://sakaiproject.org/`) is a free and open source VLE similar to products like Blackboard and Moodle. Zambeel is a PeopleSoft product, now owned by Oracle®. Some of the features of these software are to allow students and faculty to share educational material, send group emails and have discussion forums. Additional m-learning facility is provided by Panopto (`http://panopto.com/`) to host lecture videos that can be streamed and podcast.

To enable these mobile learning applications on the handheld devices and laptops of students and faculty, high-speed data services, like EDGE, are available

[1] `http://www.lums.edu.pk/`

[2] E.g. Cisco® Wireless Location Appliance: `http://www.cisco.com/en/US/products/ps6386/`

on mobile phone networks, and wired and Wi-Fi connectivity is provided by the university. At LUMS, Internet services to the computing labs, offices and hostels are via three different providers with auto-failover. This bandwidth is shared with the WLAN. Intranet services to the computing labs, offices and hostels are over high speed gigabit Ethernet and fibre channels. There is wireless connectivity in many parts of the campus, though the coverage is not complete. A Cisco® WLAN controller and various Cisco®access points are deployed. The main academic block (housing many departments and lecture halls), cafeteria, administration block, and the Rausing executive training centre are all covered.

The current bandwidth allocation policy on LUMS campus gives the highest share of the bandwidth to faculty, followed by staff and students in the ratio $2 : 1.5 : 1$, respectively. This policy is enforced uniformly throughout campus. We can think of situations where this policy might be relaxed to improve overall service. For example, when a faculty member is in the cafeteria or in a classroom, her allocated bandwidth can be reduced as compared to when she is in her office or lab. A location-aware application can easily track a user and allocate bandwidth based on the user's current location. Similarly, it might be a good idea to generally restrict the bandwidth of students in class to help them concentrate on lectures.

3 Policies and Conflicts

3.1 Defining and Verifying Policies

The policies controlling m-learning services on campus are best effected through the use of an ontological knowledge base that stores the various entities, their attributes and relationships. We can then add a rule layer atop it that defines the behaviour of the system. We now describe one such system that we have previously used in an automated home environment [13]. It uses an ontology editor for expressing ontologies in OWL with a facility to write rules in SWRL. Both the ontology and the rules are verified by their own reasoners.

SWRL [8] was developed as a rule language for the semantic web that could be processed by different reasoning engines [14,15]. Declarative rules give the administrators means for regulating system behaviour, appropriate for many forms of specification. Rule engines vary in their performance and memory use but different ontology editors provide flexibility by allowing to choose the best rule engine for a given domain. One such ontology editor and knowledge base framework is Protégé (`http://protege.stanford.edu/`).

We developed and stored the ontology of a campus environment using Protégé. It stores all the concepts, their attributes and relationships in OWL DL, which is a decidable fragment of OWL. Specifically, we use the term *ontology* to mean a data model that is used to represent different concepts or entities in a domain of knowledge along with their relationships. Stated mathematically, we define an ontology $\mathcal{O}$ as a quadruple $(\mathcal{C}, \mathcal{I}, \mathcal{P}, \mathcal{R})$ where $\mathcal{C}$ is the set of classes or concepts in the domain of interest, $\mathcal{I}$ is the set of instances or individuals of classes, $\mathcal{P}$

is the set of properties or roles of classes, and $\mathcal{R}$ is the set of relations between classes, e.g., isA, subClassOf, sameClass. Within a Campus domain, we can have a class Location with an instance Courtyard that has the property hasAccessPoint.

We are mainly concerned with two reasoning tasks for our ontological knowledge base:

- *Validation*: To detect any contradictions in the modelled information in the ontology.
- *Deduction*: To derive implicit conclusions from the information in the ontology, capturing the notion of logical consequence.

An accompanying ontology reasoner (e.g.RacerPro[3]) performs the above reasoning tasks. Thus, our ontology is formally verified to be consistent and captures all the implicit relationships. We then add a higher level rule layer to this using SWRL, as described below.

In Protégé, SWRL rules are stored as OWL individuals with their associated knowledge base. All the OWL instances and SWRL rules are automatically mapped onto facts and rules of the underlying reasoning engine, such as Jess.[4] Any new classifications or properties resulting from executing the Jess inference engine are written back to the OWL knowledge base. System administrators can subscribe to be notified of any changes in the rules and the knowledge base that they deem require immediate attention. Other changes can be taken care of in a lazy style of updating.

SWRL extends OWL axioms to include Horn-like rules.[5] Thus, one can use OWL concepts to write Horn-like rules in SWRL to reason about OWL individuals. Variables in SWRL are represented by prefixing them with a question mark. Rule atoms can be of different forms, e.g., $\mathsf{C}(\mathsf{x})$, $\mathsf{P}(\mathsf{x},\mathsf{y})$, $\mathsf{sameAs}(\mathsf{x},\mathsf{y})$, $\mathsf{builtIn}(\mathsf{r},\mathsf{x},...)$ or $\mathsf{differentFrom}(\mathsf{x},\mathsf{y})$, where C is an OWL description or data range, P is an OWL property, r is a built-in relation, and x and y are either variables, OWL individuals, or OWL data values, as appropriate [8].

We will now compose a few rules in SWRL to define the desired behaviour of our network. To reiterate, we use wireless connectivity as the network level enabler of higher level mobile learning services. We begin with the rule that assigns a certain bandwidth to a faculty member, which is assumed to be more than that of other users.

$$\mathsf{Faculty}(?\mathsf{f}) \rightarrow \mathsf{hasBW}(?\mathsf{f},2) \tag{1}$$

Rule 1 assigns a maximum symbolic bandwidth of 2, out of the total available bandwidth, to a faculty member denoted by the variable f. The bandwidth 2 only symbolises the proportion that should be assigned to faculty. The translation of this into actual mega bits per second is done at the lower device level.

[3] Renamed ABox and Concept Expression Reasoner: `http://www.racer-systems.com/`

[4] Java Expert System Shell: `http://www.jessrules.com/`

[5] A *Horn clause* is a disjunction of literals with at most one positive literal. Such a clause can be written as an implication with the positive literal as the consequent.

The next rule, rule 2, restricts the bandwidth available to students while in class to nil by assigning $\mathsf{0}$ to the hasBW relation, which denotes the currently available bandwidth to the student.

$$\mathsf{Student}(?\mathsf{s}) \wedge \mathsf{ClassRoom}(?\mathsf{c}) \wedge \mathsf{hasLocation}(?\mathsf{s}, ?\mathsf{c}) \rightarrow \mathsf{hasBW}(?\mathsf{s}, 0) \quad (2)$$

Next, we define a rule to assert that if several students are in a classroom and viewing the same lecture on Panopto, they should be able to share an annotation tool to see each other's annotations:

$$\begin{aligned} &\mathsf{Student}(?\mathsf{s}) \wedge \mathsf{ClassRoom}(?\mathsf{c}) \wedge \mathsf{hasLocation}(?\mathsf{s}, ?\mathsf{c}) \\ &\quad \wedge \mathsf{VideoLecture}(?\mathsf{v}) \wedge \mathsf{usingPanopto}(?\mathsf{s}, ?\mathsf{v}) \\ &\quad \rightarrow \mathsf{shareAnnotation}(?\mathsf{s}, \mathsf{true}) \end{aligned} \quad (3)$$

Rule 4 below is used to suggest to a user a location where higher bandwidth is available as compared to the available bandwidth in the user's present location. Here, the SWRL built-in $\mathsf{greaterThan}$ is used to compare the available bandwidths c and b.

$$\begin{aligned} &\mathsf{User}(?\mathsf{u}) \wedge \mathsf{hasBW}(?\mathsf{u}, ?\mathsf{b}) \wedge \mathsf{hasLocation}(?\mathsf{u}, ?\mathsf{x}) \\ &\quad \wedge \mathsf{Location}(?\mathsf{y}) \wedge \mathsf{hasBW}(?\mathsf{y}, ?\mathsf{c}) \\ &\quad \wedge \mathsf{swrlb} : \mathsf{greaterThan}(?\mathsf{c}, ?\mathsf{b}) \\ &\quad \rightarrow \mathsf{suggestLocation}(?\mathsf{u}, ?\mathsf{y}) \end{aligned} \quad (4)$$

This is a dynamic optimisation problem and we are working on a predictive control methodology to inform users beforehand where congestion is likely to occur.

3.2 Resolving Feature Interaction

What if there is a conflict between two rules or requirements? Consider a scenario where students are in a classroom. According to a previous rule (rule 2), the students should have no bandwidth available. But what if the class needs to have LAN and WLAN connectivity? This situation can also arise in a distance learning scenario: virtual universities usually require that students in a particular area gather in a university-provided facility to view some lectures together in the presence of a supervisor. The students can be using a mobile learning application like Panopto or Synote (`http://www.synote.org/`) to view and annotate the lectures on their laptops and desktops.

This conflict between two requirements is a form of *feature interaction*. The term feature interaction has its roots in the telecommunications and was coined by Bellcore in the early 80's [16]. In call control systems, feature interaction can be defined as [17]:

> Components enhancing a basic service are referred to as features and incompatibility between features or the influence of one feature on the behaviour of another is referred to as feature interaction.

We can think of a feature as a unit of functionality—a facet of behaviour—of a system. This means that the behaviour-defining rules can represent features and have the potential to interact in a conflicting manner. Hence, pervasive computing environments like LUMS campus certainly represent domains where feature interaction is likely to take place. Some feature interactions are harmless but the majority is severely damaging to system development and user expectations. It is a major research challenge to predict, identify and resolve feature interactions.

We can resolve the conflict between rule 2 and the opposite requirement during e-learning classes by formulating another standing rule that uses the SWRL temporal built-in during to see if it is currently the e-learning class time:

$$\begin{aligned}&\mathsf{Student}(?\mathsf{s}) \wedge \mathsf{ClassRoom}(?\mathsf{c}) \wedge \mathsf{hasLocation}(?\mathsf{s}, ?\mathsf{c})\\ &\wedge \mathsf{temporal} : \mathsf{during}(\mathsf{CurrentTime}, \mathsf{eLearningTime})\\ &\rightarrow \mathsf{hasBW}(?\mathsf{s}, 1)\end{aligned} \tag{5}$$

We now describe another feature interaction situation. Some students at the university also work as TAs and RAs. The overarching rule for students, as discussed before in rule 2, is that they would not enjoy any connectivity while in class. A TA/RA is a student but has a dual account as a university staff member as well and might need network access as staff when none is available as a student. The TAs/RAs can be given this preferential treatment through the following rule, which resolves this feature interaction:

$$\mathsf{Student}(?\mathsf{s}) \wedge \mathsf{Staff}(?\mathsf{s}) \rightarrow \mathsf{hasBW}(?\mathsf{s}, 1.5) \tag{6}$$

where 1.5 is the ratio of bandwidth provided to staff.

Finally, it is possible that two rules assign different values to the same property of an entity. This is another form of feature interaction. For example, consider rule 7 below:

$$\mathsf{User}(?\mathsf{u}) \wedge \mathsf{hasLocation}(?\mathsf{u}, \mathsf{Cafeteria}) \rightarrow \mathsf{hasBW}(?\mathsf{u}, 1.25) \tag{7}$$

It defines that all users when in the cafeteria get the same bandwidth ratio 1.25. Note that this is in conflict with rule 1 which assigned a higher bandwidth to faculty. This rule is also in conflict with the previously assigned bandwidths in the system (through the hasBW property) to students and staff. One disadvantage of SWRL, like OWL, is that it works under the *open world assumption*, which is closely related to monotonicity. Hence, we cannot modify or retract information from an ontology using SWRL. This means that we can have two rules that assign different conflicting values to the same variable. The variable will end up having two values as it is not possible to overwrite existing knowledge [8,18]. The way to detect this feature interaction is to make the property hasBW functional.[6] Protégé has the facility to define properties so. This way, an OWL reasoner (e.g. RacerPro) would indicate an inconsistency when run on this ontology. It would discover hasBW for faculty having two values 2 and 1.25 and would raise a flag. Other ontology models and rule systems have their own mechanisms to detect such conflicts and make amends.

[6] A *functional property*, also called a *single valued property*, is a property that can have only one (unique) value y for each instance x.

Another way to resolve this feature interaction is to have priorities assigned to rules. Rule 7 can be assigned a higher priority than rule 1 so that whenever faculty members are in the cafeteria, they get the same bandwidth as others.

Capturing the environment in a hierarchical manner also makes it possible for experts of a subdomain to define rules for their subdomain. For example, a professor knows better when to allow students access to the wireless network in class. They could be having a tutorial in class or a mentored distance learning session. If the professor formulates some rule that is in conflict with a general university policy (expressed as a standing rule in the ontology), then this conflict would be detected automatically.

3.3 Policy Implementation at Different Levels

To illustrate the challenge in implementing policies across heterogenous devices and subdomains, consider the wireless component of the m-learning pervasive environment. As mentioned earlier, there are commercially available WLAN controllers that allow policies to be defined and perform location tracking. They also provide a standardised interface, like SOAP/XML, for third party applications to make use of their data. But such controllers do not provide reasoning support or a shared knowledge base of domain concepts and relationships. The aim of developing ontologies is to promote a common understanding of terms in their respective namespaces. Such development takes place at both the higher global level for universal concepts and also at the lower domain-specific level for specialized concepts. Ultimately, all the concepts will be covered by the overlap and sharing of these different ontologies due to the fractal nature of these systems [19]. When expressed in formal languages like OWL DL and enhanced with rule languages like SWRL, the knowledge base has the added benefits of validation and deduction and has an abstraction to define domain behaviour. This abstraction can then be mapped to any compliant device, making the policy specification process device-agnostic. Of course, the ontology is also available in its XML syntax for doing mechanical parsing for different purposes, e.g. translation of tags and content from English to Urdu. This also lends itself to semantic matching algorithms. Elsewhere [20,21], we describe a device-level technique of using rule bundles in a bytecode format for model checking of constraints, including liveness and safety. We have also developed data structures, algorithms, and parallel computing methods for stochastic modelling and (probabilistic) model checking of large scale systems using different high-level formalisms, e.g. see [22,23]. We can use these techniques in conjunction to provide more flexible and complete low level and high level control of large-scale pervasive domains.

The ease of defining rules in SWRL or any other rule language depends on the level of expertise of the system administrator. This task needs to be simplified by using a high level language or other abstraction that matches more closely the skills of the system administrator. There is active research interest in policy definition languages [24] and it is an open research question how to automatically translate from a policy definition system, e.g. Ponder2 (`http://www.ponder2.net/`), to the OWL DL form.

4 Conclusions and Future Work

We have presented our formal knowledge-based approach towards depicting and controlling pervasive environments and have demonstrated it in a mobile learning enabled campus scenario. The benefits include semantic interoperability, reasoning support and resolution of conflicts. Our future work includes running quantitative and qualitative tests on the useability of our system. In this context, we also plan to leverage our earlier work in e-learning [25], location based services [26] and distributed virtualization [27]. Another research direction will be to enable automatic policy translation from a policy definition language to the formal ontological model.

References

1. Weiser, M.: The computer for the 21st century. SIGMOBILE Mob. Comput. Commun. Rev. 3(3), 3–11 (1999)
2. Jones, D., Bench-Capon, T., Visser, P.: Methodologies for ontology development. In: Proceedings of the IT&KNOWS Conference, XV IFIP World Computer Congress (1998)
3. Foundations. In: Grimm, S., Volz, R. (eds.) Semantics At Work, Ontology Management Tools and Techniques (June 2007)
4. Liu, L., Özsu, M.T. (eds.): Encyclopedia of Database Systems. Springer, New York (2009)
5. OWL Web Ontology Language overview, W3C recommendation (2004), `http://www.w3.org/TR/owl-features/`
6. Knowledge Interchange Format, KIF (1992), `http://www.ksl.stanford.edu/knowledge-sharing/kif/`
7. RIF use cases and requirements, W3C working draft (2006), `http://www.w3.org/TR/rif-ucr/`
8. SWRL: A semantic web rule language combining OWL and RuleML, W3C member submission (2004), `http://www.w3.org/Submission/SWRL/`
9. Cobcroft, R.S., Towers, S., Smith, J., Bruns, A.: Mobile learning in review: Opportunities and challenges for learners, teachers, and institutions. In: Online Learning and Teaching (OLT) Conference, pp. 21–30 (2006)
10. Motiwalla, L.F.: Mobile learning: A framework and evaluation. Comput. Educ. 49(3), 581–596 (2007)
11. Hightower, J., Borriello, G.: Location Systems for Ubiquitous Computing. Computer 34(8), 57–66 (2001)
12. Wi-Fi Based Real-Time Location Tracking: Solutions and Technology. White paper, Cisco Systems (2006) (available online)
13. Alvi, A., Greaves, D.J.: Checkable domain management with ontology and rules. In: ICIW 2008: Proceedings of the 3rd International Conference on Internet and Web Applications and Services, Athens, Greece, June 8-13, pp. 142–149 (2008)
14. O'Connor, M.J., Knublauch, H., Tu, S.W., Grosof, B.N., Dean, M., Grosso, W.E., Musen, M.A.: Supporting rule system interoperability on the semantic web with SWRL. In: Gil, Y., Motta, E., Benjamins, V.R., Musen, M.A. (eds.) ISWC 2005. LNCS, vol. 3729, pp. 974–986. Springer, Heidelberg (2005)

15. O'Connor, M.J., Tu, S.W., Nyulas, C., Das, A.K., Musen, M.A.: Querying the semantic web with SWRL. In: Paschke, A., Biletskiy, Y. (eds.) RuleML 2007. LNCS, vol. 4824, pp. 155–159. Springer, Heidelberg (2007)
16. Keck, D.O., Kuehn, P.J.: The feature and service interaction problem in telecommunications systems: A survey. IEEE Trans. Softw. Eng. 24(10), 779–796 (1998)
17. Bouma, W.: Feature interactions. In: Kent, A., Williams, J.G. (eds.) Encyclopedia of Computer Science and Technology, vol. 40, pp. 1–23. Marcel Dekker, Inc., New York (1999)
18. SWRLLanguageFAQ, Protégé wiki (2010), `http://protege.cim3.net/cgi-bin/wiki.pl?SWRLLanguageFAQ`
19. Berners-Lee, T., Kagal, L.: The fractal nature of the semantic web. AI Magazine 29(3), 29–34 (2008)
20. Alvi, A., Greaves, D.J.: A logical approach to home automation. In: IE 2006: Proceedings of the 2nd IET International Conference on Intelligent Environments, Athens, Greece, July 5-6, vol. 2, pp. 45–50. IET (2006)
21. Greaves, D.J., Gordon, D., Alvi, A., Omitola, T.: Using a .net checkability profile to limit interactions between embedded controllers. In: SENSORCOMM 2008: Proceedings of the 2008 Second International Conference on Sensor Technologies and Applications, Washington, DC, USA, pp. 555–561. IEEE Computer Society, Los Alamitos (2008)
22. Mehmood, R.: Disk-based techniques for efficient solution of large Markov chains. PhD thesis, University of Birmingham (2004)
23. Mehmood, R., Crowcroft, J.: Parallel iterative solution method for large sparse linear equation systems. Technical Report 650 UCAM-CL-TR-650, Computer Laboratory, University of Cambridge, UK (2005)
24. Thomas, R.K., Russello, G., Tsang, S.: Realizing the cdl cross-domain language in the ponder2 policy framework: Experiences and research directions. In: POLICY 2009: Proceedings of the 2009 IEEE International Symposium on Policies for Distributed Systems and Networks, Washington, DC, USA, pp. 76–83. IEEE Computer Society, Los Alamitos (2009)
25. Adamson, M.K., Mehmood, R.: Developing event based hierarchical middleware for e-learning. In: IADIS International Conference e-Learning 2007 (part of MCCSIS 2007), pp. 284–291 (2007)
26. Ayres, G., Mehmood, R., Mitchell, K., Race, N.J.P.: Localization to enhance security and services in wi-fi networks under privacy constraints. In: Communications Infrastructure. Systems and Applications in Europe, EuropeComm 2009, London, UK, August 11-13 (2009), Revised Selected Papers, LNICST, vol. 16, 175–188 (2009)
27. Mehmood, R., Crowcroft, J., Hand, S., Smith, S.: Grid-level computing needs pervasive debugging. In: GRID 2005: Proceedings of the 6th IEEE/ACM International Workshop on Grid Computing, Washington, DC, USA, pp. 186–193. IEEE Computer Society, Los Alamitos (2005)

Automatically Finding Answers to "Why" and "How to" Questions for Arabic Language

Ziad Salem, Jawad Sadek, Fairouz Chakkour, and Nadia Haskkour*

Aleppo University, Electrecal and Electroneic Engineering Faculty, Computer Engineering Department, *Faculty of Arts and Humanities, Aleppo, Syria
dr_ziad@hotmail.co.uk, jawad10a@yahoo.com, feirouzch@yahoo.fr

Abstract. This paper addresses the task of extracting answers to *why and how to*-questions from Arabic texts which has not been addressed yet for Arabic language in the field of question answering systems (QA). The system developed here uses one of the leading theories in computational linguistics called Rhetorical Structure Theory (RST) and based on cue phrases to both determine the elementary units and the set of rhetorical relations that is relevant to the targeted questions. Our experiment has been conducted on Arabic raw texts (automatically annotated) taken from Arabic websites and has gave a good result comparing with a one already done before to why-questions answering for English language.

Keywords: Rhetorical Structure Theory, Natural Language Processing, Question Answering for Arabic, why and how to questions, Discourse analysis.

1 Introduction

Day by day the amount of information available on the internet is growing, and it becomes more and more difficult to find answers on the WWW using standard search engines, as consequence question answering systems (QA) will become increasingly important. The main aim of QA systems is to provide the user with a flexible access to information allowing him for writing a question in natural language and presenting a short answer rather than a list of possibly relevant documents which contain the answer.

Arabic is the sixth most widely spoken language in the world [1], yet there are relatively few studies to improve Arabic information search and retrieval compared to other languages and this is true for the QA task. However few researches built QA systems oriented to the Arabic language. The systems were focused on factoid questions like who, what, where and when questions [2][3] in which named entity recognition can make a substantial contribution to identifying potential answers in a source document, but none of those system addressed *why and how to*-questions which different techniques are needed.

In the current paper the research aims at developing a system for answering *why and how to*-questions for Arabic language including a proper evaluation method as first attempt to address this type of questions. The system uses RST that has been applied in a large number of computational Linguistics applications.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 586–593, 2010.

2 Rhetorical Structure Theory

Rhetorical structure theory was developed at USC (University of Southern California) by William Mann and Sandra Thompson. The aim was finding a theory of discourse structure or function that provides enough detail to guide a computer program in generating texts. Based on their observation of edited text from a wide variety of sources, Mann & Thompson have several assumptions about how written text functions, and how it involves words, phrases, grammatical structure summarizing as following [4]:

- **Organization:** Texts consists of functionality significant parts.
- **Unity and coherence:** There must be sense of unity to which every par contributes.
- **Hierarchy:** Elementary parts of a text are composed into larger parts, which in turn are composed of yet larger parts up to the scale of the text as whole.
- **Relation Composition:** Relations hold between parts of a text. In which every part of a text has a role, a function to play, with respect to other parts in the text. A small finite set of highly recurrent relations holding between pairs of parts of text is used to link parts together to form larger parts. All rhetorical relations that can possibly occur in a text can be categorized into a finite set of relation types.
- **Asymmetry of Relations:** RST establishes two different types of units. Nuclei are the most important parts of a text, whereas satellites contribute to the nuclei and are secondary. The most common type of text structuring relation is an asymmetric class, called nucleus-satellite relations, the nucleus is considered to be the basic information, and more essential to the writer's purpose than the satellite. The satellite contains additional information about the nucleus. And it is often incomprehensible without the nucleus, whereas a text where the satellites have been deleted can be understood to a certain extent. Table 1 illustrates some of the relations identified by Mann and Thompson.

Table 1. Presents some of the relations used in RST

Relation name	Nucleus	Satellite
Background	Text whose understanding is being facilitated	text for facilitating understanding
Elaboration	basic information	Additional information
Antithesis	ideas favored by the author	ideas disfavored by the author
Enablement	An action	information intended to aid the reader in performing an action
Evaluation	A situation	an evaluative comment about the situation

Years of text analysis using RST have shown that RST is useful to capture the underlying structure of texts. Furthermore, RST has proven to be adequate in computational implementations, in the automatic analysis of texts and in the generation of coherent text [5].

3 Using Rhetorical Relations for Question Answering

Some types of rhetorical relations that might be relevant to *why and how to-* questions can help finding answers for those questions. Let us consider the two following examples, taken from Arabic websites, which clarify the method used to extract answers:

3.1 Example 1

[قالت دراسة نشرت في صحيفة بريتش ميديكال إن الشاي الاسود الذي تم إعداده عند درجة حرارة تزيد عن ٧٠ درجة مئوية يزيد من خطر الإصابة بالسرطان][١] *[وإن ذلك يفسر ارتفاع الإصابة بسرطان المري بين بعض الشعوب الغير غربية.]*[٢]

[The research published in the British Medical Journal found that black tea made at temperature greater than 70 c^o, can raise the risk of cancer,][1] *[and that may be the cause of high rates of esophageal cancer among non western people.]*[2]

In this example, $unit_1$ gives information about the cause of the problem presented in $unit_2$, so we can say that an interpretation relation holds between the two units as illustrates in Fig.1.

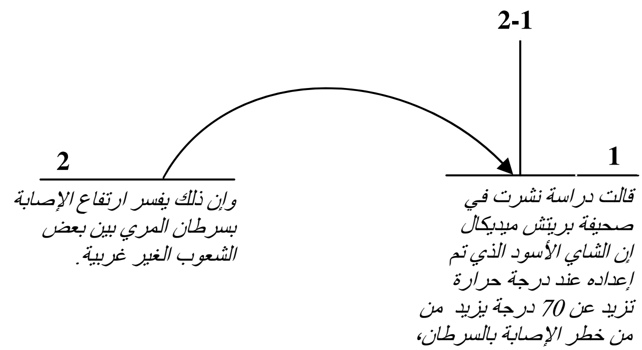


Fig. 1. The schema of the Arabic text in the example 1

Now in case of the following question:

{ لماذا تعد الإصابة بسرطان المري مرتفعة بين الشعوب الغير غربية ؟}

{Why does esophageal cancer has high rates among non western people?}

We notice that the question corresponds to the $unit_2$, so the other part of relation will be the answer for the question which is the $unit_1$.

3.2 Example 2

[يمكنك أن تصل إلى حالة من الا سترخاء العميق]1 [من خلال تنفس متعادل، يكون فيه كل شهيق و كل زفير متساويين في الطول ويساوي كل منهما الآخر في الطول، أغمض عينيك واستنشق وأنت ....]2

[You can reach a state of deep relaxation][1][through equal breathing where each inhalation and exhalation are long and of equal length. Close your eyes and inhale while……][2]

Also in this example, we notice that $unit_1$ explains the notion mentioned in $unit_2$, so we can say that an explanation relation holds between the two units as illustrates in Fig.2.

Given the following question:

{ كيف يمكن الوصو ل إلى الاسترخاء العميق ؟}

{How to reach a stage of deep relaxation?}

The question corresponds to the $unit_1$, so we can consider the other part of relation as the answer for the question.

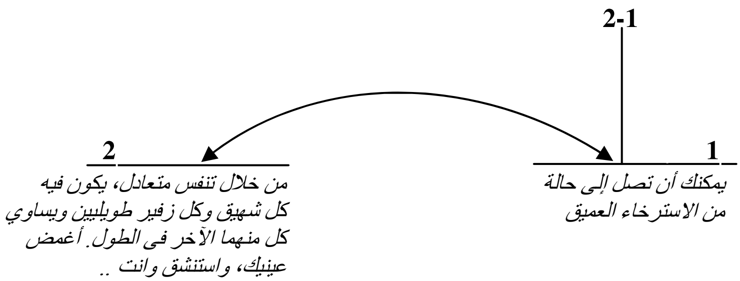


Fig. 2. The schema of the Arabic text in the example 2

We did Arabic text analysis in order to extract a set of rhetorical relations that can lead to answer *why* and *how to* questions. Identified by Al-sanie [6], eleven rhetorical relations have applied in an Arabic text summarization system. We choose four rhetorical relations from his work (Interpretation–Base–Result–Antithesis) and added other four relations (Causal–Evidence–Explanation–Purpose) to get the set of relations and its corresponding types of answer as shown in Table2.

In order to automatically derive the text structure, it first needs to determine the elementary units of a text and then find the rhetorical relations that hold between these units. Marcue [7] relied on cue phrase to perform the previous two steps as a sufficiently accurate indicator of the boundaries between elementary textual units and of the rhetorical relations that hold between them. We will use the same method in the present work. Cue phrases are words and phrases that used by writer as cohesive ties between adjacent clauses and sentences and they are crucial to the reader for understanding of the text.

Analyzing an Arabic corpus and studying the way the Arabic writer used to convey his thought to the reader [8][9] we generated a set of cue phrases that signaled each relation shown in Table2. For example the relation Explanation can be hypothesized on the basis of the occurrence of the cue phrases (......،"بواسطة"، "عن طريق"،"من خلال"). Also (...."وقال" ، "أكد" ،"وأشار") can signals an Evidence relation.

Table 2. Presents a set of the Arabic rhetorical relations used to answer why and how to Arabic questions

Question type	نوع السؤال	English equivalence	اسم العلاقة
Why - how to	لماذا – كيف	Interpretation	تفسير
Why	لماذا	Causal	سببية
Why	لماذا	Result	نتيجة
Why	لماذا	Base	قاعدة
Why - how to	لماذا – كيف	Antithesis	استدراك
Why - how to	لماذا – كيف	Purpose	غاية
how to	كيف	Explanation	شرح
Why	لماذا	Evidence	اثبات

4 Textual Units and Question Processing

Before starting the answer retrieval task we need to process and tokenize both the question and the text in which the answer may be found, this subsumes performing the following steps:

- **Normalization:** certain combinations of characters can be written in different ways in the Arabic language. For instance, glyphs that combining HAMZA or MADDA with ALEF (آ ، إ، أ) are sometimes written as a plain ALEF (ا), also the letter TAA MARBOTH (ة) is sometimes changed to HAA (ه) at the end of a word, and this will result in difficult to recognize some Arabic words, So we have to normalize all orthographic variations.
- **Stemming:** Arabic, as all Semitic languages, is a highly inflected language and has a very complex morphology; a given headword can be found in huge number of different forms. This abundance of forms results in greater likelihood of mismatch between the form of word in a question and the forms found in text relevant to the question. Thus stemming is a basic step in this context, and many are the research studies which attempt to develop Arabic stemmers. In our system we used Larkey's light- stemmer [10] in case the word's category is noun, or Khoja's root- base stemmer [11] in case of verbs which will be more efficient as proposed by Al- shammari [12].
- **Stop words removal:** due to the absence of standardized list of Arabic stop words, we dropped 300 high-frequent common words, based on Arabic literature and excluding the cue phrase list, that gives no benefits to the matching results and may save space and speed searching.

We compute the similarity between the question and the textual units by applying Vector Space Model and rank the textual units in descending order according to the similarity values using the formula shown below:

$$\mathrm{Sim}\,(Q\,,U_i) = \text{Cosine}\,\theta_{Ui} = \frac{\sum_j W_{Q,j}\;W_{i,j}}{\sqrt{\sum_j W_{Q,j}^2\;\;\sum_j W_{i,j}^2}} \qquad (1)$$

Where $WQ{,}j$, $\boldsymbol{W_{i,j}}$ are the weights of the $\boldsymbol{j^{th}}$ keyword of the question $\boldsymbol{Q}$ and textual unit $\boldsymbol{U_i}$ respectively.

The algorithm presented in Fig .3 takes as input a sequence of textual units belonging to a text and a question related to the text, and then returns a set of ranked answers.

```
Input :  A question Q ,
A sequence U[n] of textual units and a list RR of
relations that hold among the units in U.
Output: A set A of candidate answers.
1. A := null;
2. Identify the type of Q;
3. Identify a set of relations rr in RR corresponding to
    the Q type;
4. Match Q against the textual units U[n];
5. For each match Ui
6.       if ( Ui have a relation rri of one of the types in rr)
7.             sp := related span of rri;
8.             A := A ∪ sp ;
9.       else
10.          Discard the current Ui;
11.      end if
12. end for
13. Rank the answers;
```

Fig. 3. Algorithm that select answers for a given question

5 Experiments and Results

We implemented our system using java programming language. For the purpose of measuring the performance of our system we used the same experiment conducted by S.verberne [13]. We selected a number of texts of 150-350 words each. The texts were extracted from Arabic news websites. Then we distribute those texts to 15 people from different discipline and we asked them to read some of the texts and to formulate *why and how to*-questions for the answers could be found in the text, the subjects were also asked to formulate answers to each of their questions. This resulted in a set of 98 *why and how to*-questions and answers pair.

We run our system on the 98 questions we collected, and then compared the answers found by the system to the user-formulated answers; if the answer found matches the answer formulated by subject then we judged the answer found as correct. The system found the correct answer for 54 questions and this is 55% of all questions. Result is given in Table 3.

In the system created by S.vrberne, they collected a set of 336 *why*-question and answer pairs, connected to seven manually annotated English texts from the RST Treebank of 350-550 words each. When they evaluated the system, they obtained a recall of 53.3%.

Comparing our result with the one obtained by S.verberne (Table 4); it can be seen that they selected longer texts than we did. But on the other hand we dealt with raw text (the structure has automatically driven) whereas they dealt with manually annotated data. Additionally, they reported that the performance would decline if they use automatically created annotation [13]. As consequence, using the rhetorical relations proposed in this research for answering *why* and *how to* Arabic questions showing promising results.

Table 3. Shows the outcome of the system

	# questions	% of all questions
Questions handled	98	100
Correctly answered	54	55.1
Wrongly answered	44	44.9

Table 4. Presents a comparison between the two question answering systems

	Arabic QA	English QA
Questions #	98	336
Words #	150-350	350-550
Structure derivation	Automatically	manually
Source	Arabic Websites	RST Treebank
Recall	55%	53.3%

6 Conclusion and Future Work

In this paper we presented the first study for automatically finding answers to *why and how to*-questions for Arabic language based on Rhetorical Structure Theory. We performed a manual analysis on a set of Arabic texts to select a number of relation types that is relevant for those kinds of questions; we also selected some of cue phrases to signal the extracted relations. Additionally we carried out an evaluation of the system and compared it with the Suzan study. The result showed promising future in the direction of dealing with longer texts than those handled in this study.

References

1. The Bridge Language Report (2007), `http://www.bridgelanguagecenter.com`
2. Benajiba, Y., Rosso, P., Lyhyaoui, A.: Implementation of the Arabic QA Question Answering System's Computers. In: ICTC (2007)
3. Hammou, B., Abu-Salem, H., Lytinen, S., Evens, M.: QARAB: A question answering system to support the Arabic language. In: Workshop on Computational Approaches to Semitic Languages, ACL (2002)

4. Mann, W., Matthiessen, C., Thompson, S.: Rhetorical Structure Theory and Text Analysis. In: A Frame Work for the Analysis of Texts, pp. 79–195 (1992)
5. Mann, W., Taboada, M.: Rhetorical Structure Theory: Looking back and moving ahead. SAGE. Discourse Studies. 8, 423–459 (2006)
6. Al-Sanie, W., Touir, A., Mathkour, H.: Towards a Rhetorical Parsing of Arabic Text. In: International Conference on Computational Intelligence for Modeling, Web Technologies and Internet Commerce (CIMCA-IAWTIC 2005) (2005)
7. Daniel, M.: The Theory and Practice of Discourse Parsing and Summarization. The MIT Press, London (2000)
8. Jattal, M.: Nezam al-Jumlah, pp. 127–140. Aleppo University (1979)
9. Haskour, N.: Al-Sababieh fe Tarkeb al-Jumlah Al-Arabih. Aleppo University (1990)
10. Larkey, L.S., Ballesteros, L., Connell, M.E.: Improving stemming for Arabic information retrieval: Light stemming and co-occurrence analysis. In: 25^{th} SIGIR International Conference Research and Development in Information Retrieval, pp. 275–282. Tampere, Finland (2002)
11. Khoja, S., Garside, R.: Stemming Arabic Text. Computing Department. Lancaster University, Lancaster (1999)
12. Al-Shammari, E.: Towards an Error Free Stemming. In: LADIS European Conference on Data Mining (ECDM 2008). The Netherland, Amsterdam (2008)
13. Suzan, V., Lou, B., Nelleke, O.: Discourse-based answering of why-questions. Treatment Automatic Des Languages, Special Issue on Computational Approaches to Discourse and Document Processing 47(2), 21–41 (2007)

From Information to Sense-Making: Fetching and Querying Semantic Repositories

Tope Omitola, Ian C. Millard, Hugh Glaser,
Nicholas Gibbins, and Nigel Shadbolt

Intelligence, Agents, Multimedia (IAM) Group
School of Electronics and Computer Science
University of Southampton, UK
{tobo,icm,hg,nmg,nrs}@ecs.soton.ac.uk

Abstract. Information, its gathering, sharing, and storage, is growing at a very rapid rate. Information turned into knowledge leads to sense-making. Ontologies, and their representations in RDF, are increasingly being used to turn information into knowledge. This paper describes how to leverage the power of ontologies and semantic repositories to turn today's glut of information into sense-making. This would enable better applications to be built making users' lives easier and more effective.

1 Introduction

Modern digital devices have made the gathering, sharing, storing, and creating data to be a relatively painless and trivial exercise. This is evinced by the latest IDC Report[1] which reported that in 2009 alone, the amount of digital information produced in the world grew by 62% to nearly 800,000 petabytes, and in 2010 it will grow to 1.2 million petabytes. Data is being captured today as never before, while the amount of data generated by people doubles roughly every 1.5 years. The plummeting of costs of gathering and storing these data is having a tremendous impact on the expectations of individuals as they create and share information about themselves and about their relationships with others.

Humans have always gathered, stored, shared, and created data with friends and members of their tribe or clan – data on where to find good food, what places to avoid, etc. The cost and overhead of communication has always been very high and, before the advent of computation and the internet, the serendipity of someone coming across what you wanted to share has always been very low. But the ease of computation and the internet has made it easier to gather, create, and share information. Information is being generated at such a prodigious rate that the challenge now is **"sense-making"**, how do we curate information, version it, maintain it, index it, search it, query it, retrieve it, and re-use it, thereby helping people discover relevant content. The questions we should be asking ourselves are what shall we collect and what applications can we build that will help users make **sense** of these data in order to make users' lives easier, more effective, and more interesting.

[1] http://www.emc.com/collateral/demos/microsites/idc-digital-universe/iview.htm

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 594–602, 2010.

1.1 Information, Knowledge, and Sense-Making

In order to achieve adequate or useful sense-making from our information, they need to be turned into knowledge. By knowledge, we mean “information that can be sufficiently interpreted to enable action”[1]. To turn information into actionable knowledge, it is necessary to understand the connections between it and business and human processes. To manage this knowledge effectively, an important solution is the development and application of technologies useful to render information interpretable in order to enable effective action. Six challenges to effective knowledge management have been identified[1]. These are:

1. Knowledge Acquisition: Although we have a surfeit of information, the knowledge available in them are often insufficient or poorly-specified. The goal of Knowledge Acquisition is to turn information into usable knowledge. This could involve making tacit knowledge explicit, thereby identifying gaps in the knowledge already held, acquiring and integrating knowledge from multiple sources, and acquiring knowledge from unstructured media (e.g. natural language or free-flowing text).
2. Knowledge Modelling: Here, acquired knowledge is made usable for problem-solving by using model structures for its representation. These model structures, called ontologies, are specifications of the generic concepts, attributes, relations and axioms of a knowledge base or domain. They can act as a format for understanding how knowledge will be used.
3. Knowledge Publishing: The challenge here is to get the right knowledge, in the right form, in the right place, to the right person, at the right time. This will involve understanding the knowledge representation structures, the knowledge of the user(s), and of their context.
4. Knowledge Re-use: Knowledge representation are usually highly domain-specific, thereby making them more difficult to be re-purposed for different domains. A good solution to this problem will provide high leverage for effective knowledge use.
5. Knowledge Maintenance: This challenge involves making the knowledge repository functional. One of the key questions here is how does a consumer of a knowledge repository know that something has changed, and what has changed in a repository.
6. Knowledge Retrieval: In a world of distributed and federated knowledge bases, the challenge here includes finding where the relevant knowledge is, understanding its structure(s), and querying it for its values.

On the World Wide Web, ontologies are now more commonly used as the model structures to turn information into (re)usable knowledge. RDF[2] is being widely used as the representation language for ontologies, and are also the foundation for the next generation of the Web, i.e. the “Web of Data”. This Web of Data is a Web of actionable information, i.e. information derived from

[2] http://www.w3.org/TR/REC-rdf-syntax/

data through a semantic theory for interpreting the symbols. The semantic theory provides an account of meaning in which the logical connection of terms establishes interoperability between systems[2].

Increasingly, private corporations and governments are realising the potential of encoding knowledge as RDF. Facebook[3] recently announced their Open Graph Protocol[4] which allows a Facebook user to integrate other non-Facebook web pages into the user's social graph. Open Graph Protocol uses machine-processable semi-structured data to mark up web pages. Various governments, in order to improve the delivery of services to their citizens, are opening up their data and publishing these data in semi-structured format, many of them in RDF, to improve the delivery of goods and services. The United States government has set up data.gov[5] to release public data. The UK Government, keen to unlock the benefits of economic and social gain of public sector information (PSI) reuse, has set up data.gov.uk[6]. All these efforts will enable their citizens to ask questions, such as: "Where can I find a good school, hospital, investment advisor, or a good employer?" They will also lead to improvements of the delivery of services to the public.

Adequate steps need to be taken to learn how to publish and consume (public) semi-structured data.

2 A Case Study of Publishing and Consuming Public Data

Omitola et. al. [4] carried out a case study showing how United Kingdom geographical data, from the UK's Ordnance Survey, can be used as a set of "join points" to mesh public data for crime, mortality rates, and hospital waiting times. Meshing was defined as the ability to naturally merge together a dynamic set of information sources, and a join-point as a point of reference shared by all datasets. They investigated the use of disparate sets of data in an effort to better understand the challenges of their integration using Semantic Web approaches. Part of this investigation involved ascertaining the datasets that were available, their formats, and converting them into (re)usable formats. They faced a number of challenges which included:

1. Sourcing the datasets: Since many of the datasets of interest were in comma-separated-value (CSV) format, a process of converting the datasets to RDF had to be undertaken,
2. Modelling the datasets: The sources of the datasets contained little or no explicit semantic description of the data, [4] had to provide schema definitions for these data,

[3] http://www.facebook.com
[4] http://developers.facebook.com/docs/opengraph
[5] http://www.data.gov
[6] http://data.gov.uk

3. Aligning of datasets: The datasets were harvested from disparate public bodies, the problem of unintentional coreference, i.e. different names referring to the same thing, had to be solved,
4. The challenge of re-use and consumption of the data also needed to be solved.

A number of solutions used to solve some of the challenges were introduced and described in the paper. These included, inter alia,

1. the use of the SCOVO ontology [5] to model the multi-dimensional aspects of the datasets, such as time, entity types, etc,
2. the design decisions and the efforts needed to convert the (mostly) CSV-encoded datasets into RDF,
3. the selection of the appropriate "join point" used to integrate the datasets,
4. the usage of a co-reference service to resolve co-references, and
5. the use of Exhibit[7] to develop the client application used to consume the data.

3 Ontologies, Linked Data, and Linked Data Cloud

The continued growth in the adoption and usage of ontologies and semi-structured data in government and industries is bringing about the growth in datasets published in linked data format, and a growing interest in connecting these datasets together. Linked Data is a style of publishing data on the Web that emphasises data reuse and connections between related data sources. This growth and interest can be seen in the Linked Data community[8] which aims at making data freely available to everyone and to extend the Web with a data commons by publishing various open data sets as RDF and by setting RDF links between data items from different data sources. Figure 1 shows a linked data cloud of the data sets that have been published and interlinked by the community so far. Collectively, the data sets consist of over 13.1 billion RDF triples, which are interlinked by around 142 million RDF links (as of July 2009).

With these datasets in different knowledge bases and data stores, there is a paradigm shift occurring. This shift is an important one. We are moving away from the paradigm of "given a set of data, what technique(s) can I use on this dataset and gain insights" to the paradigm of "given a problem, what is the best dataset I can get to solve the problem or answer the questions". The kinds of questions we may want to ask are:

1. We could use data from a geographical database such as Geonames[9], combine this with a social network to determine how much people's locations and the distances between them affect their chance of being friends. If we combine this with census data, we could ask questions to determine if population size or population demographics affect this chance of making new friends (to ascertain if people from small towns are more open to new friendships, etc).

[7] http://simile.mit.edu/wiki/Exhibit/API
[8] http://linkeddata.org/
[9] http://www.geonames.org/

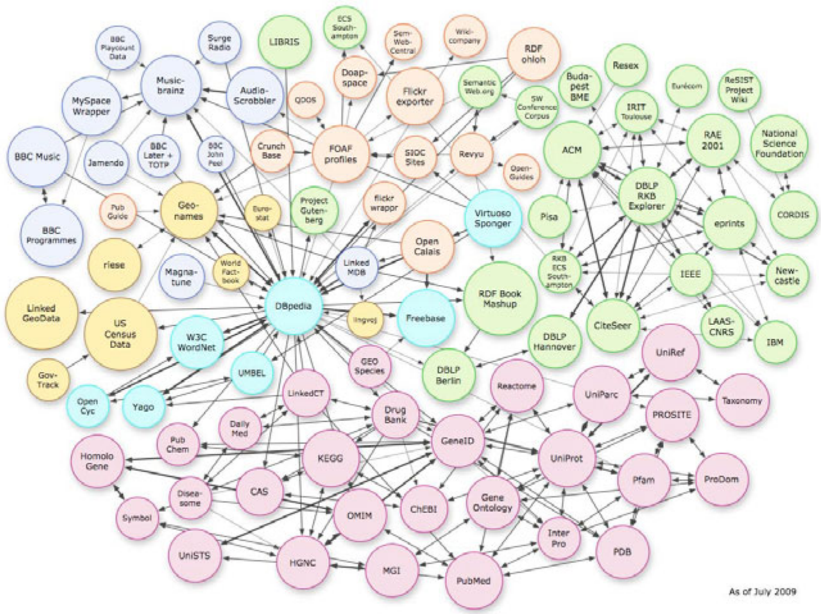


Fig. 1. The Linked Data Cloud

2. We could use CO_2 emmisions data of all the councils of the United Kingdom and check it with the different income levels of each council and with data of the average educational attainments of each council to determine if there is a correlation between income and awareness of environmental issues.

The outputs from answering questions such as the ones above may be used as inputs to an inductive reasoner or “human” expert for required validations and/or necessary recommendations.

In order to answer the aforementioned queries, we need to make it easy to find the relevant datasets and datastores in the first instance. We need to build a platform that makes it easy for data to find data.

4 Interlinking Datasets and Co-referencing Service

For data to find data, we need a dataset discovery mechanism and, after discovering relevant datasets, the selection of the best-suited ones. To make this possible, the Linked Data community has come up with **voiD**[10] [6] , a “Vocabulary of Interlinked Datasets”.

4.1 voiD

voiD is an RDF based schema used to describe the content and the interlinking between datasets. There are two core classes at the heart of voiD:

[10] http://void.rkbexplorer.com/

1. A dataset (*void:Dataset*), i.e. a collection of data, which is:
   - published and maintained by a single provider,
   - available as RDF,
   - accessible, for example, through dereferenceable HTTP URIs or a SPARQL[11] endpoint
2. The interlinking modelled by a linkset (*void:Linkset*). A linkset in voiD is a subclass of a dataset, used for describing the interlinking relationship between datasets. In each interlinking triple, the subject is a resource hosted in one dataset and the object is a resource hosted in another dataset. This modelling enables a flexible and powerful way to state the interlinking between two datasets, such as how many links there exist, the kind of links, and who made these statements.

Figure 2 depicts[12] the modelling of the interlinking in voiD.

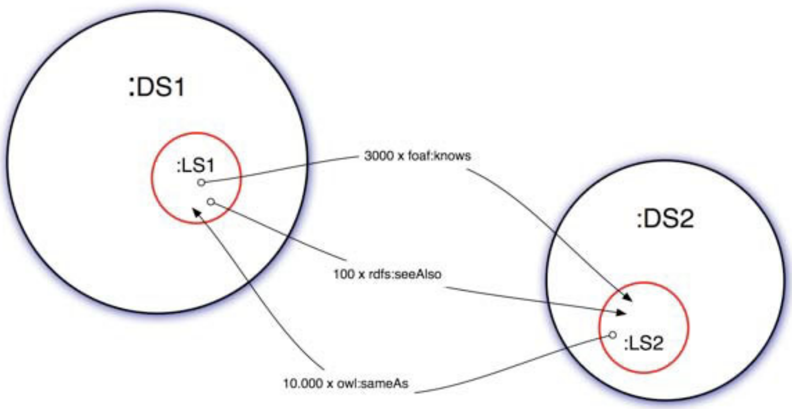


Fig. 2. Model of voiD Interlinking

4.2 Co-referencing

The acquisition of knowledge from heterogeneous sources carries with it the problem of unintentional coreference, i.e., different sources may refer to the same entities by different means. If we want to use this knowledge effectively, we need to be able to determine which entities are co-referent, and to collapse these multiple instances into a standard representation. **sameAs**[13][3] has been developed as a web service for the collection and publication of information about co-referent entities. For example, if one wants to look at co-referent entities for the UK's "county of cambridgeshire", this can be called via the following URI: *http://sameas.org/html?q=county+of+cambridgeshire*. Doing this, one can see that the "county of cambrigeshire" is defined in Geonames as

[11] http://www.w3.org/TR/rdf-sparql-query/
[12] From http://semanticweb.org/wiki/VoiD
[13] http://sameas.org/

http://www.geonames.org/2653940/ and in DBpedia[14] as *http://dbpedia.org/page/Cambridgeshire.*

Since we now have a mechanism to ascertain (un)intentional coreference between entities (the sameAs service), and to discover datasets (voiD), we need to develop a platform that can be used to query and fetch the relevant data amongst these federated datasets.

5 Federated Fetching and Querying

5.1 Federated Fetch and Contextualised Directories

Directories are locator services, they return references after being provided with query terms. Provision of references turns the next series of activities to "fetch", and on the WWW, this is equivalent to "federated fetch". For example, terms searches on the Google[15] search engine, after going to a special directory created by Google, return references to the real documents. This type of directory is an example of **contextualised directories**. Contextualised directories enable data to find data. An extension of contextualised directories are semantically reconciled directories. Semantically reconciled directories exploit synonyms (or coreferences) to return pointers to different words that mean the same thing. An example of this is the *sameAs* web service mentioned above. Further on from these are "semantically reconciled and relationship-aware directories" that provide higher degree of context by allowing users to discover additional data. One can query for any single set of terms and locate a bundle and, by using the relationship-awareness property of the directory, learn how that bundle relates to other bundles. A combination of *voiD* and *sameAs* gives us semantically reconciled and relationship-aware directory which we can use to find relevant data and datasets.

5.2 Geography of Dataset Relationships

Space is one of the principal media through which structure and form are expressed, and spatial organisation produce complex geometries of relationships and structure. A key endeavour is to determine these geometries in order to aid understanding and navigation. To determine these spatial geometries, one solution that is being explored is to leverage the power of voiD and sameAs to perform topology analysis of the datasets in the Linked Data cloud. The kinds of analysis that can be done include:

1. Degree centrality: This measures the extent to which a node or a dataset connects to all other nodes or datasets in the linked data cloud,
2. Small world: This measures the average minimum path between the nodes in different datasets,

[14] http://dbpedia.org/
[15] www.google.com

3. Clustering co-efficient/factor: This measures the probability of two datasets being neighbours of one another.

This topology relationship will act as a meta-network layer (maintaining state of networks of relationships) which can also be continually updated to cope with the changing dynamics of the datasets.

We envisage to leverage the power given us by voiD, sameAs, and the Meta-Network layer to perform more optimised query plans and as path indices to speed up query processing. This will lead to the ability to efficiently discover datasets and select the most suitable ones.

6 Conclusion

In this paper, we described efforts being made in the (Semantic) Web community to turn today's information glut into sense-making. The central element of these efforts is the representation of information in semi-structured format, (e.g. using RDF), called ontologies. We described a case study of how this was achieved using government (public) data, the attendant challenges, and the devised solutions. Such data are now being aggregated into disparate datasets and linked together forming a linked data cloud. We described services that are in use to perform the linkage (voiD) and to ascertain similarity between data concepts (sameAs). We outlined current work that is ongoing which will use voiD and sameAs to build a federated fetching and querying platform to efficiently discover relevant datasets for (Semantic) web applications to use.

Acknowledgements

This work was supported by the EnAKTing project, funded by EPSRC project number EP/G008493/1. We thank Manuel Salvadores and Gary Wills for their contributions.

References

1. Shadbolt, N., O'Hara, K.: An Overview of the Aims, Ambitions and Assumptions of the Advanced Knowledge Technologies Interdisciplinary Research Collaboration. Advanced Knowledge Technologies Selected Papers (2003)
2. Shadbolt, N., Hall, W., Berners-Lee, T.: The Semantic Web Revisited. Advanced Knowledge Technologies Selected Papers (2006-2007)
3. Glaser, H., Jaffri, A., Millard, I.C.: Managing Co-reference on the Semantic Web. In: LDOW 2009 (April 2009)
4. Omitola, T., Koumenides, C.L., Popov, I.O., Yang, Y., Salvadores, M., Szomszor, M., Berners-Lee, T., Gibbins, N., Hall, W., Schraefel, M., Shadbolt, N.: Put in your postcode, out comes the data: A Case Study. In: 7th Extended Semantic Web Conference, Greece (May 2010)

5. Hausenblas, M., Halb, W., Raimond, Y., Feigenbaum, L., Ayers, D.: SCOVO: Using Statistics on the Web of Data. In: Aroyo, L., Traverso, P., Ciravegna, F., Cimiano, P., Heath, T., Hyvönen, E., Mizoguchi, R., Oren, E., Sabou, M., Simperl, E. (eds.) ESWC 2009. LNCS, vol. 5554, pp. 708–722. Springer, Heidelberg (2009)
6. Cyganiak, R., Stenzhorn, H., Delbru, R., Decker, S., Tummarello, G.: Semantic Sitemaps: Efficient and Flexible Access to Datasets on the Semantic Web. In: Bechhofer, S., Hauswirth, M., Hoffmann, J., Koubarakis, M. (eds.) ESWC 2008. LNCS, vol. 5021, pp. 690–704. Springer, Heidelberg (2008)

A Variable Topology Analog Filter Suitable for Multi-mode Wireless Applications

Sorin Hintea[1], Doris Csipkes[1], Gabor Csipkes[1], and Hernando Fernandez-Canque[2]

[1] Technical University Cluj Napoca, Romania
doris.csipkes@bel.utcluj.ro
[2] Caledonian University Glasgow, Scotland

Abstract. A reconfigurable channel or band-select filter is one of the key elements of multi-mode analog front ends. This paper presents the system level aspects of a fully reconfigurable analog filter. The proposed circuit is based on analog array-like architecture that enables the mapping of low pass and polyphase band pass filters on a regular structure of reconfigurable unit elements. The current conveyor based topology represents a compromise between the generality of traditional field programmable analog arrays and the highly specialized requirements imposed by multi-mode wireless receiver architectures.

1 Introduction

Global roaming, worldwide equipment compatibility and the ever increasing diversity of services supported by providers are the main driving forces of emerging telecommunication technologies, such as the software defined radio (SDR), or more recently the cognitive radio (CR)[1][2].

One key element of reconfigurable analog front ends in practical SDR realizations is the channel or band select filter. The classical approach to multi-mode operation used switched signal paths in order to adapt the receiver to the requirements of a limited number of supported standards. As the number of supported standards increased, filters evolved into building blocks featuring programmable frequency parameters over a relatively wide range of frequencies [3]. The next step in the evolution process is to implement fully reconfigurable filters that can seamlessly adapt to the spectral characteristics of a wide variety of communication standards while maintaining a relatively low hardware overhead, low consumption and high cost efficiency. This adaptation may be implemented by means of a software that dynamically changes the frequency parameters, topology and even the approximation or synthesis technique.

The requirements for full reconfigurability is similar in concept with field programmable analog arrays (FPAA-s). This leads to the idea to attempt to use a FPAA-like circuit for implementing the fully reconfigurable filters.

FPAA-s have been traditionally associated with rapid prototyping and design cycle shortening in analog design. They have been mostly regarded as a monolithic collection of analog building blocks, user controlled routing resources and memory elements [4]. The block diagram of a FPAA is shown in Fig.1.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 603–612, 2010.

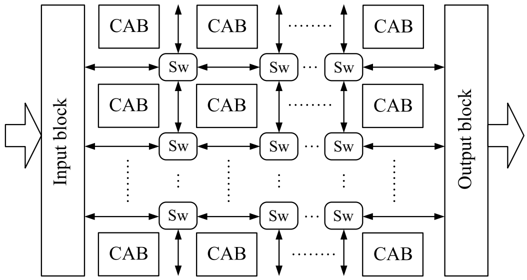


Fig. 1. The General Architecture of a FPAA

The crossbar style routing network, inherited from digital field programmable gate arrays (FPGA-s), is a feature specific to most of the FPAA architectures described in the literature. It allows routing the signal throughout the entire array and access to any of the configurable analog blocks (CAB-s) at the cost of a large number of switches connected in the signal path. However, the switch count in large scale FPAA-s becomes a limiting factor in terms of bandwidth and linearity. The structure of each CAB depends on the target application. Mapping a wide range of analog functionality to the FPAA leads to complicated CAB designs and an additional degradation of performances [5].

The main goal of the FPAA architecture described in the remainder of this paper is to investigate the system level connection between wireless technologies, required to implement filtering in SDR receivers, and the traditional flexibility of analog arrays. The proposed variable topology filter may be regarded as a compromise between the generality of the functions mapped to an analog array and the performance requirements specific to fully reconfigurable wireless receivers.

2 An FPAA Approach to Reconfigurable Filter Design

2.1 The FPAA Architecture

The survey of the existing literature shows a wide variety of proposed FPAA architectures. The implementation of the switching network, the structure of an individual CAB and the data used for configuration are highly dependent on the desired functionality. Each application defines key aspects of the circuit, such as voltage or current mode operation, continuous time or sampled processing and the fundamental active elements [4].

Most of the FPAA-s specifically designed for filtering applications are based on building blocks imposed by fixed topology filter design techniques. The fundamental active element is often the operational transconductance amplifier (OTA). Other approaches use second generation current conveyors (CCII-s), especially when paired with current mode operation [5][6][7]. The FPAA proposed in this work uses current mode signal processing along with fully differential, double output current conveyors (DOCCII-s). Unlike general architectures that support any type of filter mapped on the array [5], the FPAA allows the implementation of low pass and polyphase band pass filters often employed in direct conversion or low-IF receivers. The implemented filters can be realized with various synthesis methods in order to conserve sensitivity performance, have variable order ranging from 2 to 8 and it is suitable for developing programmable frequency parameters.

The architecture of the FPAA is based on a signal bus and a doubled linear arrangement of CAB-s, interlaced with frequency translation elements (FT) specific to complex filters. The signal bus allows the injection of the signal at different points of the cascaded CAB-s, depending on the programmed filter order. The block diagram of the circuit is illustrated in Fig.2. Each switch corresponds to a fully differential signal path.

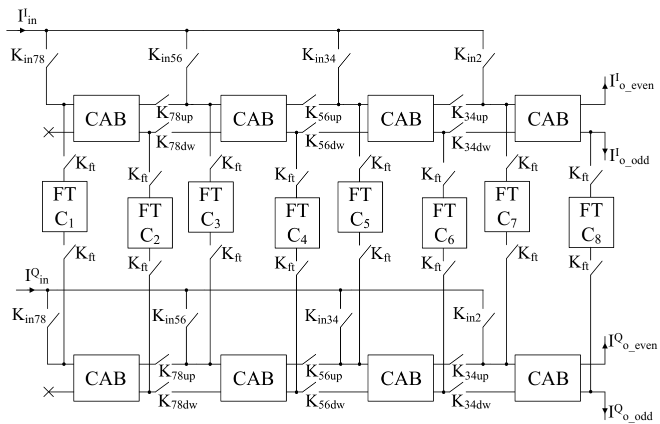


Fig. 2. The proposed FPAA block diagram

In order to select the filter order, the input signal passes through an analog demultiplexer built with switches K_{in2}, K_{in34}, K_{in56} and K_{in78}. For example, if a 6^{th} order filter is desired, the input current is injected through K_{in56} to the appropriate location in the CAB chain. The mapped filter order is controlled by the switch network according to Table 1. The output current is then extracted either at the odd or even output of the circuit. A complex polyphase filter can be obtained by closing the K_{ft} switches that connect the frequency translation network between the I and Q quadrature signal paths.

The CCII based architecture of the CAB-s may be defined according to the supported current mode filter synthesis methods. Two of the often used design techniques employ cascades of second order sections (cascade of biquads) or the functional simulation of a passive ladder prototype (leap-frog filters).

2.2 CCII Based Current Mode Filters

The most often used second order section is the Tow-Thomas (TT) biquad. Its CCII based implementation can be derived from the classical opamp version by

Table 1. Switch positions according to the filter type and order

Order	**2**	**3**	**4**	**5**	**6**	**7**	**8**
K_{in2}	ON						
K_{in34}		ON	ON				
K_{in56}				ON	ON		
K_{in78}						ON	ON
K_{34up}		ON	ON	ON	ON	ON	ON
K_{34dw}		ON	ON	ON	ON	ON	ON
K_{56up}				ON	ON	ON	ON
K_{56dw}				ON	ON	ON	ON
K_{78up}						ON	ON
K_{78dw}						ON	ON

considering voltage-current network duality. The single output CCII circuit can be simplified when double output is considered [8][9]. The resulting schematic is given in Fig.3.

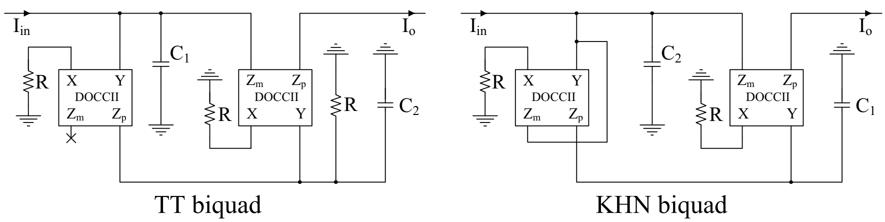


Fig. 3. The TT and KHN second order sections based on DOCCII-s

The corresponding transfer function can be written as in (1).

$$H(s) = \frac{\frac{1}{R^2 C_1 C_2}}{s^2 + \frac{1}{RC_2} s + \frac{1}{R^2 C_1 C_2}} \tag{1}$$

Another second order section, used mainly when low pass, band pass and high pass responses are required is the Kervin-Huelsman-Newcomb (KHN) biquad. Its schematic is given in Fig.3. For the present application only the low pass transfer function is of interest, which is identical to $H(s)$ written for the TT biquad in (1).

A comparison of the two second order sections shows that the only differences are the resistor connected in parallel with C_2 in the TT biquad and the local negative feedback around the first DOCCII in the KHN biquad.

Leap-frog (LF) filters are implemented by functional simulation of the equations and of the signal flow graph associated with a passive ladder prototype [10]. The filter topology is strictly related to the state variables, identified as

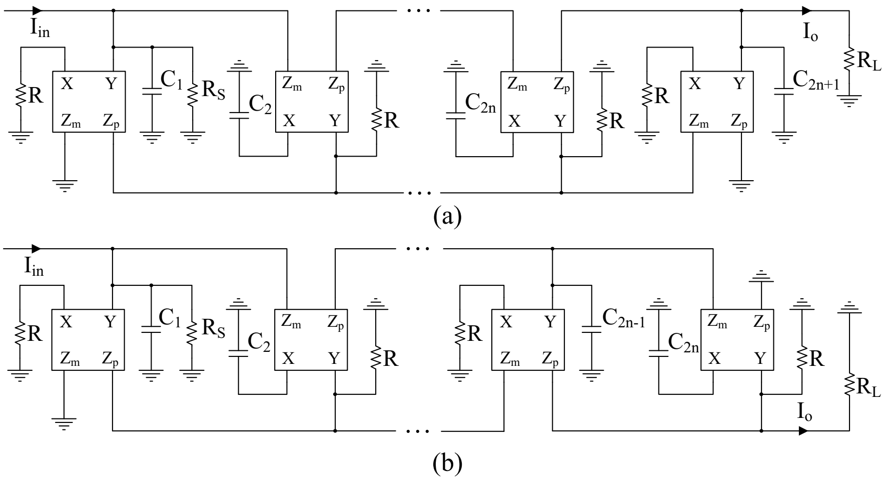


Fig. 4. The (a) odd and (b) even order low pass state variable leap-frog filter

branch currents when current mode operation is required. The generalized all-pole, DOCCII-based realization of a leap-frog low pass filter is illustrated in Fig.4.

2.3 The Configurable Analog Block

The CAB has been designed to support a cascade of TT and KHN second order sections or the leap-frog architecture. The internal structure and the switch positions required to reconfigure the circuit topology can be found by identifying similarities and differences between a second order section of the state variable filter and the biquads in Fig.3. The typical second order section of the leap-frog filter is given in Fig.5. The resistors R_{2k} and R_{2k-1}, marked with dotted lines, are normally not required, unless the section is used as the input or the output of the filter. In these cases the resistors implement the ladder terminations R_S or R_L from Fig.4.

Due to the required generality, each CAB must be able to individually implement a second order filter and behave as both input block and termination for

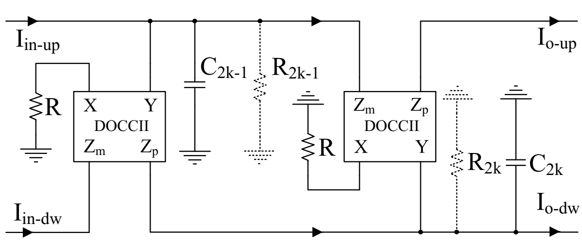


Fig. 5. A second order section in leap-frog filters

the leap-frog chain. A comparison of the second order sections from Fig.3 and Fig.5 leads to the following components of the CAB-s:

- two lossy, current mode integrators implemented with DOCCII-s [5][9];
- a local negative feedback path closed between the Y and Z outputs of the first DOCCII;
- a pair of independent inputs, I_{in-up}-I_{in-dw} and a pair of outputs, I_{o-up}-I_{o-dw}, both pairs being used for cascading

The resulting CAB structure is given in Fig.6.

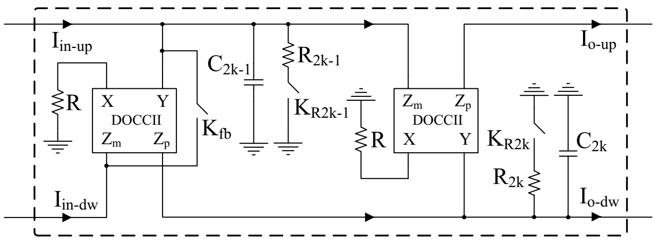


Fig. 6. The proposed CAB architecture

The circuit can be replaced directly in the overall FPAA circuit from Fig.2, regardless of the filter type and order. The internal configuration of the topology depends on the synthesis technique and it allows the selection between even and odd orders. Table 2 shows the switch positions for the considered synthesis techniques and the supported first and second order functional elements. The TT and KHN biquads may be regarded as full second order sections, while the lossy integrator is useful for building odd order cascades. The LF configuration corresponds to a complete intermediate section of the state variable filter. The input termination requires the implementation of the source resistance R_S. From the FPAA architecture in Fig.2 results that R_S is always realized as a R_{2k-1} resistance. The output termination must be differentiated depending on the imposed order. For even order filters the load resistance R_L is implemented as R_{2k} and the output current is sensed at the I_{o-up} output. For odd orders R_L is R_{2k-1} and the output current is taken from I_{o-dw}.

Table 2. Switch positions and output according to the filter type

Synthesis	K_{R2k-1}	K_{R2k}	K_{fb}	**Output**
TT biquad		ON		I_{o-up}
KHN biquad			ON	I_{o-up}
lossy integrator	ON			I_{o-dw}
LF				$I_{o-up}+I_{o-dw}$
LF input	ON			$I_{o-up}+I_{o-dw}$
LF output (even)		ON		I_{o-up}
LF output (odd)	ON			I_{o-dw}

2.4 The Frequency Translation Circuit (FT)

Polyphase band pass filters can be implemented by replacing the classical bilinear frequency transformation with the linear frequency transformation given in (2) that shifts the low pass frequency response either to positive or negative frequencies.

$$s = j\omega \ \rightarrow s^* = j\left(\omega \pm \omega_C\right) \tag{2}$$

where ω_C is the band pass filter center frequency. The bandwidth can be calculated by doubling the original low pass -3dB cut-off frequency [11][8].

The transformation is directly performed on the frequency variable ω. Consequently, it affects only capacitors in the low pass filter without changing the low pass prototype. Each capacitor can be transformed according to (3) [11].

$$j\omega C \ \rightarrow j\omega C + j\omega_C C \tag{3}$$

The frequency transformation, applied to each capacitor in the circuit, may be implemented with a gyrator like structure that is connected as a bridge between two identical, quadrature signal paths. The resulting DOCCII based circuit, usable directly in the FPAA of Fig.2, is illustrated in Fig.7.

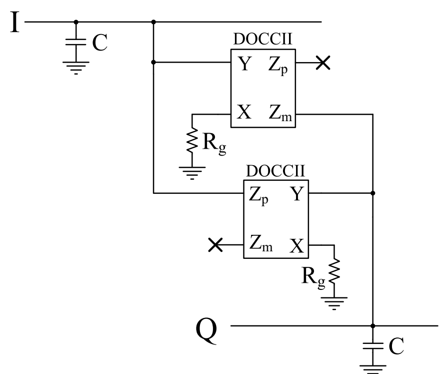


Fig. 7. The linear frequency translation circuit

3 Simulation Results

The functionality of the FPAA has been extensively simulated for all possible configurations with different orders, synthesis techniques and filter types. The simulations have been performed with the Eldo simulator from Mentor Graphics. The setup uses DOCCII-s and switch models with first order non-ideal behavior, including parasitic elements and frequency dependent transfer functions. The simulation examples given in this paper are a 5^{th} order low pass leap-frog filter and a 6^{th} order TT complex band pass filter. The low pass filter has been designed with a standard Chebyshev approximation with 0.5dB pass-band ripple and 2MHz cutoff frequency. The parameters of the band pass filter have been calculated for a Chebyshev type response, 10MHz center frequency and 4Mhz bandwidth. The configuration of the switches for different orders can be inferred from Table 1, while the CAB topologies are defined in tables 3 and 4.

Table 3. Switch positions and output for the 5^{th} order low pass leap-frog filter

	K_{R2k-1}	K_{R2k}	K_{fb}	**Output**
CAB1				
CAB2	ON			
CAB3				
CAB4	ON			I_{o-dw}

Table 4. Switch positions and output for the 6^{th} order complex band pass TT filter

	K_{R2k-1}	K_{R2k}	K_{fb}	**Output**
CAB1				
CAB2		ON		
CAB3		ON		
CAB4		ON		I_{o-up}

Additionally, in order to implement the band pass transfer function, all K_{ft} switches from the FPAA must be closed. In this case the CAB configuration is identical for the I and Q signal paths.

Fig.8 shows the magnitude response of the 5^{th} order Chebyshev low pass leap-frog filter obtained with the topology defined in tables 1 and 3. The simulated cut-off frequency and the pass-band ripple is in accordance with the imposed design specifications.

The 6^{th} order Chebyshev complex TT biquad filter has been derived from the low pass prototype by shifting the 2MHz cut-off frequency response to a 10MHz

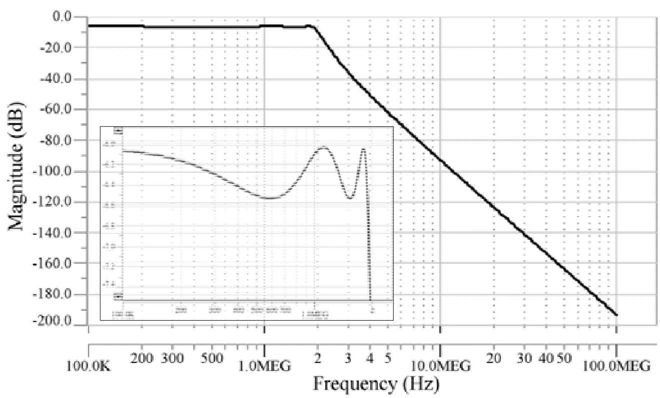


Fig. 8. Magnitude response of the 5^{th} order Chebyshev leap-frog filter

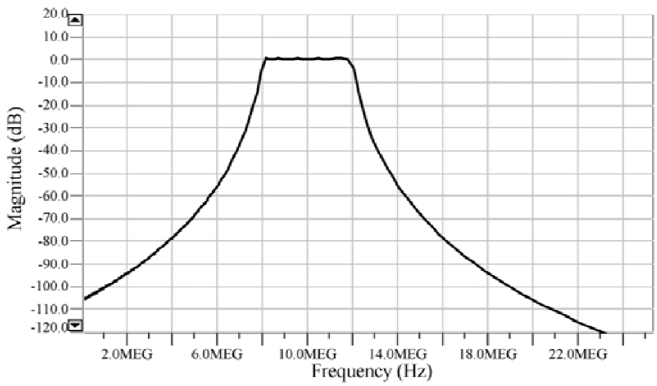


Fig. 9. Magnitude response of the 6^{th} order Chebyshev complex TT biquad filter

center frequency. The topology of the filter can be defined by using tables 1 and 4. Fig.9 illustrates the simulated magnitude response of the filter.

4 Conclusions

In this paper a reconfigurable, current mode analog array architecture has been proposed that is suitable for filtering applications in multi-mode wireless front ends. The analog array supports the implementation of low pass and complex band pass filters, synthesized either with cascades of TT and KHN biquads or by functional simulation of LC ladder prototypes. The filter order is variable between 2 and 8, accommodating the required stop band attenuation with the specifications imposed by the dynamics of the received signal and the corresponding communication standard. The fundamental building block of the array is a CAB, built around double output current conveyors. A single CAB is sufficient to implement any of the supported second order sections.

The overall architecture of the FPAA is based on a linear signal bus that acts as a multiplexer injecting the signal into different possible inputs depending on the chosen filter order. All possible switch configurations have been extensively simulated on models including first order non-idealities and have been found fully functional according to the imposed specifications.

Acknowledgements

The authors wish to thank CNCSIS Romania (National Council for Research and Higher Education) for the financial support (grant no. 657/2009). Special thanks to Mentor Graphics Co. for providing software and financial help with our research and publications.

References

1. Cafaro, G., Correal, N., Taubenheim, D., Orlando, J.: A 100 MHz-2.5 GHz CMOS Transceiver in an Experimental Cognitive Radio System. In: SDR Forum Technical Conference (2007)
2. Kenington, P.B.: RF and baseband techniques for software defined radio. Artech House Mobile Communications Series (2005)
3. D'Amico, S., Gianini, V., Baschirotto, A.: A Low-Power Reconfigurable Analog Filter for UMTS/WLAN Receivers. In: Proceedings of Norchip Conference, pp. 179–182 (2004)
4. D'Mello, D.R., Gulak, P.G.: Design Approaches to Field-Programmable Analog Integrated Circuits. Journal on Analog Integrated Circuits and Signal Processing 17, 7–34 (1998)
5. Becker, J., Henrici, F., Trendelenburg, S., Ortmanns, M., Manoli, Y.: A Hexagonal Field Programmable Analog Array Consisting of 55 Digitally Tunable OTAs. In: International Symposium on Circuits and Systems, ISCAS 2008, pp. 2897–2900 (2008)
6. Pankiewicz, B., Wojcikowski, M., Szczepanski, S., Sun, Y.: A Field Programmable Analog Array for CMOS Continuous-Time OTA-C Filter Applications. IEEE Journal of Solid-State Circuits 37, 125–136 (2002)
7. Premont, C., Grisel, R., Abouchi, N., Chante, J.-P.: Current-Conveyor Based Field Programmable Analog Array. Analog Integrated Circuits and Signal Processing 17, 105–124 (1998)
8. Soliman, E.A., Mahmoud, S.A.: New CMOS Fully Differential Current Conveyor and its Application in Realizing Sixth Order Complex Filter. In: IEEE International Symposium on Circuits and Systems (ISCAS), pp. 57–60 (2009)
9. Hwang, Y.-S., Chen, J.-J., Li, J.-P.: New Current-Mode All-Pole and Elliptic Filters Employing Current Conveyors. Springer Journal of Electrical Engineering (Archiv fur Elektrotechnik) 89, 457–459 (2006)
10. Csipkes, D., Csipkes, G.: Synthesis Method for State Variable Gm-C Filters with a Reduced Number of Active Components. In: Mixed Design of Integrated Circuits and Systems (MIXDES), pp. 292–297 (2003)
11. Csipkes, G., Csipkes, D.: Synthesis Method for Gm-C Complex Polyphase Filter Design. In: Mixed Design of Integrated Circuits and Systems (MIXDES), pp. 286–291 (2003)

A Reconfigurable Voltage Reference without Resistors

Lelia Festilă, Lorant Andras Szolga, Mihaela Cîrlugea, and Sorin Hintea

Technical University of Cluj-Napoca, The Faculty of Electronics, Telecommunications and Information Technology, Department of Bases of Electronics, Str. George Baritiu 26-28, 400027, Cluj-Napoca, Romania
{lelia.festila,lorant.szolga,mihaela.cirlugea, sorin.hintea}@bel.utcluj.ro

Abstract. We present a basic structure for a new bandgap voltage reference (BVR) that doesn't need resistors. The gain stage is based on externally linear internally nonlinear (ELIN) techniques and consists of two nonlinearized building blocks described by two generally nonlinear inverse functions F and F^{-1}. The circuit can be reconfigured to deliver a proportional to absolute temperature (PTAT) voltage, or a bandgap voltage reference.

Keywords: Bandgap voltage references, PTAT references, ELIN design, analog circuits.

1 Introduction

Many modern mixed-signal and analog circuits need high precision voltage references. The analog section in particular needs to be biased by good, performant references, which performance determines the accuracy of the whole circuit [8]. Some application specific integrated circuits (ASIC) systems also need both constant and proportional to absolute temperature (PTAT) voltages. Mostly used in different nowadays integrated circuits are bandgap voltage references (BVRs) [1][2][3]. The needed gain is usually adjusted by trimming the ratio of two resistors. In standard CMOS digital technologies, these resistors could occupy a large area or need an extra mask that increases the cost of fabrication [4]. A circuit configuration without resistors could be a good alternative for a BVR. We present in this paper a bandgap reference that does not need resistors. It is based on the externally linear internally nonlinear (ELIN) [5][6] technique.

In the second section of this paper, a general block diagram of BVRs without resistors is presented. Section III presents the proposed of implementation of a BVR and simulations for its basic mode of operation. In section IV another region of operation is revealed and this one permits a better component matching because of the opportunity to use equal bias currents. The validity of the presented model and derived schematics is proved by simulations. In both cases the functionality of the PTAT reference was also proved. Conclusions and remarks are drawn in section five.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 613–622, 2010.

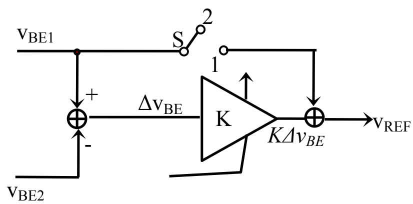
2 A General Block Diagram for Bandgap and PTAT Structures

In order to realise a temperature independent voltage reference, two voltages with equal but opposite temperature coefficients are summed. Usually the negative temperature coefficient is of v_{BE} type that is a base-emitter voltage of a bipolar transistor having a variation of about -2 mV/°C at room temperature. The second term with a positive coefficient may be proportional with thermal voltage V_T.

Fig.1 shows a block diagram based on the above principle. The difference of two base-emitter voltages is proportional to V_T:

$$\Delta v_{BE} = v_{BE1} - v_{BE2} = V_T \ln \frac{I_{C1}}{I_{S1}} - V_T \ln \frac{I_{C2}}{I_{S2}} = V_T \ln\left(n \cdot \frac{A_1}{A_2} \right) \tag{1}$$

where $I_{C1} = nI_{C2}$; $I_{S1,2} = jA_{1,2}$, V_T is the thermal voltage, $I_{C1,2}$ are collector currents, $A_{1,2}$ are base-emitter areas of the transistors, $I_{S1,2}$ are saturation currents and j is saturation current density [3].



1. $v_{REF} = v_{BE1} + K\Delta v_{BE} = v_{BE1} + GV_T$
2. $v_{REF} = K\Delta v_{BE} = G_1 T$

Fig. 1. Principle for a PTAT (2) or a bandgap (1) reference

The sum of the terms with opposite thermal coefficients is obtained if the switch S is in position 1. This results in:

$$v_{REF} = v_{BE1} + K\left(v_{BE1} - v_{BE2}\right) \tag{2}$$

K is the amplifier gain (possibly externally controlled or tuned). Taking into account (1) the two terms v_{BE} and V_T are emphasized:

$$\Rightarrow v_{REF} = v_{BE1} + GV_T \tag{3}$$

where coefficient $G = K \ln \frac{I_{C1}}{A_1}\frac{A_2}{I_{C2}}$ is independent of temperature. It will be set in order to realize an independent of temperature V_{REF}. For example performing the derivative of (3) in a point T_O and denoting $V_T(T_O) = V_{TO}$ one has:

$$\frac{\partial V_{REF}}{\partial \theta} = 0 \text{ if } G = -\frac{\partial v_{BE1}}{\partial \theta} / \frac{\partial V_T}{\partial \theta} = -T_0 \frac{\partial v_{BE1}}{\partial \theta} / V_{T_0} \tag{4}$$

$$\text{because } V_T = \frac{kT}{q} \text{ and } \left.\frac{\partial V_T}{\partial \theta}\right|_{T_O} = \frac{k}{q} = \frac{V_{T_O}}{T_O} \tag{5}$$

k is Boltzmann coefficient and q the electron charge ;

$$\text{if } v_{BE} \cong 0.6\text{V}; \quad V_{TO} \cong 26\text{mV}; \quad T_0 = 300\text{K}; \qquad \frac{v_{BE}}{\theta} \cong -2\frac{\text{mV}}{^{o}\text{C}}$$

relation (4) leads to $G \cong 24$;

$$\Rightarrow \frac{\partial v_{REF}}{\partial \theta} = 0 \;;\; v_{REF} = v_{BE} + GV_T \cong 1.2\text{V}$$

If the switch S is open, the output provides a PTAT temperature.

In order to implement the model in Fig.1, we need a summing amplifier. Traditionally, the amplifier is based on an OpAmp with negative feedback and G is adjusted by resistors.

In ELIN design [5], very popular are nonlinear transconductors and transresistors described by nonlinear continuous F and F^{-1} functions of the form:

$$i_{out} = I \cdot F\left(\frac{v_{in}}{V}\right) \quad ; \quad v_{out} = V \cdot F^{-1}\left(\frac{i_{in}}{I}\right) \tag{6}$$

where v_{in}, i_{in} and i_{out}, v_{out} are input and output signals of transconductors and transresistors respectively. I and V are reference or scaling factors having current and voltage unities respectively. Multipliers, current or voltage controlled amplifiers based only on F and F^{-1} blocks also belong to this design [6]. They have the advantage of simplicity because they do not need extra circuits for linearization. In this way these modules may be very useful in VLSI design. In this paper we propose to use such F - F^{-1} amplifiers to implement the BVR model. With this type of amplifiers realized with transconductors, the model from Fig.1 could have a general configuration as shown in Fig.2, where the first block is a nonlinear transconductor F, the second is a current mirror that introduces a current gain k, and the third block implements the F^{-1} function. This inverse function is also realized with a transconductor F having a negative feedback [6].

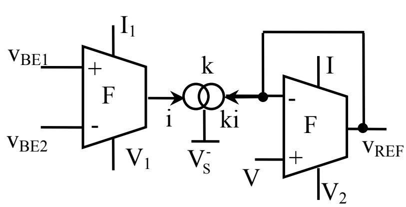


Fig. 2. Block diagram for bandgap voltage references with nonlinear transconductors

$$i = I_1 F\left(\frac{v_{BE1} - v_{BE2}}{V_1}\right) \tag{7}$$

$$ki = I_2 F\left(\frac{V - V_{REF}}{V_2}\right) \tag{8}$$

$$\text{If requirement} \quad I_2 = kI_1 \tag{9}$$

is satisfied, for any F function:

$$V_{REF} = V + \frac{V_2}{V_1}\left(v_{BE2} - v_{BE1}\right) = V + K\Delta v_{BE} \tag{10}$$

where gain $K = V_2/V_1$ depends on the function type and circuit parameters.

Remarks:

- For any function F if requirement (9) is satisfied
  - If $V = v_{BE} \Rightarrow v_{REF} = v_{BE} + GV_T$ as in (3), an operation as a BVR results
  - If $V=0 \Rightarrow v_{REF} = GV_T = G_1T$ that is an operation as a PTAT voltage reference
- If both transconductors are linear, the gain G of V_T includes the gains of these transconductors and mirror, and no requirement as in relation (9) is needed
- Requirement (9) could not be needed for particular nonlinearities F. In these cases expression of V_{REF} could include offsets which may introduce other temperature dependencies that also have to be compensated (see section 4).

3 A Bandgap Reference in the Basic Operating Mode

If the transconductors in Fig.2 are implemented with gate-coupled composed MOS transistors a BVR circuit using transconductors described by an F square law as in Fig.3 results.

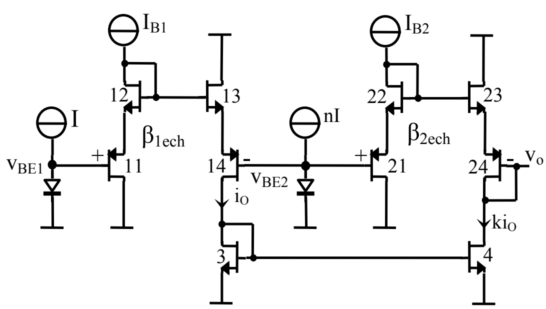
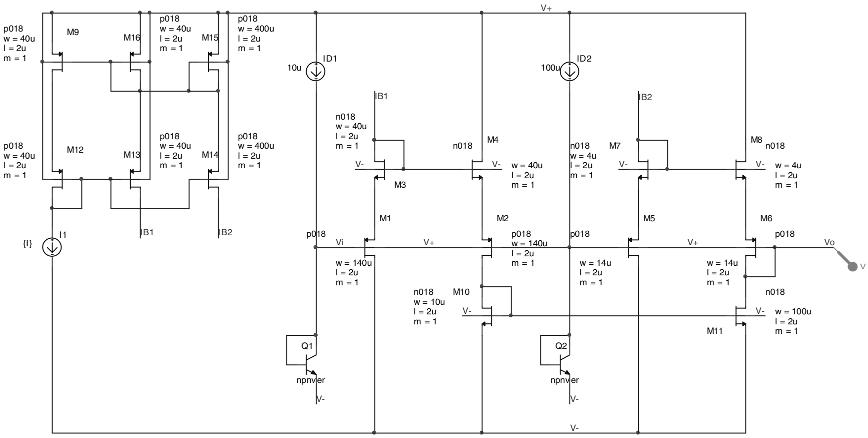
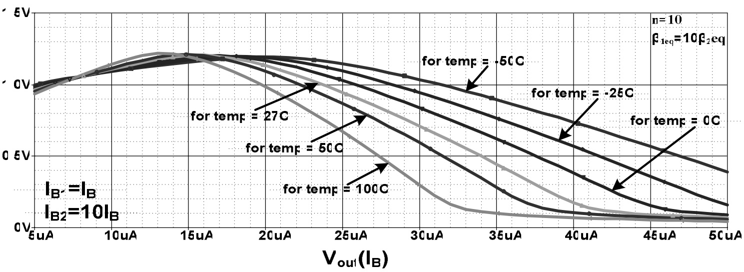


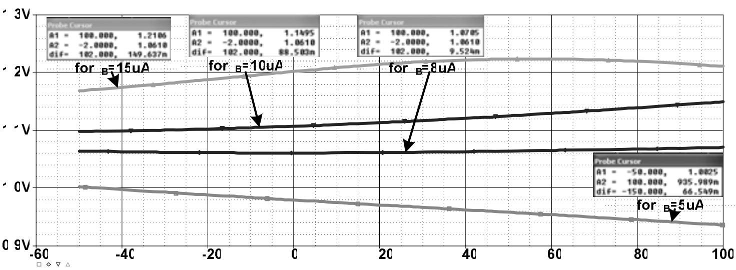
Fig. 3. BVR with gate coupled composite CMOS pairs



a)



b)



c)

Fig. 4. Simulations for I_{B2}=10 I_{B1} ; I_{D2}=10 I_{D1} ; k=10 a) schematic; b) chosing I_{B1} domain; c) temperature characteristic

In this case, functions F describing transconductors (11, 12, 13, 14) and (21, 22, 23,24) and its scaling factors I and V will be respectively:

$$i_o = I_1\left(\frac{V_{BE1} - V_{BE2}}{V_1} + 1\right)^2; \quad I_1 = I_{B1}; \quad V_1 = \sqrt{\frac{I_{B1}}{\beta_{eq1}}} \tag{11}$$

$$ki_o = I_2\left(\frac{v_{BE2} - v_o}{V_2}\right); \quad I_2 = I_{B2}; \quad V_2 = \sqrt{\frac{I_{B2}}{\beta_{eq2}}} \tag{12}$$

Equivalent transconductance factors $\sqrt{\beta_{eq}} = \frac{\sqrt{\beta_p \beta_n}}{\sqrt{\beta_p} + \sqrt{\beta_n}}$; $\beta_{n,p} = \frac{K_{n,p} W}{2L}$, refer to the n,p pairs (11) (12), (13) (14) and (21) (22), (23) (24)

When (9) is fulfilled, relation (10) is valid. Substituting V_1 and V_2 given in (11) and (12) and considering $A_1=A_2$ for diode-connected transistors results in:

$$v_o = v_{BE1} + \left(\sqrt{k\frac{\beta_{1eq}}{\beta_{2eq}}\ln n}\right)V_T \tag{13}$$

with $I_{C2} = n \cdot I_{C1}$ and requirement (8), for this case:

$$I_{B2} = k \times I_{B1} \tag{14}$$

Simulations in Fig.4.c) prove the validity of the proposed schematic. In Fig.4.b) an appropriate domain for biasing is chosing for simulated scheme a).

Table 1.

I_B[μA]	ΔV_{REF} max[mV]	V_{REF}[V]&27ºC	Sensitivity [%]
5	66,549	0,967	1,238761117
8	9,52	1,062	0,161355932
10	88,503	1,115	1,428747982
15	149,637	1,217	2,213201315

In these simulations one can notice that the output voltage can be tuned by bias currents I_B and, V_{REF} has a smaller thermal variation for a current domain between 5μA and 10 μA. A smaller dispersion of voltage reference values in this domain of bias currents can be noticed in Fig 4,c. Table 1 shows some of these results too.

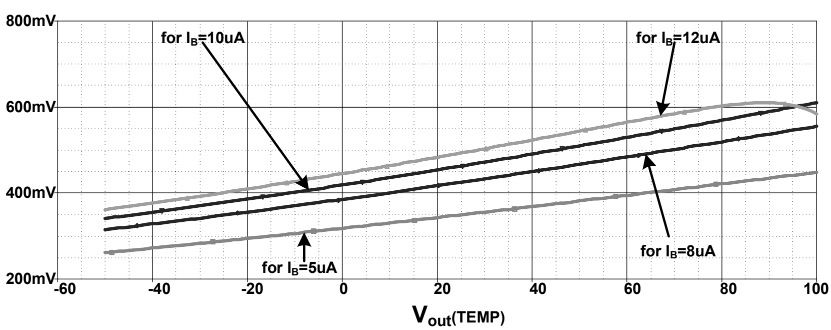


Fig. 5. PTAT reference realized with the circuit in Fig.4,a having input of M5 grounded

For a PTAT Reference the same circuit has the input of M_5 to the ground. Resulted characteristics are shown in Fig. 5. To increase sensitivity β, k, or n may be modified by reconfiguring a parallel network of transistors non figured in Fig. 4,a.

4 BVR with Offset

There are some other opportunities to bias schematic from Fig.3 as a BVR without setting the general requirement (9) because of the particular form of the function F presented in (11). If we consider relations (11) and (12), we can describe the circuit in Fig.3:

$$\frac{v_{BE1} - v_{BE2}}{V_1} + 1 > 0 \tag{15}$$

$$\frac{v_{BE2} - v_0}{V_2} + 1 > 0 \tag{16}$$

$$v_o = V_T\sqrt{\frac{k\beta_1}{\beta_2}}\ln\frac{I_{C2}}{I_{C1}}\cdot\frac{A_2}{A_1} + v_{BE2} + \sqrt{\frac{I_{B2}}{\beta_2}}\left(1 - \sqrt{\frac{kI_{B1}}{I_{B2}}}\right) \tag{17}$$

If $kI_{B1}=I_{B2}$, we obtain the conventional BVR equation (13). For this function we can also set $I_{B1}=I_{B2}$, but the output will have an offset, so that it becomes of the form:

$$v_o = GV_T + v_{BE} + V_{OS}, \quad \text{where } V_{OS} = \sqrt{\frac{2LI_{B2}}{\mu C_{ox}W}}\left(1 - \sqrt{k}\right) \tag{18}$$

Like V_T, this offset also introduces a positive slope in the temperature variation of v_o due to μ and it can be considered an extra compensation opportunity. For example one may consider the temperature variation of mobility μ of the form [7]

$$\mu = \mu_0\left(\frac{T}{T_0}\right)^{-m} \tag{19}$$

where μ_0 is the mobility at the reference temperature T_0 and m a grading factor, $m \in [1,2]$. For instance, if m=2 the requirement to have no temperature variations for the output is:

$$\begin{aligned}\frac{dv_o}{d\theta} &= G\frac{dV_T}{d\theta} + \frac{dV_{OS}}{d\theta} + \frac{dv_{BE}}{d\theta} = \\ &= G\frac{V_{T_0}}{T_0} + \frac{V_{OS}^o}{T_0} + \frac{dv_{BE}}{d\theta} = 0\ ;\quad V_{OS}^o = \sqrt{\frac{2LI_{B2}}{\mu_0 C_{ox}W}}\left(1 - \sqrt{k}\right)\end{aligned} \tag{20}$$

$$\text{If } k<1 \text{ and } \quad G\frac{V_{T_0}}{T_0} + \frac{V_{OS}^o}{T_0} = -\frac{dv_{BE}}{d\theta_0} \cong +2\frac{mV}{°C} \tag{21}$$

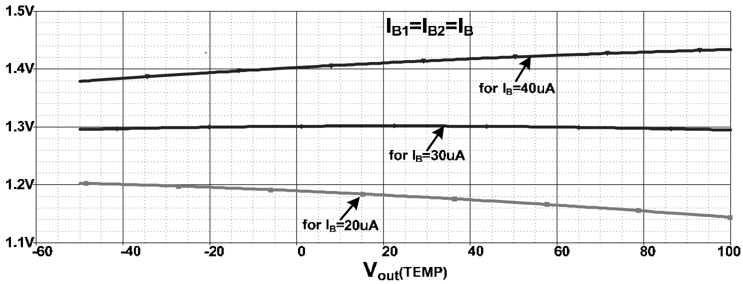


Fig. 6. Output control by IB for a BVR including an offset

there are no temperature variations and the components of output reference are approximately:

$$v_{BE} + GV_{T_0} + V_{OS}^{o} \cong 1.2V \qquad (22)$$

The reference voltage will be of the same BVR value approximately but it could be adjusted in a larger domain. The circuit could be more advantageous because it uses equal bias currents I_B. These ones can be used for tuning or controlling the reference voltage in certain domain of magnitudes. Fig.6 and Table 2 present some of results regarding influence of I_B over V_{REF} and its sensitivity.

Table 2.

I_B[μA]	ΔV_{REF}[mV]	V_{REF} N[V]	Sensitivity [%]
20	58,945	1,1797	0,899389675
30	7,2	1,3017	0,099562111
40	55,264	1,4132	0,703900368

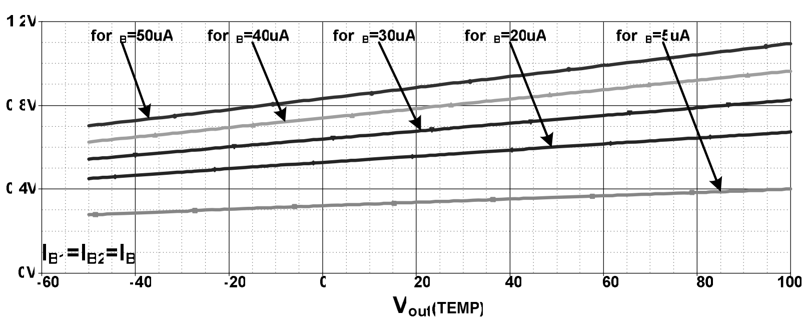


Fig. 7. PTAT reference

Remarks:

• If V=0 a PTAT references results:

$v_{REF} = \sqrt{k}V_T \ln n \frac{\beta_1}{\beta_2} + \sqrt{\frac{I_{B2}}{\beta_2}}\left(1-\sqrt{k}\right)$ for $I_{B2}=I_{B1}$ and $0< k<1$. Simulations are shown in Fig.7 A large influence of I_B over output values can be seen due to the offset

• If k>1 the slope of V_{OS} would become negative when temperature increases and is opposite to V_T variations.

5 Conclusions

A new bandgap reference circuit has been proposed. It can be switched to deliver a voltage reference stable at temperature variations or a proportional to absolute temperature (PTAT) reference.

The output voltage can be tuned to its optimum value by adjusting or controlling bias currents.

The circuit could also provide a PTAT voltage with negative slope if the current mirror of the $F\text{-}F^{-1}$ amplifier is reconfigured to have a reflexion factor k greater than 1.

We only focused in this paper on testing the validity of the proposed schematic and its opportunity to be controlled and reconfigured, without performing an optimal design for the best performance, by using a professional tool as given for example in [9].

Acknowledgement

The authors wish to thank CNCSIS (National Council for Research and High Education) for the financial support (grant no. 657/2009) and Mentor Graphics for their support in the current research.

References

1. Fayomi, C.J.B., Wirth, G.I., Achigni, H.F., Matsuzawa, A.: Sub 1V CMOS Bandgap Reference design techniques: A Survey. Analog Integrated Circuits Signal Processings 62, 141–157 (2010), doi:10.1007/s10470-009-9352-4
2. Guan, X., Wang, X., Wang, A., Zhao, B.: A 3V 110μW 3.1ppm/°C Curvature-Compensated CMOS Bandgap Reference. Analog Integrated Circuits Signal Processings 62, 113–119 (2010), doi:10.1007//s10470-009-9333-7
3. Sansen, W.M.C.: Analog Design Essentials. Springer, Heidelberg (2008), ISBN-13 978-0-387-25746-4
4. Buck, A.E., McDonald, C.L., Lewis, S.H., Viswanathan, T.R.: A CMOS Bandgap Reference Without Resistors. IEEE Journal of Solid-State Circuits 37(1), 81–84 (2002)
5. Tsividis, Y.: Externally Linear, Time Invariant Systems and Their application to Companding Signal Processing. IEEE Transactions on CSII 44(2), 65–84 (1997)
6. Feştilă, L., Szolga, L., Cîrlugea, M., Groza, R.: Analog Multiplying/Weighting VLSI cells for SVM classifiers. In: Lovrek, I., Howlett, R.J., Jain, L.C. (eds.) KES 2008, Part III. LNCS (LNAI), vol. 5179, pp. 360–367. Springer, Heidelberg (2008), ISSN 3029743, ISBN-978-3-540-855675

7. Guo-Yi, Y., Xue-Cheng, Z.: A novel low-voltage high precision current reference based on subthreshold. MOSFETS Journal of Zhejang University Science A 8(1), 50–55 (2007), http://www.springerlink.com
8. Ramos, F.G.R., Ferreira, L.H.c., Pimenta, T.C.: A Programmable Voltage Reference Optimized for Power managemet applications. Analog Integrated Circuits Signal Processing 57, 3–9 (2008), doi:10.1007/s10470-008-9208-3, http://www.springerlink.com,
9. Biagrtti, G., Orcioni, S., Turchetti, C., Alexandrini, M.: SiSMA-A tool for efficient analysis of analog CMOS integrated circuits affected by design mismatch. IEEE Trans. On Computer-Aided Design of Integrated Circuit and Systems 23(2), 192–207 (2004)

A Double-Layer Genetic Algorithm for Gm-C Filter Design

Paul Farago, Sorin Hintea, Gabriel Oltean, and Lelia Festila

Technical University Cluj Napoca, Romania
paul.farago@bel.utcluj.ro
www.bel.utcluj.ro

Abstract. Although analog circuits play an important role in Systems-on-a-chip, their design is effort and time consuming. Automated design methodologies are elaborated to overcome drawbacks resulting from human design. This paper proposes a double-layer on-line genetic algorithm-based optimization method for use in the automated design of Gm-C filters. To accomplish on-line circuit evolution, a Matlab-Eldo interface is proposed for communication of the GA with the circuit simulation environment. After a presentation of the Gm-C filter with an analysis of filter tunability, the two layers of the evolution are presented: raw filter design and fine-tuning of the filter characteristic. Simulation of the evolutionary algorithm proves the efficiency of the double-layer approach in reducing design time for a GA-only optimization technique.

1 Introduction

In mixed-signal systems-on-a-chip (SOC) an important role is played by the analog part, for interfacing functions with the analog world [1]. Although the analog part is a small fraction of the complete system, its design requires increased time and effort compared to the digital counterpart, thus leading to increased design costs [1]. Analog circuit design translates to a complex, high-dimensional, multi-objective and multi-constrained optimization problem [2]. It is still performed by qualified engineers whose training is also expensive and time-consuming.

For digital circuit design, specifications are written in a hardware description language, e.g. VHDL or Verilog, and circuit synthesis is done with minimal human intervention [3]. For analog circuit design however, automated design tools are still an issue of research. The answer may lie in evolutionary computation [4], namely Genetic Algorithms (GA), which simulate natural evolution in order to handle multi objective optimization tasks.

Analog circuit design can be divided into topological-level and parameter-level design [5]. Topological-level design has found a solution in Genetic Programming [6]. This article deals with the parameter-level design, that is, with circuit sizing. Several solutions for optimizing the automated circuit design have been formerly proposed, among which: hybrid optimization [3], fuzzy optimization [7], differential genetic algorithms [5], etc. Each solution is based on a GA and exhibits several layers of optimization.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 623–632, 2010.

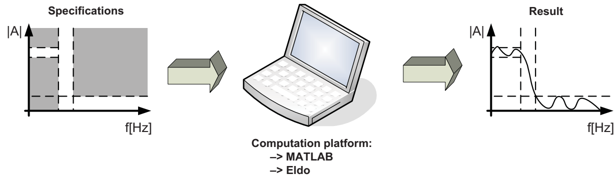


Fig. 1. The diagram of the GA-based filter design application

This article proposes a double-layer on-line GA-based optimization method for the design of programmable active Gm-C filters. A diagram of the targeted application is shown in Fig. 1. Based on a number of design specifications the algorithm generates a fully designed filter.

The proposed double-layer algorithm makes use of the programmable aspect of Gm-C filters [8]. Two design goals are set: an approximate circuit sizing to approach the desired characteristic and a fine tuning of the resulting structure to fully meet the design specifications. Compared to [3], [5] and [7], the proposed solution implements a GA-only optimisation technique.

The paper is organized as follows. Paragraph 2 presents a Matlab-Eldo interface implemented for solving the on-line circuit design problem. Paragraph 3 describes the Gm-C filter topology, attention being paid to filter programmability. Paragraph 4 presents the implementation of the proposed double-layer GA. Finally, paragraph 5 presents the experimental results which lead to a number of conclusions.

2 The Matlab-Eldo Interface

The main concern with on-line circuit evolution is that evolution and circuit simulation are performed within two different software media, both hosted on the

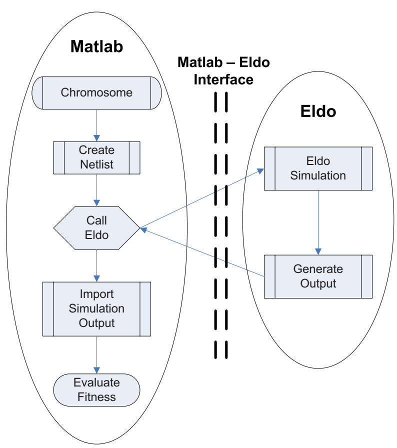


Fig. 2. The evaluation of the fitness function over the Matlab-Eldo interface

same platform as shown in Fig.1. The GA is implemented in Matlab [9] whereas circuit simulation is performed with Eldo, available from Mentor Graphics [10]. Thus, a communication needs to be established between the two media.

In case of on-line evolution, circuit simulation within the evolutionary loop is required. Three steps need to be performed: encoding the evaluated chromosome into a circuit netlist, call of the simulator over a Matlab-Eldo interface, and importing the simulation results into Matlab. A block diagram of the fitness evaluation algorithm is given in figure Fig.2.

The Matlab-Eldo interface is implemented as follows. The circuit netlist and the simulation definitions are saved into a file format that can be interpreted by Eldo, that is a $*.cir$ file [10]. The simulator is then called from Matlab. Eldo interprets the circuit netlist and the simulation definitions and outputs the desired data. The output data is imported to Matlab and is subject for further analysis.

3 The Gm-C Filter Topology

The active filter topology is implemented by simulation of a passive LC filter prototype. An inductor from the passive prototype is simulated using operational transconductance amplifiers (OTA) in a gyrator + capacitor structure [11], as shown in Fig.3 (a) and (b). The equivalent inductance value is given by equation (1).

$$L = \frac{C}{g_m^2} \tag{1}$$

As such, implementation of a Gm-C active filter topology resumes to the synthesis of a passive LC network. This relies on a few ground-rules listed in Table 1 [11]. A 5th order ladder Gm-C LPF with elliptic approximation is shown in Fig.4.

The frequency response of the Gm-C filter depends on both the OTA gains and the capacitor values [11]. The OTA gain is in a direct dependence with the OTA bias current [1]. The OTA bias current has thus a direct influence on the filter characteristic, specifically the pass-band gain and the cutoff frequency. In [8], a solution is presented that provides a digital control of the OTA gain, by having its bias current sourced with a current division network (CDN). Such a structure will be further on referred to as a programmable gain OTA (PG-OTA) and is shown in Fig.5. The expressions of the bias current and the transconductance are given in equations (2) and (3), where $a = a_1 a_2 ... a_n$ is the digital control word. The design

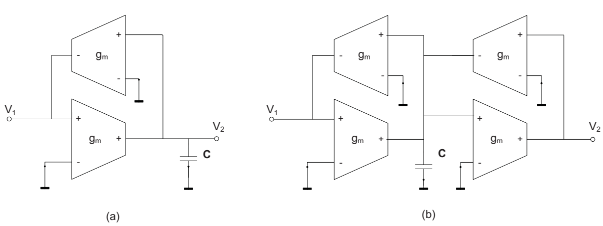


Fig. 3. Gm-C simulation of a (a) grounded inductor (b) floating inductor

Table 1. Ground rules for the implementation of passive LC filters

Filter Type	Longitudinal Impedance	Transversal Impedance
LPF - all pole	L	C
HPF - all pole	C	L
BPF - all pole	Series LC	Parallel LC
BSF - all pole	Parallel LC	Series LC
Arbitrary Transmission Zeros	Parallel LC	Series LC

of the PG-OTA is detailed in [12]. By use of PG-OTAs for the implementation of the Gm-C filter, a digital control of the filtering characteristic is achieved.

$$I_{bias} = \sum \bar{a}_i \cdot \frac{I_{in}}{2^i} \tag{2}$$

$$g_m = \frac{2 \cdot I_{bias}}{V_{GS} - V_{th}} = \frac{2 \cdot \sum \bar{a}_i \cdot \frac{I_{in}}{2^i}}{V_{GS} - V_{th}} \tag{3}$$

The design of the Gm-C filter consists in determining the appropriate OTA gains and capacitor values. Admitting that the OTA gain values are known, as is the case in practice, the filter design reduces to only determining the capacitor

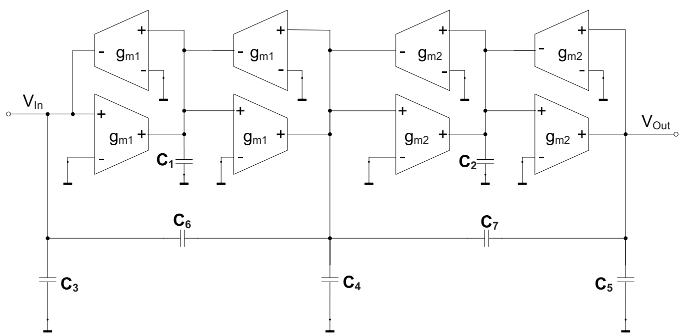


Fig. 4. The Gm-C lowpass ladder filter

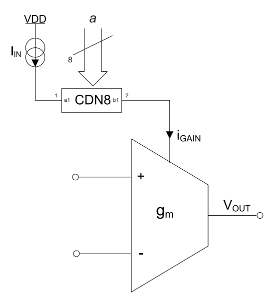


Fig. 5. The schematic of the Programmable Gain OTA

values. However, employment of the PG-OTA allows a fine-tuning of the filter characteristic by appropriate change of the digital control words [8]. This splits the design procedure in two steps: the determination of the capacitor values - hard tuning, and the adjustment of the PG-OTA control words - fine tuning.

4 The Double-Layer Genetic Algorithm for Filter Design

This paragraph presents the implementation of a GA suitable to solve the design problem of the tunable Gm-C filter. A block diagram of the proposed double-layer GA is shown in Fig.6. The implementation of the two evolutionary loops is based on [3].

In both GA layers, the genetic operators are implemented as follows. Selection is done based on a stochastic universal sampling scheme, after individuals have been arranged with a rank-based scaling function. A number of 2 elite individuals are chosen to survive to the next generation. Reproduction is then performed with a crossover factor of 0.8 and a mutation factor of 0.2. The crossover operator is implemented based on a multi-point crossover scheme. The mutation operator is implemented with a Gaussian probability function, with an initial variance of 1, that linearly decreases to 0 after 100 generations. A detailed description of the genetic operators is given in [4].

The first evolutionary loop performs a raw design of the Gm-C filter. The objective of this loop is to generate a raw approximation of the desired frequency response. At this stage, the algorithm determines the appropriate capacitor values. The function tolerance is set to a relatively high value, i.e. $10^{-5}...10^{-4}$, in order to break the optimization process when the design criteria are partially met. The partially fit solution is subject to a second evolutionary loop. The algorithm performs a fine-tuning of the filter characteristic by adjusting the PG-OTA control words. Finally, the GA returns the fully optimized solution to the filter design problem.

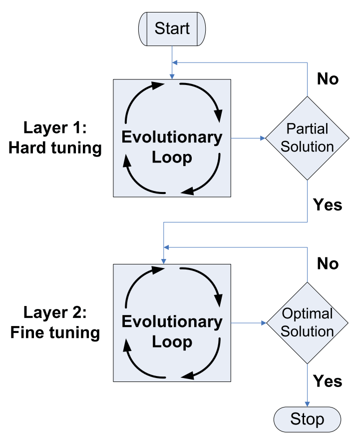


Fig. 6. The double-layer Genetic Algorithm structure

Both GA loops run with a population of 50 individuals. The parameters of the individuals which form the population of solutions are encoded as follows. The capacitor values are represented with double variables. Out of practical considerations regarding circuit integration, the capacitor values are constrained to a maximum of $35pF$. The OTA gains, given by the corresponding CDN control words, are encoded as binary strings and constrained to vary around $200\mu S$. The objective function is to minimize the square error between the desired (*spec(f)*) and the available (*amp(f)*) frequency characteristics. The cost function is given in equation (4). A penalty factor (*p=20*) has been added to increase the cost of the pass-band gain error. The evaluation of the objective function is done on a frequency sweep basis, over the Matlab-Eldo interface proposed in paragraph 2.

$$cost = p \cdot \sum (amp(f) - spec(f))^2 \tag{4}$$

5 Simulation Results

The filter design specifications are depicted in Fig.7. The simulation of a ladder filter requires an input and a terminating resistance [11] which explains the 6dB attenuation in the pass-band. The aim is to optimize the frequency response in order to fit into the set of specifications. The optimization approach is to minimize the error between the targeted and the available filter characteristic.

Several applications will be presented and the simulation time will be measured for each. The simulation time includes: Matlab computation for the GA iterations, circuit simulation and entering and returning from the simulator which is actually the most time-consuming ($\approx$ 3 seconds/individual).

A first GA application consists of a single-layer GA for designing the Gm-C elliptic filter. The OTA gains are set to $200\mu S$ and only the capacitor values need to be determined. The chromosome is defined as a vector of 7 double variables constrained to $35 \cdot 10^{-12}$, each representing the value of a capacitor. The stopping criterion is the fulfillment of the optimization criteria with a tolerance of 10^{-6} for a number of 50 generations.

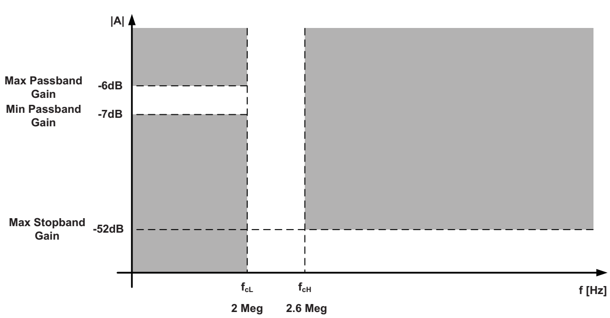


Fig. 7. The filter design specifications

(a)

(b)

(c)

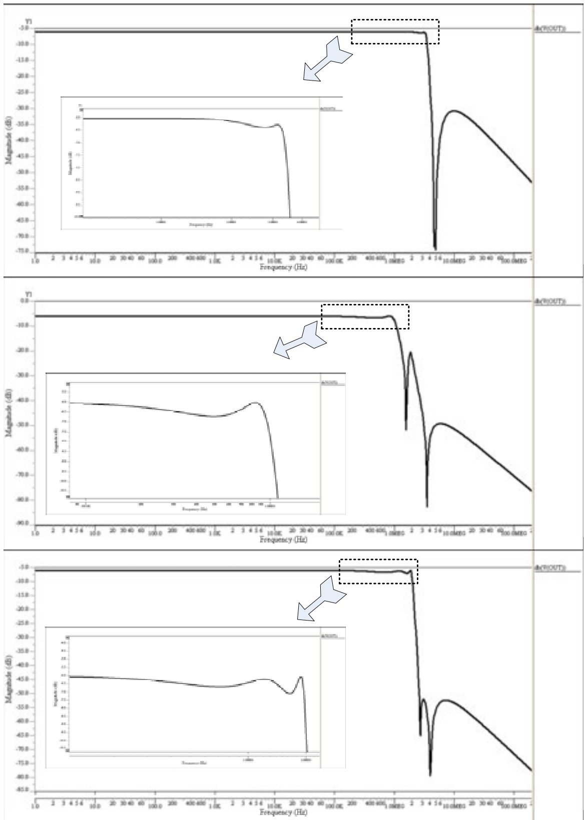


Fig. 8. The Gm-C filter charatcteristic in the (a) 2^{nd} generation, (b) 14^{th} generation, (c) final generation

Figures 8 show 3 different steps in the evolutionary process. Figure 8 (a) shows the frequency characteristic of the fittest filter in the second generation. Similarly, Fig.8 (b) stands for the 14^{th} generation and Fig.8 (c) for the last generation. The resulting capacitor values are: C1=11.5pF; C2=18.5pF; C3=28.2pF; C4=33.7pF; C5=23.4pF; C6=11.5pF; C7=3.37pF.

The evolution lasted for 4 hours and 12 minutes. Figure 9 shows the evolution of the best fitness and the mean fitness scores vs. generations. As shown, after a number of 30 - 40 generations, the GA saturates around a partial solution which isn't significantly improved for the next generations. It is thus sensible to consider breaking the evolutionary loop and move to the next optimization layer.

A second GA application consists of optimizing the Gm-C filter from a stage where the former GA would be halted. For demonstration purpose, consider that point being the 30^{th} generation where the capacitors are fixed to the values: C1=11.5pF; C2=18.8pF; C3=28.3pF; C4=33.8pF; C5=23.4pF; C6=11.8pF; C7=3.41pF.

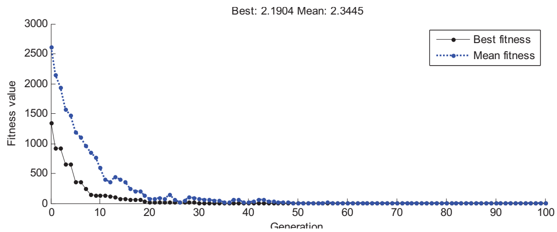


Fig. 9. Evolution of the best fitness and the mean fitness values vs. generations for Application 1

The GA needs to determine the appropriate OTA gains. Considering equation (3),the problem reduces to determining the CDN control words. The chromosome consists of a vector of 16 binary variables, 8 variables corresponding for a PG-OTA gain, noting that 4 PG-OTAs have the same gain as required in (2). This gives an OTA gain variation from $150\mu S$ to $200\mu S$. This application is again implemented as a single-layer GA.

The evolution lasted for 2 hours and 15 minutes. Figure 10 shows the evolution of the best fitness and the mean fitness scores vs. generations. The frequency characteristic of the optimized filter is plotted in Fig.11. The resulting OTA gains are: $g_{m1} = 232\mu$S; $g_{m2} = 186\mu$S.

A third application consists of combining the two evolutionary loops formerly described into one single application, as shown in Fig.6. For the first GA layer, the function tolerance has been increased to 10^{-5} and the number of stall generations was reduced to 30, so the GA loop would break after 40-50 generations. For the second GA layer, the chromosome representation was changed to a vector of two double variables, which were converted into the PG-OTA control words by means of a $double-to-binary$ conversion. As such, the chromosome length was significantly reduced, leading to a smaller number of individuals per generation, i.e. 30 individuals, and a shorter evolution time.

The evolution lasted for 2 hours and 52 minutes. The evolution of the best fitness and the mean fitness vs. number of generations is shown in Fig.12. The increased value of the stall function tolerance caused the first layer loop to break after 38 generations. The next 51 generations, the GA tuned the filter in order

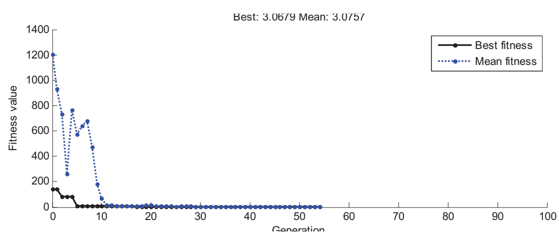


Fig. 10. Evolution of the best fitness and the mean fitness values vs. generations for Application 2

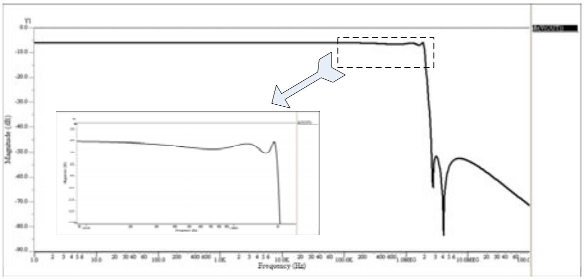


Fig. 11. The Gm-C filter characteristic after fine-tuning

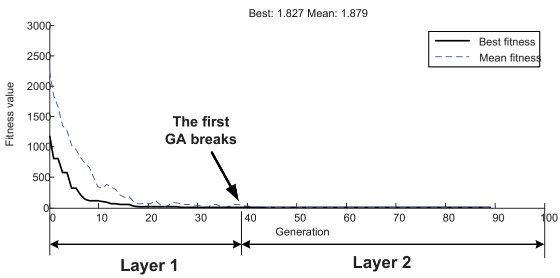


Fig. 12. Evolution of the best fitness and the mean fitness values vs. generations for Application 3

to fully comply with the design specifications. As expected, the result of the double-layer GA is a considerable reduction in evolution time.

6 Conclusions

In this paper, a double-layer GA-based on-line evolutionary algorithm was proposed and used for the automated design of an active lowpass filter. The Gm-C filter networks allow a hard-tuning and a fine-tuning of the filter characteristic. The optimization algorithm makes use of this particularity and splits the design problem in two layers: raw design and fine-tuning. Online evolution required circuit simulation within the evolutionary loop, which was solved by the implementation of a Matlab-Eldo interface. Both layers of the proposed algorithm have first been tested separately to prove their functionality. Then, the complete double-layer GA was run to solve the filter design problem. By the employment of the proposed algorithm, the design time has been considerably reduced.

Acknowledgements

The authors would like to thank CNCSIS (National Council for Research and High Education), grant no 657/2009, and Mentor Graphics, for their support in the current research.

References

1. Laker, K.R., Sansen, W.: Design of Analog Integrated Circuits and Systems. McGraw-Hill, New York (1994)
2. Barros, M., Guilherme, J., Horta, N.: GA-SVM Optimization Kernel Applied to Analog IC Design Automation. In: Proceedings of the 13th IEEE International Conference on Electronics, Circuits and Systems, ICECS 2006, pp. 486–489 (2006)
3. Grimbleby, J.B.: Hybrid Genetic Algorithms for Analogue Network Synthesis. In: IEE/IEEE International Congress on Evolutionary Computation, CEC 1999, Washington DC, pp. 1781–1787 (July 1999)
4. Zebulum, R.S., Pacheco, M.A.C., Vellasco, M.M.B.R.: Evolutionary Electronics - Automatic Design of Electronic Circuits and Systems by Genetic Algorithms. CRC Press, Boca Raton (2002)
5. Liu, B., Wang, Y., Yu, Z., Liu, L., Li, M., Wang, Z.: Analog circuit optimization system based on hybrid evolutionary algorithms. INTEGRATION, The VLSI Journal 42, 137–148 (2009), Jing Lu a, FranciscoV.Fernaǹdez
6. Koza, J.R., Bennett III, F.H., Andre, D., Keane, M.A.: Automated WYWIWYG Design for Both Topology and Component Values of Electrical Circuits using Genetic Programming. In: Genetic Programming 1996: Proceedings of the First Annual Conference, pp. 123–131 (1996)
7. Oltean, G., Hintea, S., Sipos, E.: Analog Circuit Desdign Based on Computational Intelligence Techniques. Journal of Automation, Mobile Robotics and Intelligent Systems 3, 1–7 (2009)
8. Hintea, S., Farago, P., Festila, L., Soser, P.: Reconfigurable Filter Design for Implantable Auditory Prosthesis. Electronics and Electrical Engineering 3(99), 7–12 (2010)
9. `http://www.mathworks.com/products/matlab/`
10. `http://www.mentor.com/`
11. Schaumann, R., Van Valkenburg, M.E.: Design of Analog Filters. Oxford University Press, Oxford (2001)
12. Farago, P.: Design of Reconfigurable Analog Integrated Circuits for Cochlear Implants, Diploma Thesis, Cluj-Napoca (June 2009)

Optimised Dielectric Totally Internally Reflecting Concentrator for the Solar Photonic Optoelectronic Transformer System: Maximum Concentration Method

Firdaus Muhammad-Sukki[1], Roberto Ramirez-Iniguez[1], Scott G. McMeekin[1], Brian G. Stewart[1], and Barry Clive[2]

[1] School of Engineering and Computing, Glasgow Caledonian University, Cowcaddens Road, Glasgow, G4 0BA Scotland, UK
[2] SolarEmpower Ltd,74-75 Brunner Road, London E17 7NW, England, UK
firdaus.muhammadsukki@gcu.ac.uk

Abstract. Renewable energies have become a matter of global attention during the last few years, with solar energy being a major focus. Building Integrated Photovoltaic (BIPV) system is an example of solar energy applications. This paper describes the Solar Photonic Optoelectronic Transformer (SPOT) system, one of the components in the SolarBrane, a BIPV system developed by SolarEmpower Ltd. The general design of a dielectric totally internally reflecting concentrator (DTIRC) is described. The process of designing an optimized solar concentrator using the maximum concentration method (MCM) for the SPOT system is also outlined with detailed steps on the simulation setup presented. The paper then compares the performance of the new concentrator with the current concentrator design. The results from the MATLAB simulation shows that the MCM offers a higher geometrical concentration gain, with a slight increase in the concentrator size.

Keywords: Dielectric totally internally reflecting concentrator; maximum concentration method.

1 Introduction

In the last decade, there has been significant change in the use of the world's energy resources. Both governments and private sectors across the world are trying to cut the dependency on fossil fuels by optimizing the existing systems to minimize the energy consumption, as well as trying to venture into alternative energies. A lot of countries are now focusing their efforts on renewable energy. It has been reported that in 2008 alone, the total global investment in renewable energy reached approximately USD120 billion [1].

Solar energy is one of the alternative energies that has vast potential for meeting energy requirements of the future. One of the solar energy applications is the BIPV system. It consists of integrating PV modules into a building envelope such as the roof or the façade [2]. It is one of the applications that allows collection of sun energy to generate power for a building.

R. Setchi et al. (Eds.): KES 2010, Part IV, LNAI 6279, pp. 633–641, 2010.

2 SPOT System

A good example of a BIPV system is proposed by SolarEmpower Ltd. and is called the SolarBrane [3]. It is a new static solar device. One of the small components that form this device is the SPOT system. Unlike conventional solar photovoltaic (PV) systems which only generates electricity, SolarBrane is able to make full use of direct and indirect solar radiation to produce electricity, hot water, space heating, illumination, ventilation and shade for a building. This will reduce the power consumption of the building, as well as generating carbon credits. Almost 75% of the sunlight is converted to useful energy as compared to 15%[1] in the conventional solar PV [3].

Fig. 1. SolarBrane mounted on the wall of a house

Another distinct advantage of a SolarBrane is that it could be retrofitted in any part of the building as compared to the traditional solar PV which could only be deployed at specified parts of the building (e.g. the roof or the window) and needs to be placed at an optimum angle for maximum efficiency.

To reduce the dependency on expensive silicon as well as to reduce the cost of the system, a solar concentrator is used in the design. A solar concentrator is one of the devices used in the BIPV system that maximizes the collection of solar light and focuses the light to a smaller exit area, at which a PV cell will be attached. While traditional solar PV uses a large area of silicon cell, SolarBrane only uses around 30% of the total silicon whilst maintaining the same output power.

SolarEmpower is currently working closely with the Fraunhofer Institute for the DESOL project [4], with the aim to desalinate sea water entirely using solar technology. By using the same principle, it is possible to demonstrate that enabled by a SolarBrane, water could also be extracted from air. This will make a building become energy and water independent [5],[6].

SolarBrane is also being designed to be used in the deployment of SolarEmpower Greenhouses [5],[6]. While crops are grown in them, these greenhouses also act as the power generator by converting the sunlight into electricity. The water in the air, even in arid conditions, can be extracted using a particularly hydrophilic saline solution

[1] The numerical number varies depending on the type of PV cell used.

which is then desalinated by the heat cogenerated by the photovoltaic effect employed by the SolarBrane. In this way the SolarBrane Greenhouse will simultaneously generate food, power and water which could be ideal for the developing world, remote areas and even cities. Table 1 summarises the advantages of SolarBrane over a conventional Solar PV.

Table 1. The advantages of SolarBrane over a conventional Solar Panel [3], [5]

Subject	Traditional Solar PV	SolarBrane
Electrical power/efficiency	Depend on silicon efficiency	Depend on silicon efficiency
Sunlight conversion to useful energy	15%	75%
Amount of PV	100%	30%
Space Heating	No	Yes
Hot water	No	Yes
Shade	No	Yes
Ambient light	No	Yes
Cooling/Ventilation	No	Yes
Further	No	Water Technology
Deployment	Only at optimum angle	All:East/West/North/ South

The current solution used in the SPOT system employs an optical concentrator extrusion of a DTIRC profile (see Figure 2), and is deposited on a rectangular PV cell cooled by water and air depending upon the application. This concentrator is proven to provide higher optical gains than alternative optical elements such as a conical concentrator or a compound parabolic concentrator (CPC), which reduces the amount of PV cell required. In addition, their large field-of-views (FOVs) can be used to eliminate the need for mechanical tracking [7].

However, further optimization on the existing concentrator is desirable. The current DTIRC design is based on the phase conserving method (PCM) developed by Ning et. al [7]. The optimized design of the concentrator discussed in this paper is based on the DTIRC design using the maximum concentration method (MCM).

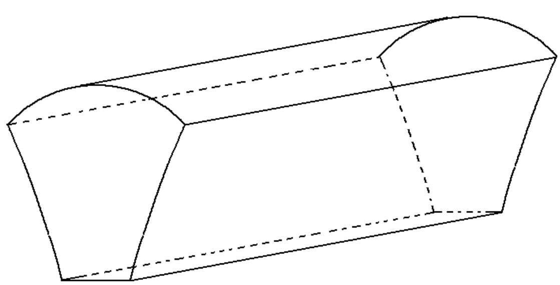


Fig. 2. Solar concentrator in the SPOT system

3 General DTIRC Design

The concept of DTIRC was first introduced by Ning et al. [7] more than two decades ago. This new class of optical element has the capability to achieve concentrations close to the theoretical maximum limits. There are two ways to produce the DTIRC; maximum concentration method and phase conserving method. Although both methods can be used to create similar structure, the first technique offers slightly higher concentration and therefore is more suitable for solar application.

DTIRC consists of three parts; a curved front surface, a totally internally reflecting side profile and an exit aperture (see Figure 3). When the rays hit the front curved surface, they are refracted and directed to the side profile. Upon hitting the sidewall, they are totally internally reflected to the exit aperture. Outside the range of the acceptance angle, the rays exit from the side profile, missing the PV cell [8]. The front aperture can be a hemisphere, but different designs such as parabola and eclipse have been developed recently [9].

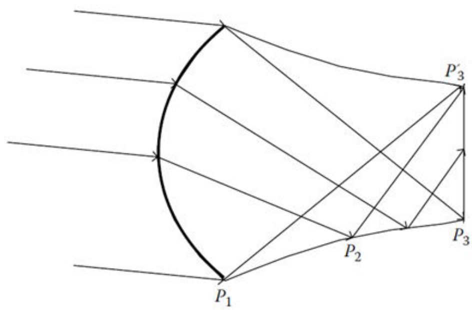


Fig. 3. General design of a DTIRC [7]

The side profile of a DTIRC consists of two parts (see Figure 3); from P1 to P2 and from P2 to P3. Any extreme rays hitting portion P1 to P2 will be directed to the point P3' after a single total internal reflection (TIR). The ray hitting P2 exits at P3' although it barely satisfies the TIR. However, for portion P2 to P3, there are two ways to create this profile.

1. Based on the MCM:
   - Impose a condition such that all the rays will experience a single TIR and arrive at the exit aperture, without restricting the exit angle and the incident angle.
   - This will result in achieving the maximum possible concentration.
2. Based on the PCM:
   - Create a profile that will allow the rays to exit in parallel.
   - This will enable the rays to exit with a well-defined wavefront.

The DTIRC is often compared to the dielectric CPC. The advantage of a DTIRC over a CPC is that it offers higher gain and smaller sizes [7],[8]. Figure 4 compares the

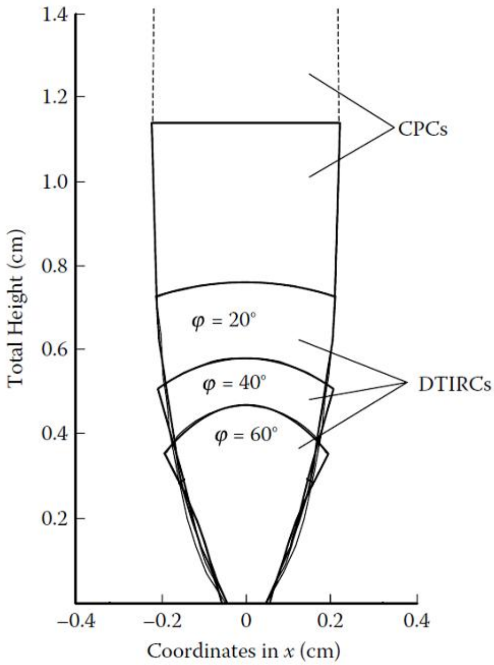


Fig. 4. Comparison of three DTIRCs with different arc angle and a CPC [8]

total height of a DCPC and a variety of DTIRCs. As the front surface arc angle increases the total height of a DTIRC shrinks dramatically. This is useful in producing a more compact concentrator design. An increase in the front surface curvature will also change the side profile from convex to concave [7].

4 Maximum Concentration Method (MCM)

4.1 Design Procedure

Designing a DTIRC based on the MCM is explained in detail by Ning et. al in [7]. To simulate this method in the Cartesian coordinate system, the 2-D solution is obtained first to produce the x and y-coordinates of the profile. The 3-D version is obtained by extending the 2-D coordinates over the z-plane.

The 2-D representation requires six input parameters, the front surface arc angle (φ), the acceptance angle (θ_a), the dimension of exit aperture (d_0), the trial length of the entrance aperture (d_1), the index of refraction of the dielectric (n) and the number of extreme rays (N).

Based on the input variables, a computer program is used to calculate the trial height, which is later used to calculate the coordinates of the side wall. This calculation takes into account the number of extreme rays entering the concentrator at the critical angle. Once it is completed, the program compares the trial entrance aperture with the calculated aperture. A new entrance aperture is computed from the difference

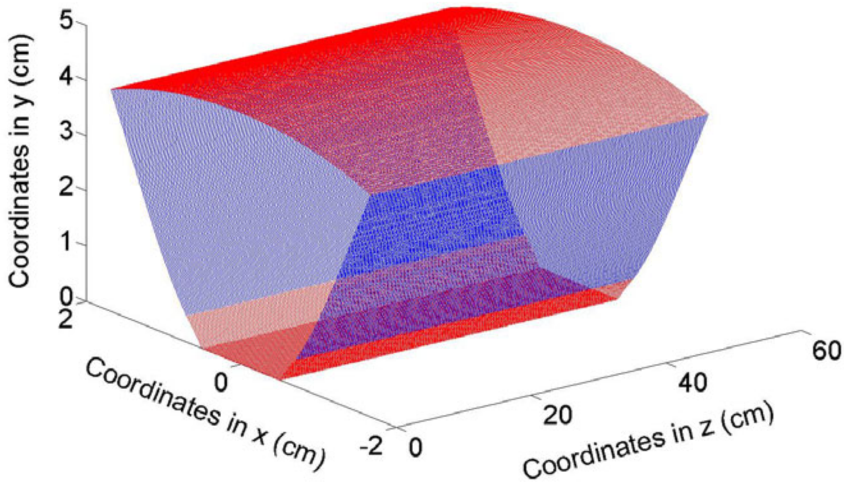


Fig. 5. 3-D design for a DTIRC based on the MCM

between the two apertures. A number of iteration take place until the difference between both apertures is within an acceptable error value [7]. Figure 5 shows the simulated diagram of the new concentrator design based on the MCM.

4.2 Design Analysis

Despite having to include six input parameters in designing a DTIRC, the exit aperture and the refractive index are normally fixed in many cases. The number of extreme rays is used to provide the number of x-y coordinates as well as the numerical precision. The effect that the input parameters have on the geometrical concentration gain, the entrance diameter and the total height is presented in Table 2 and 3. Table 2

Table 2. Comparison of Geometrical Concentration Gain vs Acceptance Angle (Index=1.5, Arc Angle=30^o, Exit Aperture =1cm)

Acceptance Angle (^{o})	Phase Conserving Method [10]			Maximum Concentration Method		
	Entrance diameter (cm)	Total height (cm)	Gain	Entrance diameter (cm)	Total height (cm)	Gain
18.00	4.6740	7.6365	4.6740	4.6995	7.6714	4.6995
20.00	4.1946	6.5778	4.1946	4.2224	6.6136	4.2224
22.00	3.8018	5.7350	3.8018	3.8335	5.7737	3.8335
24.00	3.4737	5.0508	3.4737	3.5092	5.0919	3.5092
26.00	3.1955	4.4864	3.1955	3.2365	4.5315	3.2365
28.00	2.9564	4.0143	2.9564	3.0026	4.0627	3.0026

Table 3. Comparison of Geometrical Concentration Gain vs Arc Angle (Index=1.5, Acceptance Angle=20°, Exit Aperture =1cm)

Arc Angle (°)	Phase Conserving Method [10]			Maximum Concentration Method		
	Entrance diameter (cm)	Total height (cm)	Gain	Entrance diameter (cm)	Total height (cm)	Gain
10.00	4.3154	9.0768	4.3379	4.3175	9.0789	4.3392
20.00	4.2623	7.6048	4.3507	4.2718	7.6170	4.3591
30.00	4.1946	6.5778	4.3931	4.2224	6.6081	4.4178
35.00	4.1545	6.1658	4.4252	4.1999	6.2146	4.4675
40.00	4.1095	5.8011	4.4639	4.1818	5.8815	4.5384
45.00	4.0589	5.4731	4.5089	4.1714	5.5943	4.6288

shows the effect of different acceptance angle while Table 3 indicates the effect of varying the front surface arc angle. The geometrical concentration gain for a 2-D concentrator is defined as the diameter ratio of the entrance to exit aperture [8].

$$\text{Geometrical Concentration Gain, } C = d_1/d_0 \tag{1}$$

In general, the DTIRC gain is inversely proportional to the acceptance angle and to the arc angle. As the front surface becomes more curved, rays are bent more sharply. It is harder to satisfy the total internal reflection, thus the gain is reduced [7]. From Table 2 and Table 3, in both cases, the MCM offers a higher gain than the PCM. Both the total height and the entrance diameter of the new concentrator design are slightly greater than the current concentrator design. This implies that the MCM produces a slightly larger concentrator structure. Tables 4 and 5 show the increase as a percentage in terms of the entrance diameter, total height and the geometrical concentration gain when using the MCM. Table 4 shows the effect of different acceptance angle while Table 5 indicates the effect of varying the front surface arc angle.

Table 4. Percentage of increase in terms of the entrance diameter, total height and the geometrical concentration gain (Index=1.5, Arc Angle=30°, Exit Aperture =1cm)

Acceptance Angle (°)	Percentage of Change (%)		
	Entrance diameter	Total height	Gain
18.00	0.5456	0.4570	0.5456
20.00	0.6628	0.5443	0.6628
22.00	0.8338	0.6748	0.8338
24.00	1.0220	0.8137	1.0220
26.00	1.2831	1.0053	1.2831
28.00	1.5627	1.2057	1.5627

Table 5. Percentage of increase in terms of the entrance diameter, total height and the geometrical concentration gain (Index=1.5, Acceptance Angle=20^o, Exit Aperture =1cm)

Arc Angle (^{o})	Percentage of Change (%)		
	Entrance diameter	Total height	Gain
10.00	0.0487	0.0386	0.0487
20.00	0.2229	0.1828	0.2229
30.00	0.6628	0.5443	0.6628
35.00	1.0928	0.9034	1.0928
40.00	1.7593	1.4583	1.7593
45.00	2.7717	2.3077	2.7717

From both tables, it is evident that the new method is able to improve the gain of the concentrator, which agrees with the result in [7]. The new concentrator however, has a slight increase in terms of size as compared to the existing one.

5 Conclusion

For any BIPV system, it is always desirable to improve the design to obtain the optimised performance. This is also true to the SolarEmpower's BIPV system, the SolarBrane. The main conclusions of this paper are summarised as follows:

1. The SPOT System is one of the components that form the SolarBrane. SolarBrane not only is capable to make full use of solar energy i.e. produce electricity, hot water, space heating, illumination, ventilation and shade for a building, but is also possible to make the building water independent.
2. The SPOT system uses a solar concentrator which focuses the solar energy into a smaller area as well as reducing the dependency on expensive PV, hence reducing the total cost of the system. The current concentrator in the SPOT system is a DTIRC design based on the PCM.
3. An optimized design for the SPOT system based on the MCM has been explored. From the simulation results, this new design provides slightly higher gains, but presents a larger size.

Future work. The next stage of the project involves the optimisation on the other plane of the concentrator. By including this factor, it should be possible to create a different concentrator capable of achieving higher gain.

Acknowledgment. The authors would like to acknowledge the collaboration of SolarEmpower Ltd. for its contribution to this project.

References

1. Renewable Energy Policy Network for the 21st Century (REN21 (2009), `http://www.ren21.net/pdf/RE_GSR_2009_Update.pdf`
2. Eiffert, P., Kiss, J.G.: Building-Integrated Photovoltaic Designs for Commercial and Institutional Structures: A Sourcebook for Architects. US Department of Energy's (2000)
3. SolarEmpower Ltd, `http://www.solarempower.com/`
4. DESOL Project, `http://cric-projects.com/front/index.jsp?idProject=10`
5. 3B Nexus, `http://www.3bnexus.com/`
6. Pre-X Capital Management Ltd, `http://www.pre-x.com/`
7. Ning, X., Winston, R., O'Gallagher, J.: Dielectric Totally Internally Reflecting Concentrator. Applied Optics 26(2), 300–305 (1987)
8. Ramirez-Iniguez, R., Idrus, S.M., Sun, Z.: Optical Wireless Communications: IR for Wireless Connectivity. Aurbach Publication, Florida (2007)
9. Ramirez-Iniguez, R., Green, R.J.: Optical Antenna Design for Indoor Optical Wireless Communication Systems. International Journal of Communication Systems 18, 229–245 (2005)
10. Ramirez-Iniguez, R.: Optical Antenna Design Software (OADES), Manual for OADES Version 1.0 (2007)

Author Index