Proceedings of the Eighteenth International Conference on Distributed Multimedia Systems

PROCEEDINGS

DMS 2012

The 18th International Conference on Distributed Multimedia Systems

Sponsored by

Knowledge Systems Institute Graduate School, USA

Technical Program

August 9-11, 2012
Eden Roc Renaissance Miami Beach, USA

Organized by

Knowledge Systems Institute Graduate School

Copyright © 2012 by Knowledge Systems Institute Graduate School

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written consent of the publisher.

ISBN 1-891706-32-2 (Paper)

Additional Copies can be ordered from:
Knowledge Systems Institute Graduate School
3420 Main Street
Skokie, IL 60076, USA
Tel:+1-847-679-3135
Fax:+1-847-679-3166
Email:off ce@ksi.edu
http://www.ksi.edu

Proceedings preparation, editing and printing are sponsored by
Knowledge Systems Institute Graduate School

Printed by Knowledge Systems Institute Graduate School

DMS 2012 Foreword

Welcome to DMS 2012, the 18th International Conference on Distributed Multimedia Systems.

With today's proliferation of multimedia data (e.g., images, animations, video, and sound), comes the challenge of using such information to facilitate data analysis, modeling, presentation, interaction, and programming. This is particularly important for the ever increasing number of end-users who are domain experts, but not IT professionals, allowing productive activities to cope with economic recession.

The conference is organized into eleven sessions on a variety of specialized themes, including: Slow Intelligence Systems, Mobile Intelligent Application, Human-Computer Interaction, Visual Languages and Computing, Distance Education Technology, Semantic Computing, Cultural and Social Multimedia, and Software Engineering and Software Security. The selection of papers to be presented at the conference was based upon a rigorous review process, with an acceptance rate of 42% of the submissions received in the category of full research papers. The conference program also includes short papers that report on ongoing research activities and applications.

In addition, we are very pleased to have Dr Borko Furht (Florida Atlantic University, USA), a distinguished researcher in the area of multimedia, delivering an invited keynote speech on industry multimedia projects.

The DMS Conference continues to be an internationally diverse research gathering. This year we are expecting authors and guests from 17 countries: Brazil, Canada, China, Egypt, France, Germany, India, Italy, Japan, Pakistan, Portugal, Saudi Arabia, Taiwan, Thailand, Korea, UK, and the United States.

As program co-chairs, we sincerely express our gratitude to the dedicated program committee members and conference support staff who have contributed to making DMS 2012 a success. We hope that you find this year's conference to be an invigorating exchange of research ideas, and that you include some time to enjoy the beautiful sights, sounds, and tastes of Miami.

Giorgio Valle, Shi-Hong Huang and Haiping Xu
DMS 2012 Program Co-Chairs

The 18th International Conference on Distributed Multimedia Systems (DMS 2012)

August 9-11, 2012
Eden Roc Renaissance Miami Beach, USA

Conference Organization

Steering Committee Chair

Shi-Kuo Chang, *University of Pittsburgh, USA*

Conference Co-Chairs

Prof. G. Costagliola, *University of Salerno, Italy*
Prof. S. Sitharama Iyengar, *Florida Int'l Univ., USA*
Prof. Jennifer Leopold, *Missouri University of Science & Technology, USA*

Program Co-Chairs

Prof. Giorgio Valle, *Univ. of Milan, Italy*
Prof. ShiHong Huang, *Florida Atlantic University, USA*
Prof. Haiping Xu, *University of Massachusetts Dartmouth, USA*

Program Committee

Junia Coutinho Anacleto, *Federal University of Sao Carlos, Brazil*
Carolina Lagos Aguirre, *University of Santiago, Chile*
Arvind K. Bansal, *Kent State University, USA*
Tolga Capin, *Bilkent University, Turkey*
Augusto Celentano, *Universita Ca Foscari di Venezia, Italy*
Ing-Ray Chen, *Virginia Tech (VPI&SU), USA*
Shu-Ching Chen, *Florida International University, USA*
William Cheng-Chung Chu, *Tunghai University, Taiwan*
Gennaro Costagliola, *Univ of Salerno, Italy*
Alfredo Cuzzocrea, *ICAR-CNR and University of Calabria, Italy*
Andrea De Lucia, *Univ. of Salerno, Italy*
Sergio Di Martino, *Universita degli Studi di Napoli Federico II, Italy*

Tiansi Dong, *FernUniversity-Hagen, Germany*
David H. C. Du, *Univ. of Minnesota, USA*
Larbi Esmahi, *National Research Council of Canada, Canada*
Daniela Fogli, *Universita degli Studi di Brescia, Italy*
Farshad Fotouhi, *Wayne State University, USA*
Kaori Fujinami, *Tokyo University of Agriculture and Technology, Japan*
Ombretta Gaggi, *Univ. of Padova, Italy*
Nikolaos Gkalelis, *Informatics & Telematic Institute, Greece*
Angela Guercio, *Kent State University, USA*
Niklas Hallberg, *FOI, Sweden*
Erland Jungert, *Linkoping University, Sweden*
Yau-Hwang Kuo, *National Cheng Kung University, Taiwan*
Fuhua Lin, *Athabasca University, Canada*
Alan Liu, *National Chung Cheng Univeristy, Taiwan*
Jonathan Liu, *University of Florida, USA*
Paolo Nesi, *University of Florence, Italy*
Kia Ng, *University of Leeds, UK*
Sethuraman Panchanathan, *Arizona State Univ., USA*
Antonio Piccinno, *Univ. of Bari, Italy*
Fabio Pittarello, *University of Venice, Italy*
Giuseppe Polese, *University of Salerno, Italy*
Buntarou Shizuki, *University of Tsukuba, Japan*
Peter Stanchev, *Kettering University, USA*
Mahbubur Syed, *Minnesota State University, USA*
Genny Tortora, *University of Salerno, Italy*
Ellen Walker, *Hiram College, USA*
Atsuo Yoshitaka, *JAIST, Japan*
Ing Tomas Zeman, *Czech Technical University, Czech Republic*
Kang Zhang, *University of Texas at Dallas, USA*
Roberto Zicari, *Goethe University Frankfurt, Germany*

Publicity Chair

Kao-Shing Hwang, *National Chung Cheng University, Taiwan*

Proceedings Cover Design

Gabriel Smith, *Knowledge Systems Institute Graduate School, USA*

Conference Secretariat

Judy Pan, *Chair, Knowledge Systems Institute Graduate School, USA*
Noorjhan Ali, *Knowledge Systems Institute Graduate School, USA*
Dennis Chi, *Knowledge Systems Institute Graduate School, USA*
David Huang, *Knowledge Systems Institute Graduate School, USA*
Alice Wang, *Knowledge Systems Institute Graduate School, USA*

International Workshop on Distance Education Technologies (DET 2012)

August 9-11, 2012
Eden Roc Renaissance Miami Beach, USA

Conference Organization

Steering Committee Chair

Shi-Kuo Chang, *University of Pittsburgh, USA*

Workshop Co-Chairs

Prof. Oscar Fuhua Lin, *Athabasca University, Canada*
Prof. Teresa Roselli, *Univ. of Bari, Italy*
Prof. Veronica Rossano, *Univ. of Bari, Italy*

Program Co-Chairs

Prof. Qun Jin, *Univ. of Waseda, Japan*
Prof. Luigi Collazzo, *Univ. of Trento, Italy*
Dr. Nicola Villa, *Univ. of Trento, Italy*
Dr. Pierpaolo Di Bitonto, *Univ. of Bari, Italy*

Program Committee

Pasquina Campanella, *University of Bari, Italy*
Maiga Chang, *Athabasca University, Canada*
Yuan-Sun Chu, *National Chung Cheng University, Taiwan*
Pierpaolo Di Bitonto, *University of Bari, Italy*
Larbi Esmahi, *Athabasca University, Canada*
Rita Francese, *University of Salerno, Italy*
Robert Heller, *Athabasca University, Canada*
Pedro Isaias, *Open University, Portugal*
Maria Laterza, *University of Bari, Italy*

Lian Li, *HeFei University of Technology, China*
Hong Lin, *University of Houston-Downtown, USA*
Paolo Maresca, *University Federico II, Napoli, Italy*
Andrea Molinari, *University of Trento, Trento, Italy*
Ignazio Passero, *University of Salerno, Italy*
Michele Risi, *University of Salerno, Italy*
Veronica Rossano, *University of Bari, Italy*
Giuseppe Scanniello, *University of Salerno, Italy*
Mahbub Syed, *Minnesota State University, USA*
Kazuo Yana, *Hosei University, Japan*
Chengjiu Yin, *Kyushu University, Japan*

International Workshop on Visual Languages and Computing (VLC 2012)

August 9-11, 2012
Eden Roc Renaissance Miami Beach, USA

Conference Organization

Steering Committee Chair

Shi-Kuo Chang, *University of Pittsburgh, USA*

Workshop Chair

Levent Burak Kara, *Carnegie Mellon University, USA*

Program Chair

A. J. Delaney, *University of Brighton, UK*
Metin Sezgin, *Koc University, Turkey*

Program Committee(tentative)

Danilo Avola, *University of Rome, Italy*
Dorothea Blostein, *Queen's University, Canada*
Jeff Browne, *UC Santa Barbara, USA*
Alfonso F. Cardenas, *University of California, USA*
Gennaro Costagliola, *University of Salerno, Italy*
Philip Cox, *Dalhousie University, Canada*
Sergiu Dascalu, *University of Nevada, USA*
Aidan Delaney, *University of Brighton, UK*
Vincenzo Deufemia, *University of Salerno, Italy*
Filomena Ferrucci, *University of Salerno, Italy*
Manuel J. Fonseca, *INESC-ID, Portugal*
Vittorio Fuccella, *University of Salerno, Italy*
Joaquim A Jorge, *Instituto Suoerior Tecnico, Portugal*

Levent Burak Kara, *Carnegie Mellon University, USA*
Robert Laurini, *University of Lyon, France*
Ian Oliver, *Nokia, Finland*
Joseph J. Pfeiffer, *New Mexico State University, USA*
Beryl Plimmer, *University of Auckland, New Zealand*
Tevf k Metin Sezgin, *Koc University, Turkey*
Gem Stapleton, *University of Brighton, UK*
Stefano Valtolina, *University of Milan, Italy*
Giuliana Vitiello, *University of Salerno, Italy*

Table of Contents

Foreword ..... **iii**

Conference Organization ..... **iv**

Keynote

Industry Multimedia Projects
Borko Furht ..... **xv**

Slow Intelligence Systems

Relating Slow Intelligence Research to Bilingualism
Tiansi Dong and Ingo Glöckner **3**

Design and Implementation of Image Analysis System by Applying Component-based Slow Intelligence System Framework
Liqun Kuang, Shi-Kuo Chang, Yingze Wang and Yao Sun **9**

A Framework for Intrusion Detection System based on the Slow Intelligent Approach
Francesco Colace, Massimo De Santo, Salvatore Ferrandino and S.K. Chang **17**

Cloud Services for Medical Image Processing: Application of Service Science and Slow Intelligence Systems (S)
Kang-Lin Peng, Shi-Kuo Chang, Yu-Chee Tseng and Thomas M. Deserno **25**

A Slow Intelligence Framework for Human Behavior Recognition Using Wearable Cameras (S)
Wen-Hui Chen and Shi-Kuo Chang **29**

Mobile Intelligent Application

Development of A Real-Time Air Pollution Alert System Using Smart Phones (S)
Nuri Yilmazer, Young Lee, Juan Mora, William Webb and Kuo-Jen Liao **39**

Human-Computer Interaction

Sharing Digital and Physical Contents on and above Distributed Surfaces using Active Infrared Keying
Buntarou Shizuki, Yuichi Murata and Jiro Tanaka **42**

Fuzzy Color Space Segmentation to Identify the Same Dominant Colors as Users
José C. Amante and Manuel J. Fonseca **48**

Insights on the development of visual tools for analysis of pollution data
Paolo Buono and Maria Francesca Costabile **54**

Semantic Computing and Processing

Ontological Filtering for Sentiment Analysis
Francesco Colace, Massimo De Santo, Paolo Napoletano, C. Becchi and S.K. Chang **60**

Eff cient Computation of Object Boundary Intersection and Error Tolerance in VRCC-3D+ (S)
Nathan Eloe, Jennifer Leopold, Chaman Sabharwal and Zhaozheng Yin **67**

Determining Minimal Transitions Between VRCC-3D+ Relations (S)
Adaleigh Martin, Ian Kottman, Robyn Littleton, Jennifer Leopold, Chaman Sabharwal and Nathan Eloe **71**

Cultural and Social Multimedia

EEG-Based BCI Data Analysis on Visual-Perceptual Priming in the Context of a Museum of Fine (S)
Annalisa Banzi and Raffaella Folgieri **75**

Use of large multi-touch screens for informal learning (S)
Rosa Lanzilotti and Maria Francesca Costabile **79**

Software Engineering and Software Security

Software Product Line Evaluation: Categorization and Evolution over the Years
Edson A. Oliveira Junior, Itana M. S. Gimenes and José C. Maldonado **83**

Variability Management in Software Product Line Activity Diagrams
Diego R. Fiori, Itana M. S. Gimenes, José C. Maldonado and Edson A. Oliveira Junior **89**

Improving the identif cation of traceability links between source code and requirements
Daniel Hanspeter, Andrea Janes, Alberto Sillitti and Giancarlo Succi **95**

Intelligent Multimedia Computing

An Integrated Architecture for Multiagent Virtual Worlds for Performing Adaptive Testing Games
Sandeep Virwaney, Fuhua Lin, Grant McClure and Robert Heller **101**

Optimimal resources utilization for indexing within a distributed multimedia retrieval system: an implementation for a video surveillance use case
Dana Codreanu, Ana-Maria Manzat and Florence Sedes **107**

Functionalities and Flow Analyses of Knowledge Oriented Web Portals
Daniele Cenni, Paolo Nesi and Michela Paolucci **113**

Artif cial Intelligence Approaches to Software Engineering

Enhanced Fuzzy Multiple Regression Neural Model for Software Effort Estimation (S)
Khaled Mohamed Shams, Dr. Amr Kamel **119**

Automating UML Sequence Diagram Generation by Treating it as a Planning Problem
Yaser Sulaiman and Moataz Ahmed **124**

Knowledge Management to Support the Use of Agile Methodologies (S)
Jordão E. and Furtado F. **130**

DET Workshop

Evolution of Learning Environments

An educational game to learn type 1 diabetes management
Pierpaolo Di Bitonto, Teresa Roselli, Veronica Rossano, Elda Frezza and Elvira Piccinno.... **139**

Building Wider Team Cooperation Projects from Lessons Learned in Open Communities of Practice
Paolo Maresca, Angela Guercio and Lidia Stanganelli **144**

e-Learning Applications: medicine & mobility

Information systems and e-Learning platforms: from loose coupling to deep integration
Luigi Colazzo and Andrea Molinari **150**

Revisiting design of learning and mobility services
Bruno Apolloni, Raffaella Folgieri and Giorgio Valle **155**

Blended learning in Continuing Medical Education (S)
Francesco Epifania **161**

VLC Workshop

Visual Languages and Computing

DEViL3D – A Generator Framework for Three-Dimensional Visual Languages
Jan Wolter **171**

Semi-Automated Generation of Domain-Specif c, Natural Language-Based, Visual Programming Languages (S)
Michael Wisely and Jennifer L. Leopold **177**

Formalizing the Syntax of Codecharts (S)
Jon Nicholson, Aidan Delaney and Gem Stapleton **181**

Does the Orientation of an Euler Diagram Affect User Comprehension?
Andrew Blake, Gem Stapleton, Peter Rodgers, Liz Cheek and John Howse **185**

A Video Game to Learn KeyScretch (S)
Gennaro Costagliola, Mattia De Rosa, Vittorio Fuccella and Fabrizio Torre **191**

Poster/Demo: Multimedia Arts & DMS/VLC Demos

Semi-automatic requirement tracing in modif ed code: An Eclipse Plugin (P)
Daniel Hanspeter, Andrea Janes, Alberto Sillitti and Giancarlo Succi **A-3**

An Interactive Learning and Assessment System for Simulation-based Science Education Using Cloud Computing Technology (P)
Mengxia (Michelle) Zhu and Frackson Mumba **A-5**

Reviewer's Index **A-6**

Author's Index **A-8**

Poster/Demo Presenter's Index **A-10**

Note: (S) indicates a short paper.
(P) indiecates a poster or demo, which is not a refereed paper.

Keynote
Industry Multimedia Projects

Borko Furht
Florida Atlantic University, Florida

Abstract

In this talk we will first introduce the NSF-sponsored Industry/University Cooperative Center for Advanced Knowledge Enablement at FAU, which presently has 16 industry members with about $1.5 million memberships. The Center is successfully building a bridge linking academia, industry, and government in a coordinated research initiative. We describe several applied multimedia projects conducted within the Center including video and image mining for coastline security, 3D image reconstruction and segmentation of brain cells, augmented reality methods for hearing augmentation, automatic asset tracking in datacenter, and a few others. All these projects are initiated by industry partners who are the members of the Center and who are interested to use the obtained research results and create successful commercial products. The talk will complete with our prediction where the multimedia computing is heading to the next several years.

About the Speaker

Borko Furht is a professor and chairman of the Department of Computer & Electrical Engineering and Computer Science at Florida Atlantic University (FAU) in Boca Raton, Florida. He is also Director of the NSF-sponsored Industry/University Cooperative Research Center (I/UCRC) on Advanced Knowledge Enablement. Before joining FAU, he was a vice president of research and a senior director of development at Modcomp (Ft. Lauderdale), a computer company of Daimler Benz, Germany; a professor at University of Miami in Coral Gables, Florida; and a senior researcher in the Institute Boris Kidric-Vinca, Yugoslavia. Professor Furht received a Ph.D. degree in electrical and computer engineering from the University of Belgrade.

His current research is in multimedia systems, video coding and compression, 3D video and image systems, wireless multimedia, and Internet and cloud computing. He is presently Principal Investigator and Co-PI of several multiyear, multimillion-dollar projects, including NSF PIRE project and NSF High-Performance Computing Center. He is the author of numerous books and articles in the areas of multimedia, computer architecture, real-time computing, and operating systems. He is also editor of two encyclopedias – Encyclopedia of Wireless and Mobile Communications, CRC Press, 2007, 2012 (2nd edition), and Encyclopedia of Multimedia (Springer, 2009).

He is a founder and editor-in-chief of the Journal of Multimedia Tools and Applications (Springer). He has received several technical and publishing awards, and has consulted for many high-tech companies including IBM, Hewlett-Packard, Xerox, General Electric, JPL, NASA, Honeywell, and RCA. He has also served as a consultant to various colleges and universities. He has given many invited talks, keynote lectures, seminars, and tutorials. He served on the Board of Directors of several high-tech companies.

Proceedings
International Conference on
Distributed Multimedia Systems
(DMS 2012)

Relating Slow Intelligence Research to Bilingualism

Tiansi Dong and Ingo Glöckner
Department of Mathematics and Informatics
FernUniversiät in Hagen
Universitätstr. 1, 58084 Hagen, Germany
Email: {tiansi.dong|ingo.gloeckner}@fernuni-hagen.de

***Abstract*—Late-learned-L2 Bilinguals are people who learn the second language (L2) after the acquisition of the mother language (L1). We relate Slow Intelligent Systems (SIS) research to bilingualism, in particular late-learned-L2 bilinguals. The separation of and the relation between the quick and the slow decision cycles are supported by the linguistic models of the late-learned-L2 bilinguals and by the human memory structure. The competition relation is proposed as a new relation between quick/slow decision cycle based on the research on bilingualism. The cognitive mechanism of bilinguals extends the application area of slow intelligent systems into machine translation. Within the SIS framework, translation starts with a slow decision cycle whose critical process is cross-linguistic cue switching. We argue that an efficient way to evaluate this new translation method is to carry out German into Chinese translation in education and in computational simulation. We review three main machine translation technologies in AI, and show the advantages of the *slow intelligent* translation.**

I. Introduction

Almost all of the intelligence systems in nature, even roses, have *slow intelligence* [11]. Following the "minimalist" strategy in cognitive science that *cognitive processes needed by other areas of cognitive functioning can also be involved in language processing* [38], we explore *slow intelligence* in the domain of second language acquisition. This is a slow process for adults to acquire a second language – They will start with memorizing basic words and expressions, and apply grammar rules. The performance of using the second language is slowly improved through practicing, and can reach to the level of the mother tongue. It is a slow intelligence system whose performance improves gradually over time, whose slow decision cycle supports the acquisition of a new language, and whose quick decision cycle supports mother-tongue-level language processing.

The aim of this paper is to show that the second language acquisition of adults is a joint topic for research in slow intelligence, linguistics (bilingualism), and AI. Research in psycholinguistics, neurolinguistics, and aphasia not only supports the general framework of slow intelligent systems, but also provides data to identify new features. Research on slow intelligent systems propose novel approaches to machine translation in AI.

The rest of the paper is structured as follow: Section 2 relates SIS in bilingualism research by associating the SIS framework with the hierarchical model of late-learned-L2 bilinguals, and with memory structures of mind. The separation of two different decision cycles in slow intelligence systems is consistent with the two branches in the revised hierarchical model of late-learned-L2 bilinguals [35], and also consistent with declarative/procedural memory structure of mind, e.g. [53], [54], [55]. Section 3 reviews some brain-impaired bilinguals and proposes a new general feature of SIS that make the quick/slow decision cycles competitive. Section 4 proposes that the critical process of translation is cross-linguistic cue switching, which is a cognitive processing independent of understanding and production within a single language. Section 5 shows the envisioned German into Chinese translation is an efficient way to evaluate this novel translation method. Section 6 critically reviews three main approaches to machine translation in AI and demonstrates the advantage of the slow intelligence approach.

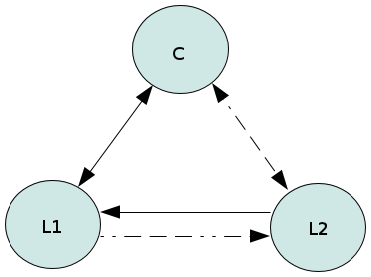


Fig. 1. The *Revised Hierarchical Model* by Kroll and Stewart [35] (1994). 'C' is the conceptual representation; 'L1' is the lexical representation of the mother language; 'L2' is the lexical representation of the second language

II. Anchor SIS research in Bilingualism

Slow Intelligent Systems (SIS), as a general framework for intelligent systems, shall be applicable for the second language acquisition of adults who fluently speak their mother tongue and whose performance of the second language improves gradually. Do they have two separate language systems in mind, each supports one language, or one system for both languages? How shall we use the SIS framework to simulate bilingualism?

A. The Aggregated Language Model

Sufficient research in psycholinguistics supports an aggregated language model, that is, later learned languages are superimposed on the earlier learned language. This idea dates back to Freud [20] (1891). For example, researchers found

non-selective access to lexemes in both languages, e.g., [3], [16], [15]. That is, lexemes in different language are simultaneously activated. This explains why many bilinguals often mix lexemes in different languages. The aggregation can reach to such a degree that some Spanish-English bilinguals may pronounce neither in Spanish nor in English, as observed by Walters [61] (2005).

A general agreement in psycholinguistics is on the separation between meanings and forms in language processing. Meanings (concepts) are shared by both languages, e.g., [10], [31], [13], while forms are specific to each language, e.g., [31], [51], [21]. A hierarchical model of meanings and forms for lexeme acquisition is illustrated in Figure 1. For non-proficient late-learned L2 speakers, there will be no direct link between L2 and C, and for bilinguals who learned two languages from birth, there will be no direct link between L1 and L2, i.e., [45], [35]. From this, we can infer how proficient bilinguals carry out translation: firstly, meaning-acquisition from the source language, secondly, language-production from the acquired meaning into the target language.

B. Linguistic Mental Storage of late-learned-L2 Bilinguals

With regards to the mental storage of linguistic knowledge, neurolinguists conducted a serial of experiments and found that lexemes of both languages of late-learned-L2 bilinguals are macroscopically represented in the same cerebral area e.g., [32], [12], [19]. This area is associated with the declarative memory system, whose main function is for knowledge representation and processing of facts and events, e.g. [55].

Native speakers use their mother tongues with high proficiency, but may fail to explain the grammar rules. The reason lies in the fact that mother tongues are acquired as skills through practice, instead of as knowledge in the form of facts and rules. Cognitive skills are managed by the procedural memory system, e.g., [1], [52], [63], which can also support the syntactical knowledge of the mother tongue, e.g. [53], [54].

The declarative memory system has a rapid learning ability and can acquire new knowledge fast. Knowledge acquisition of the procedure memory system is carried out gradually, e.g., [42], [44]. Late-learned-L2 bilinguals may have more difficulty than early-learned-L2 bilinguals to reach the same level of proficiency in using L2. The reason lies in the fact that late-learned-L2 bilinguals learn grammars of the second language as rules and relations, which are stored in their declarative memory systems, e.g., [55]. They can master their second language at a high-level of proficiency, and result in a neurocognitive pattern similar to that of mother tongue speakers, i.e., [55, p.153], for the performance of a skill can be gradually improved through practicing.

C. Slow Intelligent Framework of Late-Learned-L2 Bilinguals

From the neurolinguistic research on bilingualism, we see two different ways in language acquisition: one is to view language as a cognitive skill, and learn languages purely through practicing; the other is to view language as a knowledge system, and learn languages through remembering grammatical rules and lexemes. The first way associates with the activities of the procedural memory system. Language acquisition in this way is slow, but can reach to a high efficiency in using languages. The second way associates with the activities of the declarative memory system. Language acquisition in this way is much faster, but results in inefficient usage of languages.

To model the second language acquisition of late-learned-L2 bilinguals, the slow cycle models the activities of the declarative memory system: learning lexical knowledge and grammatical rules. This cycle is quick to learn language knowledge, but may not use language in high-level of proficiency. Continuous practice gradually nurtures the quick cycle, which models activities of the procedural memory system.

The slow cycle for language acquisition of late-learned L2 speakers consists of two layers: the meaning layer, including basic concepts and relations among basic concepts; the form layer, including lexemes and syntactical rules. Language comprehension is modeled by the process of acquiring meaning representation from forms; language production is modeled by the process of developing forms from meaning representations. Translation is modeled by the sequence of comprehension of the source language, and production in the target language.

One advantage of relating SIS research to bilingualism research is that the former will be benefited from the fruitful research results from bilingualism, as shown below.

III. Competitive relation between quick/slow decision cycles

The patient, E.M., in [2] spoke Venetan as the mother language (L1), and learned Italian as her second language (L2) for three years in an elementary school with poor performance. After a stroke, she was not able to speak her mother language, but could proficiently speak her second language. Translation from L1 into L2 was more accurate and quick than vice versa; sentence translation from L2 into L1 was hardly possible.

The patient, A.D, in [43] spoke French as L1, and learned Arabic (L2) at a French school at the age of 10. Her both languages were very fluent. After a traffic accident, she lost consciousness for 15 minutes. Four days later, she could only speak Arabic. Her French was slowly recovered. On the 19th day after the accident, translation from L2 into L1 was good; and translation from L1 into L2 was a total failure. On the next day, translation from L1 into L2 was excellent, while translation from L2 into L1 was poor. On the 28th day after the accident, translation from L1 into L2 was very poor, while translation from L2 into L1 was very good. On the 41st day after the accident, translation in each direction was poor. On the next day, translation from L1 into L2 was poor, while translation from L2 into L1 was good.

Both E.M. and A.D. are later-learned L2 bilinguals. Processing of L1 is supported by quick decision cycle, while processing of L2 is supported by slow intelligence cycle. Their recovery pattern suggested a new feature to the slow intelligent framework: the competitive relation between quick/slow decision cycles. That is, the dysfunction of cycle results in enhanced performance of the other cycle.

This new feature is supported by studies on the memory system of animal and human, e.g., [41], [44], [54]. For example, women are better than men in verbal memory tasks [34], because verbal memory tasks depend on the declarative memory system (slow decision cycle) [59] and estrogen betters the performances of the declarative memory system [54]. Men show better performance in skills, e.g. aimed throwing, mental rotation [30], which depend on the procedural memory system (quick decision cycle) [56]. Performances in these tasks are bettered with the decreased estrogen, and worse with the increased estrogen [23], [30].

IV. Translation as cross-linguistic cue switching

Research on bilingualism enhances the understanding of the nature of second language acquisition, and provides a new application area for SIS engineering: machine translation.

A. Linguistic Cues

A cue is a piece of information that a speaker or listener can use to determine the relationship between meanings and forms [37, pp. 169]. Cross-linguistic study showed that each language uses a particular set of cues: Italian extremely relies on agreement cues, German relies on both agreement cues and animacy cues, English relies overwhelmingly on word order [39], Chinese relies on cues of passive marker 被/by, animacy, word order, object marker 把/hold, indefinite marker 一/one [37].

B. Translation as the Cross-linguistic Cue Switching

Translation can be understood as a cognitive functioning of an aggregated language model, whose central task is the cross-linguistic cue switching. When a sentence in language A is received by the model, cues of language A are used to acquire the meaning of the sentence; then the model generates cues of language B from the acquired meaning to produce sentence in language B.

Some researchers view this cross-linguistic cue switching as a pragmatic control [46], [27], or as inhibitory control [22]. Such cue-switching may not be easy for late-learned second language bilinguals. They may use the cues of their mother tongue to process L2 sentences [28]. Even proficient second language speakers might be insensitive to cues which do not exist in their mother language [29].

Bilinguals with brain lesion may understand both languages, and fail to translate from one into the other language [2], [43]. Cross-linguistic cue switching is an linguistic function independent of functions of a single language model and is the central activity of translation.

Cross-linguistic cue switching may not be trivial. The Dani people, living in the central highlands of western New Guinea, only have two terms, “mili” (roughly means “dark”) and “mola” (roughly means “light”), for colors [47]. Translation between Dani color terms and English color terms would not be easy. The Dani color terms are only determined by the brightness of a color; while English color terms are determined by hues, brightness, and saturation. Translating color terms requires switching of granularity (zoom-in or zoom-out) and perspective (e.g. shrink, broaden). The Guugu Yimithirr people, living in North Queensland, Australia, only use absolute orientations, e.g., “I left it on the southern edge of the western table in your house” [36]. English speakers would use relative orientations in terms of “left”, “right”, “front”, and “back”. Translation of spatial descriptions in the two languages demands a total perspective switching in orientation reference frameworks. The problem is that preferred perspective and the granularity differs, as noted in Rosch et al.[48, p. 430](1976), and sometimes even incompatible without extra information. The precondition for cross-linguistic cue switching is that their semantic representations must be made compatible.

V. Efficient Evaluation through German into Chinese Translation

The SIS translation method shall be applicable to any two languages. To evaluate this method efficiently, we need to choose two languages with different cues. The more different, the better.

Chinese is notorious for its lack of inflection informations. In this aspect, Chinese is different from almost all other western languages. Among all western languages, which one shall we choose? We can choose German, for the reason that it has a very systematical case based inflection system. Some western languages, however, even have more systematical inflection systems than German. For example, the Turkish inflection system is much more regular than German, and “is perhaps the most regular, semantically transparent, and unambiguous system in the world” [5, pp. 173].

Due to some technical and social considerations, we prefer German to Turkey. Technically, we shall focus on the central topic, i.e. cross-linguistic cue switching. Ideally, the meaning representation of L1 has been intensively researched and there are computational tools for automatic acquisition. This happens to the German language, i.e., [25], [26]. Socially, the more L1 people interested in L2 learning, the better. There are much more German people interested in learning Chinese and traveling in China, than Turkey people.

Why not choose English, instead of German, for there are definitely more English people interested in Chinese than German people? The answer is: Because the English language has the poorest inflection system among all of Indo-European and Ural-Altaic languages, including Dutch, English, Italian, German, Hebrew, Hungarian, Italian, Japanese, Turkish. In English, the word-order plays the dominant role in conveying meanings, though it has some inflectional devices, such as number, person, and gender of nouns, and agreement relations between nouns and verbs. This isolates English from all other Indo-European and Ural-Altaic languages and demonstrates some similarity to Chinese, [6], [37].

The cue-difference between German and Chinese results in tremendous difficulties for native German speakers to construct Chinese sentences. For the phrase *im Alter von 82 Jahren*, a

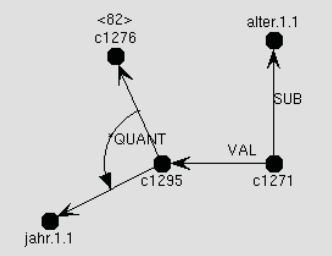


Fig. 2. The meaning representation of the phrase 'im Alter von 82 Jahren'

German student translated it into *歲82年*[1], which is meaningless in Chinese. Though she knew Chinese translations of single German words, she did not know the order to construct them into a meaningful Chinese phrase. What has not been fully researched in second language education is a systematic cue-switching rules from German into Chinese, which can be represented in the three types of schemas as follows: the meaning schema, the single translation schema, and the Chinese production schema. The meaning schema is the semantic representation of the input. The meaning schema for the above example is illustrated in Figure 2. The semantic representation of German has been intensively researched [25], and can be automatically acquired by the WOCADI parser [26], [24]. The single translation schema lists possible Chinese translations of all leaf nodes of the meaning schema. The single translation schema of the German student can be listed as follows.

$$(1) \quad \texttt{jahr.1.1} \rightarrow 年$$
$$(2) \quad \texttt{alter.1.1} \rightarrow 歲$$

The Chinese production schema lists rules, which transform relations in the meaning schema into meaningful Chinese phrases, based on cues of the Chinese language. The production schema of the above example can be listed as follows.

$$(3) \quad *\texttt{QUANT}(\texttt{X},\texttt{Y}) \rightarrow [\texttt{X}] \triangleright [\texttt{Y}]$$
$$(4) \quad \texttt{c}[\texttt{SUB}(\texttt{alter.1.1}), \texttt{VAL}(\texttt{X})] \rightarrow [\texttt{X}] \triangleright [\texttt{alter.1.1}]$$
$$(5) \quad [\texttt{X}] \triangleright [\texttt{Y}], [\texttt{Y}] \triangleright [\texttt{Z}] \rightarrow [\texttt{X}] \triangleright [\texttt{Z}]$$

'[X]' is the single Chinese translation of 'X'. Rule (3) tells the word order in Chinese and the logical relation: `[X]` is restricted to domain `[Y]`: Following Russell [49] (1919), we can say that *'82' refers to any domain which has 82 elements*. Rule (4) is understood similarly. Rule (5) tells that $\triangleright$ is transitive. These rules fullfill the cue-switching from German into Chinese. The German student could translate *im Alter von 82 Jahren* into *82* $\triangleright$*年*$\triangleright$*歲*, and into *82*$\triangleright$*歲* with Rule (5). The Chinese translation *82歲* will be obtained by suppressing (removing) all logical relations, here only one $\triangleright$.

VI. Contribution to machine translation in AI

The computational simulation of human translation is normally called "Machine translation (MT)" in AI. Rule-based Machine Translation, Example-based Machine Translation, and Statistical Machine Translation are the three main technologies. Is there anything new of SIS translation to machine translation in AI?

A. Statistical Machine Translation

The Statistical Machine Translation (SMT) is the most popular and the most widely-studied method. The idea dates back to Weaver[62](1949), and was re-introduced in early 90s [9]. The central idea can be stated as follows: Given a sentence f in language A, we look for the sentence e in language B which maximizes $\mathbf{P}(e \mid f)$, written as $argmax_e \mathbf{P}(e \mid f)$. That is, the sentence e in language A is *most likely* the translation of the sentence f in language B. The term '*most likely*' can be approached by the Bayes rule: $argmax_e \mathbf{P}(e \mid f) = argmax_e \mathbf{P}(e) * \mathbf{P}(f \mid e)$. $\mathbf{P}(e)$ is the probability that e occurs in language A. $\mathbf{P}(f \mid e)$ is the probability that f occurs as a sentence in language B, if e is a sentence in language A. $\mathbf{P}(e)$ is called the 'Language Model' (LM). $\mathbf{P}(f \mid e)$ is called the 'Translation Model'. The basic idea is that firstly all possible alignments between e and f are enumerated, and assigned to an equal probability; then the Estimation-Maximization (EM) bootstrapping algorithm, i.e., [7] is applied to update these probabilities for all possible alignments (including multiple (1 to many) or zero (1 to 0) alignments) again and again, till one alignment with a significantly higher probability appears. The statistical translation approach is therefore based on such a mind model that each language is a separate system, e.g. $\mathbf{P}(e)$, and there is a transition system between the language models, i.e. $\mathbf{P}(f \mid e)$. With this model, it would be hard to simulate the fact that bilinguals often produce utterances containing lexemes in different languages.

B. Example-based Machine Translation

The example-based approach, e.g. [40], [8], decomposes the input into several phrases with the assumption that each phrase can be directly translated into phrases in the target language by analogy. The output in the target language is composed using these translated fragments, which, as admitted by Nagao [40], is by no means trivial. Analogy is implemented as a case-based reasoning. This approach was targeted to translate languages with different structures, i.e., English and Japanese. This approach requires a huge sample bilingual corpus as examples, which is not always available. Another difficulty is the problem of ambiguity. As the main tenet of the approach is not to do deep linguistic analysis (i.e. meaning representation), ambiguity becomes a foreign concept to this approach.

The Example-based approach was inspired from a naive observation of the second language acquisition of human adults, and views translation as a process of analogy, as described in [40, pp.173]: "Let us reflect about the mechanism of human translation of elementary sentences at the beginning of foreign language learning. A student memorizes the elementary English sentences with the corresponding Japanese sentences. The first stage is completely a drill of memorizing lots of similar sentences and words in English, and the corresponding

[1]The meaning in English is: at the age of 82. This example is taken from a semester examination of the Chinese Department, Bonn University, Germany

Japanese. Here we have no translation theory at all to give to the student. He has to get the translation mechanism through his own instinct. He has to compare several different English sentences with the corresponding Japanese. He has to guess, make inferences about the structure of sentences from a lot of examples". By saying "memorizes the elementary English sentences with the corresponding Japanese sentences", we know that it is carried out in the declarative memory; by saying "a drill of memorizing lots of similar sentences and words", we know that it is carried out in the procedural memory. Therefore, this approach could be extended, taking benefit of concrete research findings in second language acquisition of human adults.

C. Rule-based Machine Translation

The classical approach to MT is the Rule-Based Machine Translation (RBMT). 'Rule' is a general term, including linguistic information acquired from dictionaries and grammars. The main principle of the RBMT approach is to find a structural link between the input in the source language and the output in the target language, while preserving the meaning. Normally we distinguish two kinds of RBMT systems: *Transfer* RBMT and *Interlingual* RBMT. The *Transfer* RBMT approach works better for two languages with a similar structure. The *Inter-lingual* RBMT has an intermediate representation independent of the source and the target languages, which somehow serves as the meaning of the source sentence. Is such intermediate meaning representation really adequate for all natural languages? One reason against such canonical meaning representation is that there is no canonical meaning representation for formal languages, i.e., [64]. Another reason is from neurolinguistics and psycholinguists: instead of one-for-all model, they advocate an aggregated model.

D. Comments

From the perspective of pragmatism, it is not so important whether a machine translation system simulates the cognitive activity of human translation, as long as the system can be reliably used in real applications. To improve the translation reliability, researchers have attempted to combine different approaches. For example, example-based translation approach with rule-based translation approach, e.g., [57], [50], rule-based translation approach with statistical translation approach, e.g., [14], adding formal domain knowledge into translation system, e.g., [18]. In one of the largest research project, Verbmobil, [60] researchers tried almost all possible hybrid method and also developed new methods within techniques. For example, using templates within example-based machine translation, i.e., [4], using *stattrans* module within statistical translation, i.e., [58]. However, Uszkoreit, Xu, and Liu observed that "although the successful project [Verbmobil] resulted in a strong push to face-to-face speech translation, the base technologies could not be lifted up to the level of coverage and robustness required for a product in the demanding tourist market, ... machine translation is still far from being reliable", [57, pp. 135-137].

Something might be still being neglected in the machine translation theories, as Knight [33, p.82] said, "If we pour in lots of theories from computer science, linguistics, statistics and AI–and still get wrong translations–then we know we need better theories". One lesson learned from neurolinguistics is that there is a specific neuro-functional system supporting translation, as remarked by Fabbro [19, pp. 219]. Based on research of bilingualism, we propose that this specific system is the process of cross-linguistic cue production, which is clearly not needed for single language comprehension and production. The second lesson learned from neuro- and psycho-linguistics is the aggregated, hierarchical, two-branched language model of bilinguals. This model is consistent with the separation of the slow/quick decision cycles of the slow intelligence framework. The third lesson learned from human memory structure is that late-learned-L2 bilinguals start learning L2 with declarative knowledge. This is consistent with properties of the slow/quick decision cycles. These three points are not taken into consideration in the three main machine translation technologies in AI, and can be well modeled in the proposed slow intelligence translation approach. A sample translation system within the SIS framework was described in [17].

REFERENCES

[1] J. W. Addridge and K. C. Berridge. Coding of serial order by neostriatal neurons: A "natural action" approach to movement sequence. *Journal of Neuroscience*, 18:2777–2787, 1998.

[2] S. Aglioti, A. Beltramello, F. Girardi, and F. Fabbro. Neurolinguistic and follow-up study of an unusual pattern of recovery from bilingual subcortical aphasia. *Brain*, 119:1551–1564, 1996.

[3] E. P. Altenberg and H. S. Cairns. The effect of phonosyntactics contraints on lexical processing in bilingual and monolingual subjects. *Journal of Verbal Learning and Verbal Behavior*, 22(2):174–188, 1983.

[4] M. Auerswald. Example-Based Machine Translation with Templates. In W. Wahlster, editor, *Verbmobil: Foundations of Speech-to-Speech Translation*, pages 418–427. Springer Verlag, 2000.

[5] E. Bates and B. MacWhinney. Competition, Variation, and Language Learning. In B. MacWhinney, editor, *Mechanisms of Language Acquisition*, pages 157–193. Lawrence Erlbaum, Hillsdale, NJ, 1987.

[6] E. Bates, B. MacWhinney, C. Caselli, A. Devescovi, F. Tatale, and V. Venza. A Cross-linguistic Study of the Development of Sentence Interpretation Strategies. *Child Development*, 55:341–354, 1984.

[7] L. E. Baum. An inequality and associated maximization technique in statistical estimation of probabilistic functions of a markov process. *Inequalities*, 3:1–8, 1972.

[8] H. U. Block. Example-based Incremental Synchronous Interpretation. In W. Wahlster, editor, *Verbmobil: Foundations of Speech-to-Speech Translation*, pages 411–417. Springer Verlag, 2000.

[9] P. F. Brown, J. Cocke, and S. A. Della Pietra. A Statistical Approach to Machine Translation. *Computational Linguistics*, 16(2):79–85, 1990.

[10] A. Caramazza and I. Brones. Semantic classification by bilinguals. *Canadian Journal of Psychology*, 34(1):77–81, 1980.

[11] S. K. Chang. A general framework for slow intelligence systems. *International Journal of Software Engineering and Knowledge Engineering*, 20(1):1–18, 2010.

[12] M. W. L. Chee, E. W. L. Tan, and T. Thiel. Mandarian and English single word processing studied with functional magnetic resonance imaging. *The Journal of Neuroscience*, 19(8):3050–3056, 1999.

[13] H-C. Chen. Lexical processing in a non-native language: Effects of language proficiency and learning strategy. *Memory and Cognition*, 18(3):279–288, 1990.

[14] Y. Chen, M. Jellinghaus, A. Eisele, Y. Zhang, S. Hunsicker, S. Theison, C. Federmann, and H. Uszkoreit. Combining multi-engine translations with Moses. In *Proceedings of the Fourth Workshop on Statistical Machine Translation*, StatMT '09, pages 42–46, Stroudsburg, PA, USA, 2009. Association for Computational Linguistics.

[15] À. Colomé. Lexical activation in bilinguals' speech production: language specific or language independent? *Journal of Memory and Language*, 45(4):721–736, 2001.

[16] A. Costa and A. Caramazza. Is lexical selection in bilinguals language-specific? further evidence from spanish-english and english-spanish bilinguals. *Bilingualism Language and Cognition*, 2(3):231–244, 1999.

[17] T. Dong and I. Glöckner. Word Expert Translation from German into Chinese in the Slow Intelligence Framework. In *The 17th International Conference on Distributed Multimedia Systems*, pages 16–21, Florence, Italy, 2011.

[18] M. C. Emele, M. Dorna, A. Lüdeling, H. Zinsmeister, and C. Rohrer. Semantic-Based Transfer. In W. Wahlster, editor, *Verbmobil: Foundations of Speech-to-Speech Translation*, pages 359–376. Springer Verlag, 2000.

[19] F. Fabbro. The Bilingual Brain: Cerebral Representation of Languages. *Brain and Language*, 79:211–222, 2001.

[20] S. Freud. *Zur Auffassung der Aphasien. Eine kritische Studie*. Franz, Leipzig und Wien, 1891. Translated by E. Stengel (1953). New York: International University Press.

[21] L. Gerard and D. L. Scarborough. Language-specific lexical access of homographs by bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 15(2):305–315, 1989.

[22] D. W. Green. Mental control of the bilingual lexico-semantic system. *Bilingualism Language and Cognition*, 1:67–81, 1998.

[23] E. Hampson. Variations in sex-related cognitive abilities across the menstrual cycle. *Brain and Cognition*, 14(1):26–43, 1990.

[24] S. Hartrumpf. *Hybrid Disambiguation in Natural Language Analysis*. Der Andere Verlag, Osnabrück, Germany, 2003.

[25] H. Helbig. *Knowledge Knowledge Representation and the Semantics of Natural Language*. Springer-Verlag, Berlin Heidelberg, 2006.

[26] H. Helbig and S. Hartrumpf. Word class functions for syntactic-semantic analysis. In *Proceedings of the 2nd International Conference on Recent Advances in Natural Language Processing*, pages 312–317. Tzigov Chark, Bulgaria, 1997.

[27] A. E. Hernandez, A. Martinez, and K. Kohnert. In search of the language switch: An fMRI study of picture naming in Spanish-English bilinguals. *Brain and Language*, 73:421–431, 2000.

[28] C. N. Jackson and P. E. Dussias. Cross-linguistic differences and their impact on L2 sentence processing. *Bilingulism: Language and Cognition*, 12(1):65–82, 2008.

[29] N. Jiang. Morphologial insensitivity in second language processing. *Applied Psycholinguistics*, 25:603–634, 2004.

[30] D. Kimura. *Sex and Cognition*. MIT Press, Cambridge, MA, 1999.

[31] L. Kirsner, M. C. Smith, R. S. Lockhart, M. L. King, and M. Jain. The bilingual lexicon: Language-specific units in an integrated network. *Journal of Verbal Learning and Verbal Behavior*, 23(4):519–539, 1984.

[32] D. Klein, B. Milner, R. J. Zatorre, V. Zhao, and J. Nikelski. Cerebral organization in bilinguals: A PET study of Chinese-English verb generagion. *NeuroReport*, 10:2841–2846, 1999.

[33] K. Knight. Automating Knowledge Acquisition for Machine Translation. *AI Magazine*, 18(4):81–96, 1997.

[34] J. H. Kramer, D. C. Delis, E. Kaplan, L. O'Donnell, and A. Prifitera. Developmental sex differences in verbal learning. *Neuropsychology*, 11(4):577–584, 1997.

[35] J. F. Kroll and E. Stewart. Category Interference in Translation and Picture Naming: Evidence for Asymmetric Connections between Bilingual Memory Representations. *Journal of Memory and Language*, 33:149–174, 1994.

[36] S. C. Levinson. Language and Cognition: The Cognitive Consequences of Spatial Description in Guugu Yimithirr. *Journal of Linguistic Anthropology*, 7(1):98–131, 1997.

[37] P. Li, E. Bates, and B. MacWhinney. Processing a Language without Inflections: A Reaction Time Study of Sentence Interpretation in Chinese. *Journal of Memory and Language*, 32:169–192, 1993.

[38] B. MacWhinney. The Competition Model. In B. MacWhinney, editor, *Mechanisms of Language Acquisition*, pages 249–308. Lawrence Erlbaum, Hillsdale, NJ, 1987.

[39] B. MacWhinney, E. Bates, and R. Kliegl. Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, 23:127–150, 1984.

[40] M. Nagao. A framework of a mechanical translation between Japanese and English by analogy principle. In A. Elithorn and R. R. Banerji, editors, *Artificial and Human Intelligence*. Elsevier Science Publishers, 1984.

[41] M. G. Packard and B. J. Knowlton. Learning and Memory Functions of the Basal Ganglia. *Annual Review of Neuroscience*, 25:563–593, 2002.

[42] M. G. Packard and J. L. McGaugh. Inactivation of Hippocampus or Caudate Nucleus with Lidocaine Differentially Affects Expression of Place and Response Learning. *Neurobiology of Learning and Memory*, 65(1):65–72, 1996.

[43] M. Paradis, M. C. Goldblum, and R. Abidi. Alternate antagonism with paradoxical translation behavior in two bilingual aphasic patients. *Brain and Language*, 15:55–69, 1982.

[44] R. A. Poldrack and M. G. Packard. Competition between memory systems: Converging evidence from animal and human studies. *Neuropsychologia*, 41:245–251, 2003.

[45] M. C. Potter, K-F. So, B. van Eckardt, and Feldman L. B. Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior*, 23(1):23–38, 1984.

[46] C. J. Price, D. W. Green, and R. von Studnitz. A functional imaging study of translation and language switching. *Brain*, 122:2221–2235, 1999.

[47] E. Rosch. Natural categories. *Cognitive Psychology*, 3:328–350, 1976.

[48] E. Rosch, C. B. Mervis, W. Gray, D. Johnson, and P. Boyes-Braem. Basic objects in natural categories. *Cognitive Psychology*, 8:382–439, 1976.

[49] B. Russell. *Introduction to Mathematical Philosophy*. George Allen & Unwin, Ltd., London and The Macmillan Co., New York, 1919. Citation is based on the reprint by Dover Publications, Inc. 1993.

[50] F. Sánchez-Martínez, M. L. Forcada, and Way. A. Hybrid rule-based –example-based MT: Feeding apertium with sub-sentential translation units. In M. L. Forcada and A. Way, editors, *Proceedings of the 3rd Workshop on Example-Based Machine Translation*, pages 11–18, Dublin, Ireland, 2009.

[51] D. L. Scarborough, L. Gerard, and C. Cortese. Independence of lexical access in bilingual word recognition. *Journal of Verbal Learning and Verbal Behavior*, 23(1):84–99, 1984.

[52] D. L. Schacter and E. Tulving, editors. *Memory systems*. MIT Press, Cambridge, MA., 1994.

[53] M. T. Ullman. The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, 30(1):37–69, 2001.

[54] M. T. Ullman. Contributions of neural memory circuits to language: The declarative/procedural model. *Cognition*, 92(1-2):231–270, 2004.

[55] M. T. Ullman. A Cognitive Neuroscience Perspective on Second Language Acquisition: The Declarative/Procedural Model. In C. Sanz, editor, *Mind and Context in Adult SLA: Methods, Theory, and Practice*, pages 141–178, Washington, 2005. Georgetwon University Press.

[56] M. T. Ullman and E. I. Pierpont. Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, 41(3):399–433, 2005.

[57] H. Uszkoreit, F. Xu, and W. Liu. Challenges and Solutions of Multilingual and Translingual Information Service Systems. In *12th International Conference on Human-Computer Interaction*, pages 132–141, 2007.

[58] S. Vogel, F. J. Och, C. Tillmann, S. Nießen, H. Sawaf, and H. Ney. Statistical Methods for Machine Translation. In W. Wahlster, editor, *Verbmobil: Foundations of Speech-to-Speech Translation*, pages 377–393. Springer Verlag, Berlin, Heidelberg, New York, 2000.

[59] A.D Wagner, D.L. Schacter, M Rotte, W Koutstaal, A Maril, A.M Dale, B.R Rosen, and RL Buckner. Building memories: Remembering and forgetting of verbal experiences as predicted by brain activity. *Science*, 281:1188–1190, 1998.

[60] W. Wahlster, editor. *Verbmobil: Foundations of Speech-to-Speech Translation*. Springer Verlag, 2000.

[61] J. Walters. *Bilingualism: The Sociopragmatic-Psycholinguistic Interface*. LEA, Mahwar, New Jersey, 2005.

[62] W. Weaver. Translation (1949). In W. N. Locke and A. D. Booth, editors, *Machine Translation of Languages: Fourteen Essays*, pages 15–22. MIT Press, Cambridge, MA, 1955.

[63] D. B. Willingham. A Neuropsychological Theory of Motor Skill Learning. *Psychological Review*, 105:558, 1998.

[64] W. A. Woods. Semantics and quantification in natural language question answering. In D.G. Bobrow and A. Collins, editors, *Respresentation and Understanding*, pages 35–82. Academic Press, New York, 1975.

Design and Implementation of Image Analysis System by Applying Component-based Slow Intelligence System Framework

Liqun Kuang[12], Shi-Kuo Chang[2], Yingze Wang[2], Yao Sun[2]
[1] School of Electronics and Computer Science and Technology, North University of China, Taiyuan, Shanxi 030051, China
{kuang}@nuc.edu.cn
and
[2] Department of Computer Science, University of Pittsburgh, Pittsburgh, PA 15260, USA
{chang, yingzewang, yaosun }@cs.pitt.edu

***Abstract*— In order to improve the accuracy and adaptation ability of image analysis methods in accordance with surroundings, an image analysis system with different combinations of various parameterized algorithms is presented. This system is designed and implemented by a general framework of Component-based Slow Intelligence System. A face recognition experiment is conducted on this system. We split each face database to several parts, depending on some quality characteristics of the facial images (light variation, expression variation, angle variation, and normal) and apply different face recognition algorithms in each database part. The overall recognition performance is improved, and the experimental results show that our method is superior to individual image analysis methods.**

Keywords-Component-based software system; slow intelligence system; image analysis; super component

I. INTRODUCTION

In order to facilitate a deeper understanding of image and interpretation, image analysis employs mathematical models with image processing techniques to analyze the underlying characteristics and the upper structure, and to extract various parameters which representing the image characteristics and demonstrating certain intelligent behaviors. The research of image analysis has been widely conducted into multiple applications of face recognition, emotion recognition, handwriting recognition, OCR-Optical character recognition, biomedical image analysis, CBIR-Content based image retrieval, and video object extraction, etc [1]. There are many kinds of images analysis methods to solve diversifor m problems, such as Bayesian decision making, linear classification, nearest neighbor classification, clustering algorithms, neural networks and support vector machines, etc [1]. Although image analysis has achieved a lot of research fruits, and has applied in specific applications in many areas, there are still many problems to be solved, e.g. establishing a common theoretical basis. In fact there are many kinds of image analysis algorithms on hand, but one might face the difficulty of choosing the appropriate algorithm in practice, as each of them has its own advantage in different case.

Recently, Chang [2] proposed a general intelligent framework called slow intelligence system (SIS). Chang et. al also presented the design of component based SIS and the user interface design to produce and manage the generic SIS system [3-6]. In this paper, we use the idea of SIS and apply it to the image analysis task. We propose a new image analysis system, which uses different image analysis algorithms as the candidate methods. Our system extracts image features using different methods, searches for the best method according to the status quo and propagates the learned knowledge over iterations. As we show, it gives the superior performance and can combine different image analysis methods into one system.

The paper is organized as follows. In Section II, we present a method of dual visual representations and a tool for general component-based system. In Section III, the principle of SIS and the design tool for the specific component-based SIS are introduced. In Section IV, our image analysis system based specific component-based SIS is presented and its virtual development tool is utilized. In Section V, we demonstrate the superiority of our system in human learning by solving the face recognition problems. In Section VI, we give the conclusion of our work and point out some possible future works.

II. A GENERAL COMPONENT-BASED SYSTEM

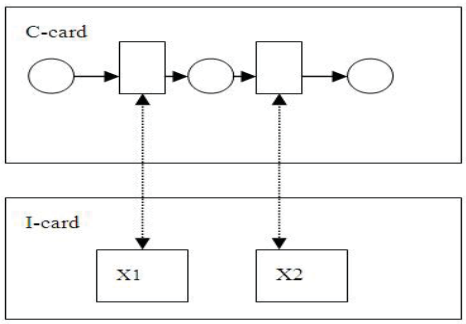


Figure 1. Dual visual representations.

A component-base system is characterized by splitting all processes into separate components so that all of the data and functions inside each component are semantically related. We have designed and implemented a general component-based system which can create component-based system framework rapidly by designing the system's control-card (C-card) and information-card (I-card) [4]. A generic example is illustrated in figure 1. The main idea is to employ two visual representations, the I-card and the C-card together to specify a component-based software system. The I-card specifies the

logical relationships among the co mponents, and the C-card specifies the control and t iming relationships am ong the components as Petri net or sequence diagra m in UML. Their inter-relationships are re presented by bi-directional arcs (the dotted lines) connecting components, or other entities, in the I-card and the C-card. With regard to system-wide co-ordination, components communicate w ith each other via XML format messages which should be authorized ahead of sendi ng or receiving by any component.

We have provided a visual developm ent tool (SISProjectCreator) for spec ifying and creating a tem plate project of ge neral component-based system or/and specific component-based SIS with a specific application as show n in figure 2. This tool can achie ves three tasks while creating a general component-based system: (1) specifying C-Card, which can define functional depe ndency repertoire by s pecifying messages transferred among components; (2) specifying I-Card, which can define components; (3) If a c omponent is a SIS application, it can be replaced by such the specific SIS system at run-time via specifying it as a specific SIS component at design time.

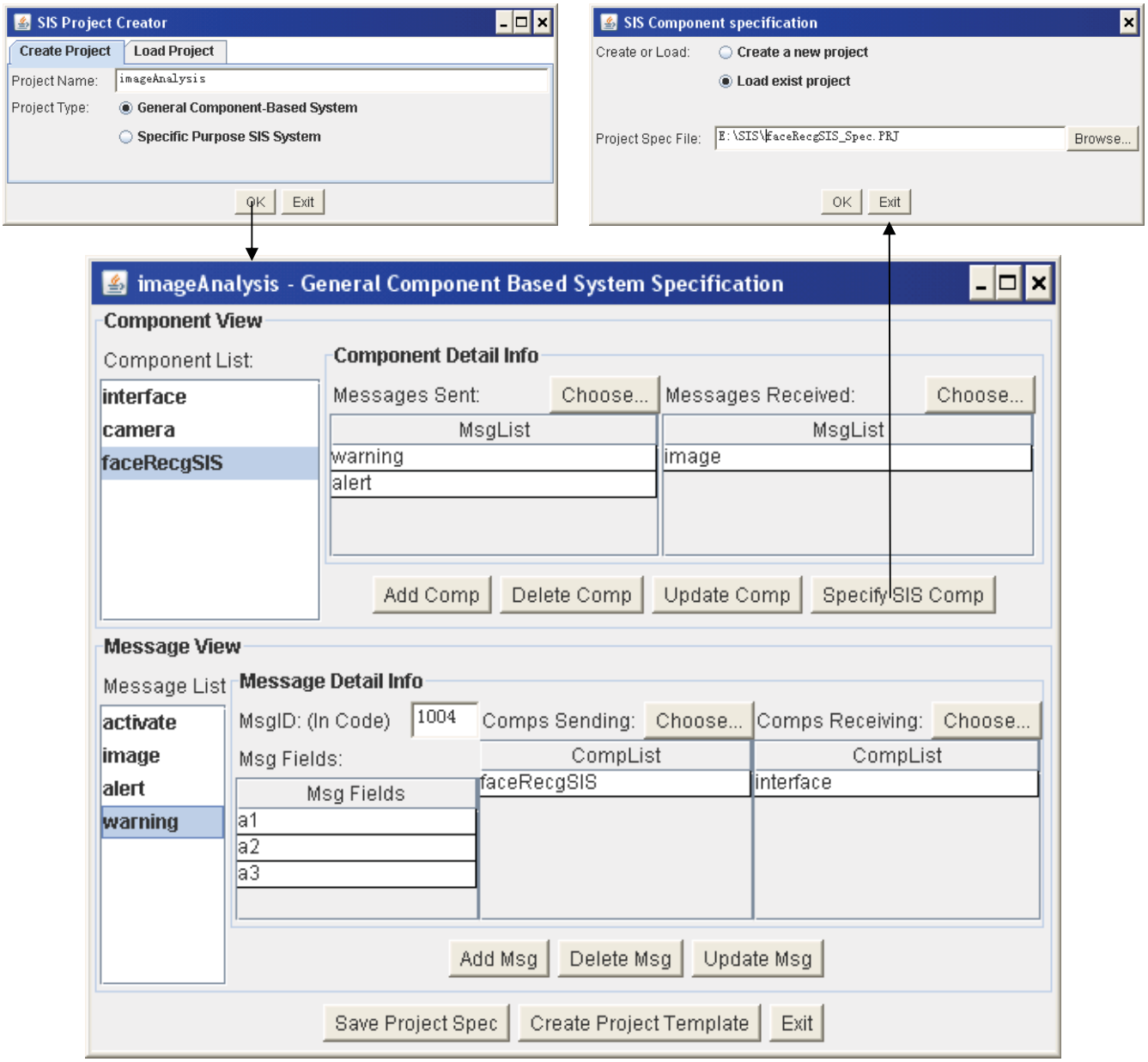


Figure 2. Visual development tool, SISProjectCreator, for general component-based system.

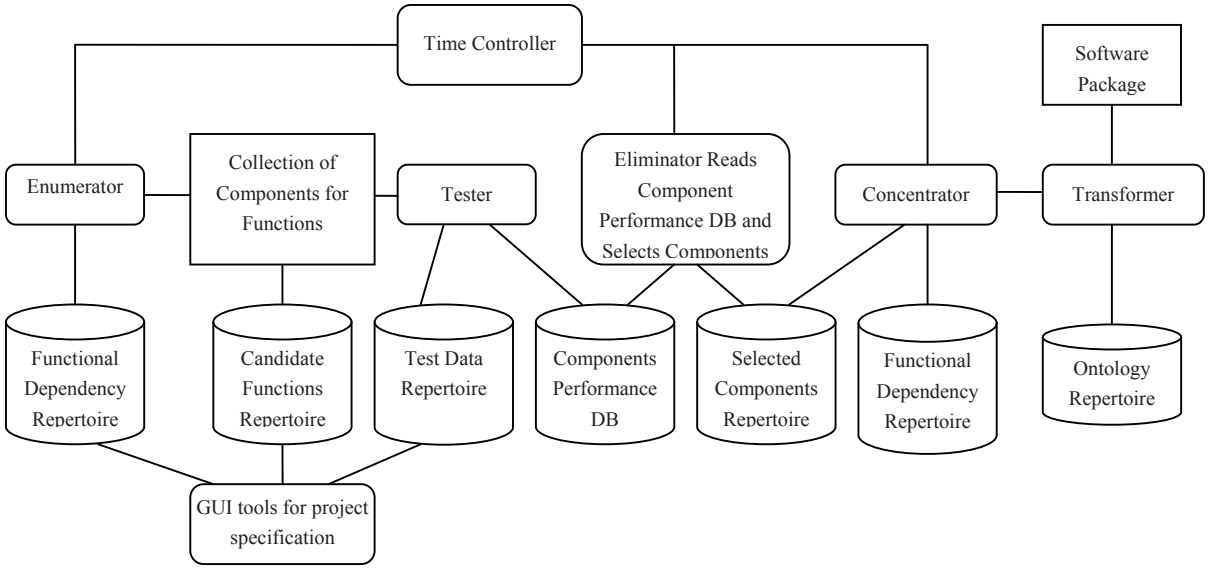


Figure 3. System architecture of component-based Slow Intelligence System.

III. A SPECIFIC COMPONENT-BASED SIS

A specific component-based SIS is characterized by being able to improve performance over time through a process involving enumeration, propagation, adaptation, elimination and concentration. Figure 3 illustrates the system architecture containing the key components of the specific SIS system. A SIS continuously learns, searches for new solutions and propagates and shares its experience with other peers. It has multiple decision cycles such that actions of slow decision cycle(s) may override actions of quick decision cycle(s), resulting in poorer performance in the short run but better performance in the long-run. [2]

We have implemented a specific component-based SIS in Java language [3]. First a SIS server is implemented. It deals with all the messages transferring among components. Messages are the only way that components could communicate with each other and all the messages within the SIS system are routed through the SIS server. Components could be classified into two classes: system components and candidate components. System components include Enumerator, Eliminator, Propagator, Adaptator, Concentrator. They are responsible for initializing and managing the SIS system. Candidate components wrap various algorithms to provide multiple performance for different circumstances. Upon receiving the data message from the SIS server, they will work on the dataset and send back the results to SIS server. The SIS server could forward the results to corresponding components like Eliminator. Unless Eliminator close a certain candidate component, candidate components are always running and servicing.

Like general component-base system, the visual development tool (SIS ProjectCreator) can also be applied to specify and create a template project of specific component-based SIS while developing a new SIS application. The GUI of the tool is shown in figure 4. This tool can accomplish three specific tasks: (1) specifying a more sophisticated C-Card, which can define functional dependency repertoire in Time Controller, by specifying multiple cycles and their executive rules; environmental variable can also be introduced to affect the route. However, the messages transferred among components can not be specified as they are hard-coded in SIS; (2) specifying I-Card, which can define candidate functions repertories in each cycle where the functional dependency can be parallel or sequential; (3) specifying testing data, which is stored in test data repertoire.

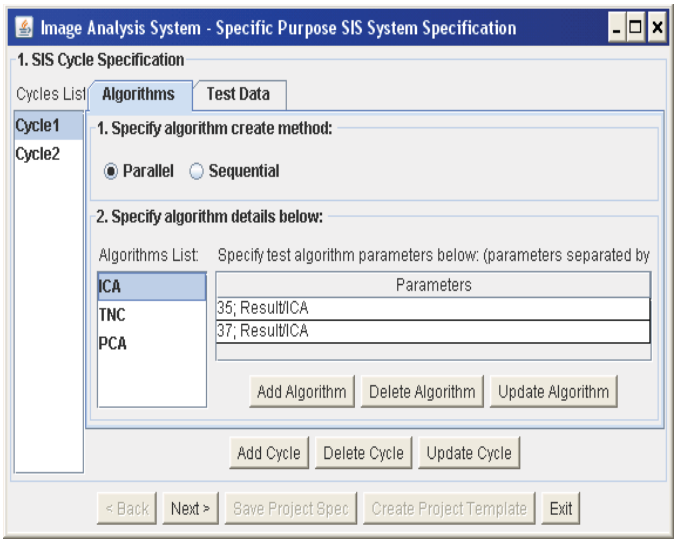


Figure 4. Visual development tool, SISProjectCreator, for specific component-based SIS.

A specific component-based SIS can run independently. It can also be wrapped into a general component-based system as a SIS component, as shown in figure 2. An important feature of components is that they are substitutable, so that a component can replace another (at design time or run-time), if the successor component meets the requirements of the initial

component. Consequently, components can be replaced with either an updated version or an alternative without affecting the general system. In this case, the faceRecgSIS component defined in figure 2 can be replaced by the specific component-based SIS in figure 4 while the general component-base system is running.

IV. Design and Implementation of Image Analysis System

Firstly, we design the C-card and I-card for the image analysis system with dual visual representations approach. The partial visual representations for SIA are shown in figure 5. The images being analyzed are sent to Image Extractor fromImage Sender, so that the main part of image is extracted. Next the Enumerator invokes the only one functional block (super component) containing different candidate components for image analysis. All the candidate components in this super component are the different specified algorithms or the generic algorithm with different parameters doing the same job for analyzing images. We have implemented three parameterized algorithms: principal component analysis (PCA) [7-8], independent component analysis (ICA) [9] and 3-neighbor classification (TNC) [10] respectively. In each model, there are different parameters to control its performance. Then the Tester collects and presents the test results. The Time Controller restarts the image analysis cycle with a different super component. The Eliminator eliminates the inferior parameterized algorithms, and finally the Concentrator selects the best parameterized algorithm based on relative increase in correctness.

Then we use the virtual tool SISProjectCreator to implement our previous design of I-Card, C-Card and testing data respectively. Initially, we specify testing data as shown in figure 6 for testing each candidate component. All testing data are images and are stored under a folder. Then I-card for each component (super component and/or simple component) is required to define as shown in figure 7. For the super component containing a number of candidate algorithms, we need to use I-Card to add each candidate component from the template algorithm with different parameters to this super component. We have implemented three algorithms for image analysis super component, namely, TIN, PCA and ICA. For a simple component in Petri-net, although it includes only one component, we still need to add it from database or specific algorithm. In the end, we create a template component-based SIS project by the virtual tool SISProjectCreator. All required components are created and communicate with each other based on C-Card. We use Petri Net to define the dependencies among different functional components as shown in Figure 8. In the system, there is only one supper component "Image Analysis" which depends on Image Extractor component and sends its result to Verifier component. Places in Petri Net represent the messages sending from one transition to another, since SIS uses messages to communicate among all components.

In addition, Cycle switch rules and environment variables can be defined in SIS Time Controller. There are two methods to specify cycle switch rules, mapping table and matrix, which can control the sequence of decision cycles. We define a rule by mapping table that controls Cycle 2 running after Cycle 1 shown as figure 9. We can also define environment variables to specify whether a rule is effective or not. As we know the image quality is very important to any image analysis system, here we define an environment variable, imageQuality, to control whether the defined rule is applicable. Only when the processed images meet the given condition (imageQuality>0.5), then rule1 can be effective, which means cycle1 can proceed to cycle2.

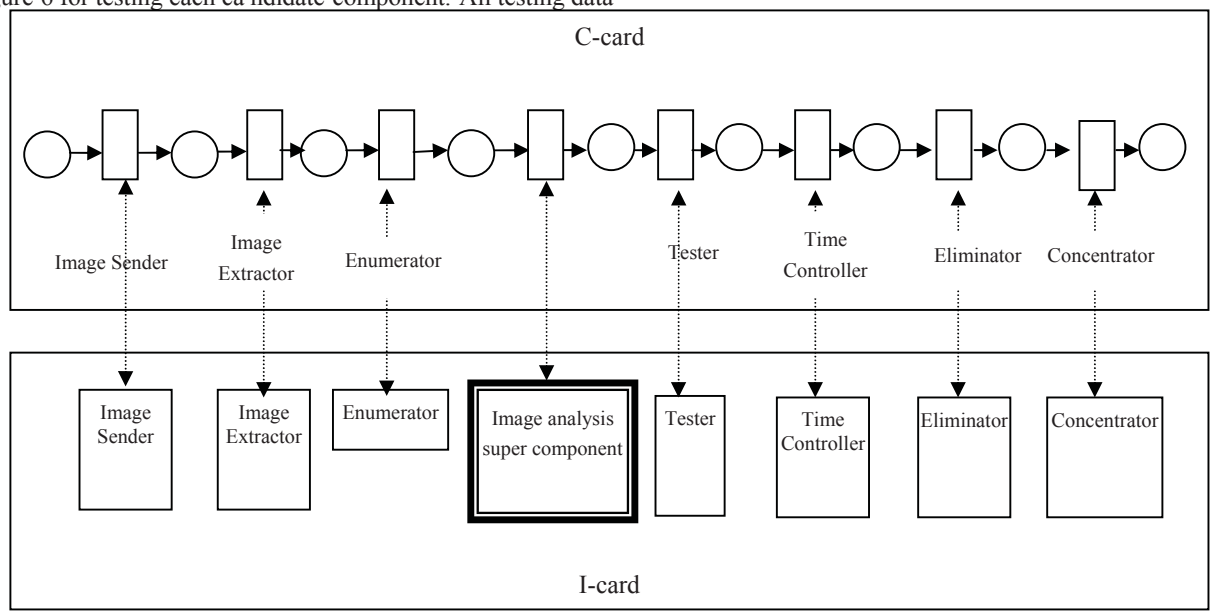


Figure 5. Dual visual representations for image analysis system.

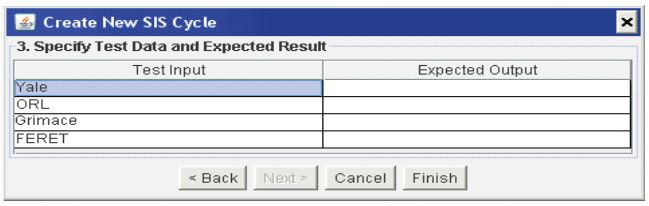


Figure 6. Specification of testing data with SISProjectCreator.

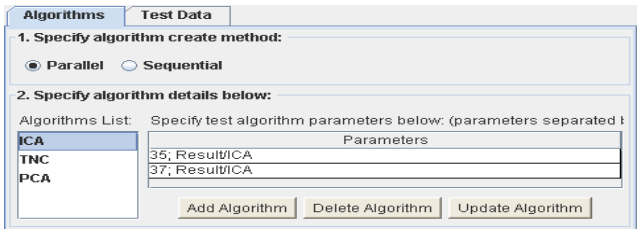


Figure 7. Specification of I-card with SISProjectCreator.

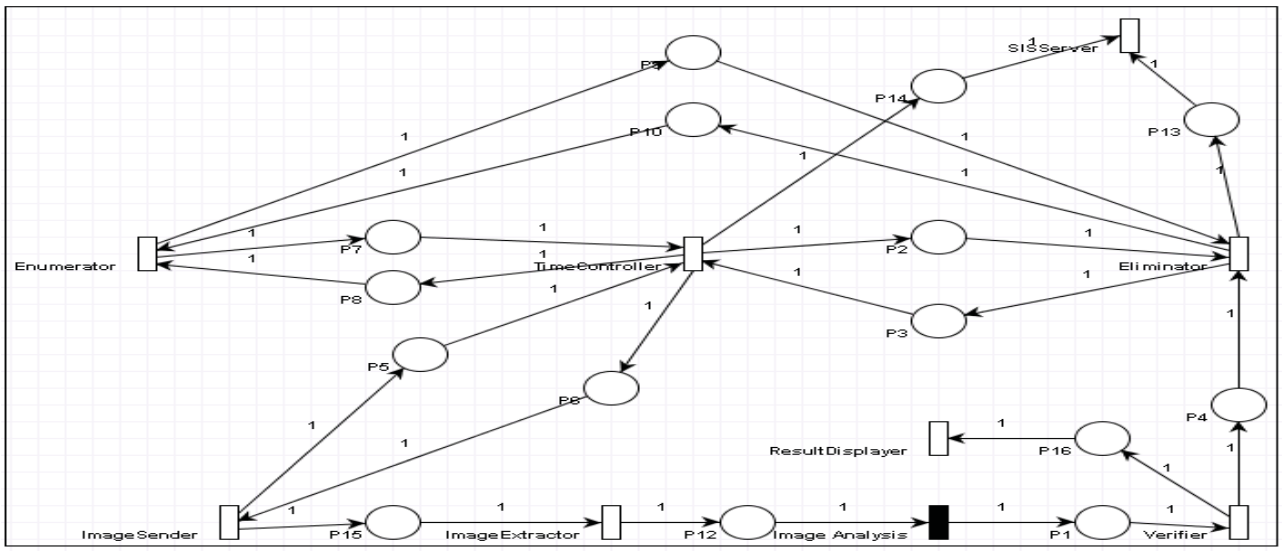


Figure 8. Design of C-card for dependencies among different functional components

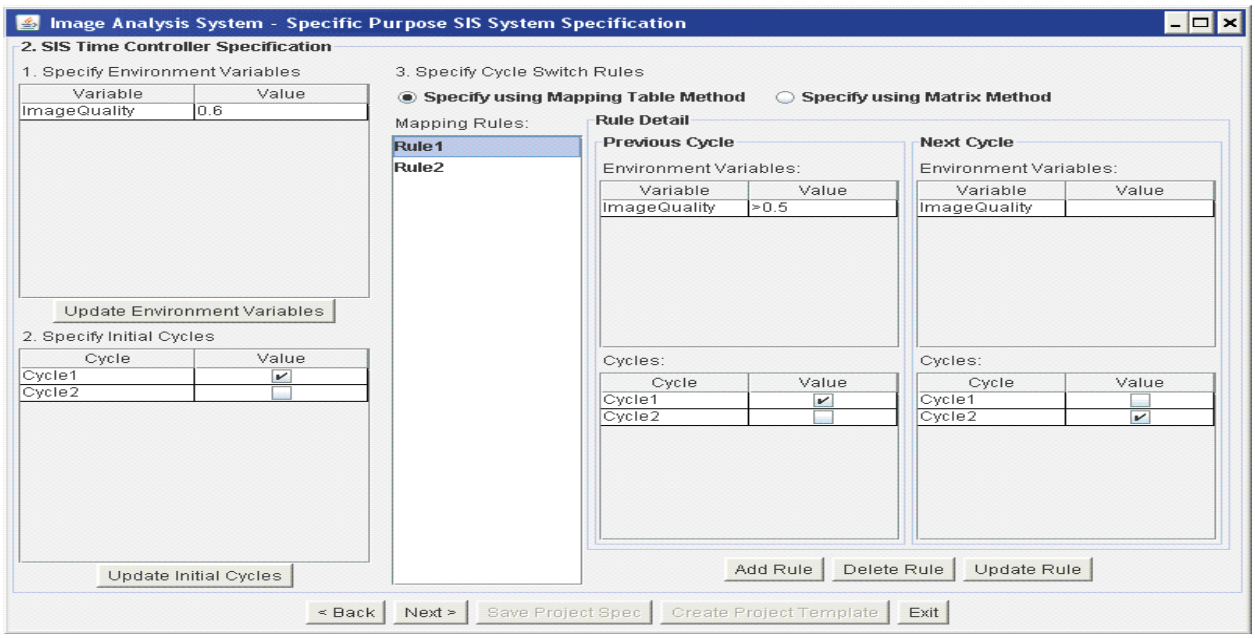


Figure 9. screenshot for time controller specification.

V. EXPERIMENT OF FACE RECOGNITION

Face recognition is a hot research topic in pattern recognition area. It analyzes images of human faces and extracts the effective features to identify status. It has a wide range of applications in identity verification, criminal identification, scene monitoring, human-computer interaction, visual communication, etc. In the practice of human-computer interaction, although there are a lot of image recognition algorithms available, the end user does not know which method is the best one for the particular illumination, facial expressions and gestures. Thus, we conduct a face recognition experiment by our image analysis system to help human learning. In this section, we describe the experimental data and the experiment steps, and then analyze the face recognition results .

A. Experimental Data Description

Four human face databases named FERET, Yale, ORL and Grimace are used in this experiment [11], as shown in figure 10-13 respectively. Firstly, we use the FERET face database [12]. Harry Wechsler et al. build the FERET face database. The facial images were collected independently from the FERET program's developers. There was some minor variation in images collected on different dates. The FERET database contains 1564 sets of images for a total of 14,126 images that includes 1199 individuals and 365 duplicate sets of images. For some individuals, over two years had elapsed between their first and last sittings, with some subjects being photographed multiple times. This time lapse was important because it enabled researchers to study, for the first time, changes in a subject's appearance that occur over a year.

Next database we use is the Yale face database [8]. The Yale Face Database contains 165 grayscale images in GIF format of 15 individuals. There are 11 images per subject, one per different facial expression or configuration: center-light, w/glasses, happy, left-light, w/no glasses, normal, right-light, sad, sleepy, surprised, and wink.

Then we use the ORL face database [7] [10], which contains a set of face images taken between April 1992 and April 1994 at AT&T Laboratories Cambridge. There are ten different images of each of 40 distinct subjects. For some subjects, the images were taken at different times, varying the lighting, facial expressions (open / closed eyes, smiling / not smiling) and facial details (glasses / no glasses). All the images were taken against a dark homogeneous background with the subjects in an upright, frontal position (with tolerance for some side movement).

At last we use the Grimace face database [13]. A sequence of 20 images per subject was taken, using a fixed camera. During the sequence the subject moves his/her head and makes grimaces which get more extreme towards the end of the sequence. There is about 0.5 seconds between successive frames in the sequence.

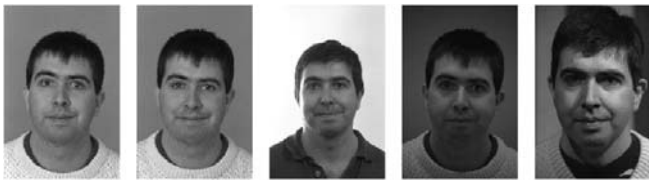


Figure 10. FERET face database.

Figure 11. Yale face database.

Figure 12. ORL face database.

Figure 13. Grimace face database.

B. Experimental Protocol

We have three face recognition algorithms to deal with those face database, named PCA, ICA and TNC. The users would be confused as they don't know which algorithm is the best or most appropriate in recognizing a given subject. We set up this experiment to help users improve their skill in the application of face recognition.

Users may reorganize those face databases according to their experience. As we know the images effects are very important to any face recognition method, users can reorganize those four face databases into another four face databases in accordance with the variation degree of light, facial expression and face angle. The default database is Normal if there is no obvious variation of facial images. The groups of reorganized face database from original databases are shown in Table I.

TABLE I. SUBJECT NUMBER IN REORGANIZED FACE DATABASES

Original Face Databases	**Reorganized Face Databases**			
	Light Variation	***Expression Variation***	***Angle Variation***	***Normal***
FERET	40	15	10	17
Yale	3	7	3	2
ORL	8	6	23	3
Grimace	3	10	3	2

Then the image analysis system runs all three algorithms in super component for those re organized face databases to find out which parameterized algorithm is most efficient to particular kind of face images. For each subject, the first half images are used as train samples and the rem ains are testing samples. A feature matrix is obtained by each face recogni tion algorithm for each subject and then is used to verify the testing samples. The accurate rate in each group of testing samples is shown in table II. The algorithm with highest accurate rate in each group of reorganized face database is chosen to match a circumstance. For example, if the a ccurate rate of a parameterized algorithm is Max $\{L_1, L_2, L_3\}$, this algorithm is chosen as the most efficient algorith m in processing face images of light variation. It is helpful for the users as they learn which method (including its param eter value) is most appropriate to select in a specified circumstance. The users will become more and more proficient in utilizing im age analysis methods by reorganizing face databases in different ways, which can improve the human-machine interaction learning.

In the end we compare the accurate rate of the learned method with that of indi vidual face recognition method in original four face databases.

C. *Experimental Results Analysis*

We have implemented SIS appr oach successfully with TNC, PCA a nd ICA. In our system, the Tim e Controller component is responsible for starting instances of all three face recognition algorithms with various para meters. While the system is running, we can observe the messages communication among all co mponents through an administrator tool of SIS system as shown in figure 14 which is helpful to understand the designed I-card and C-card [3]. The experimental results are displayed in ResultDis player component as shown in figure 15.

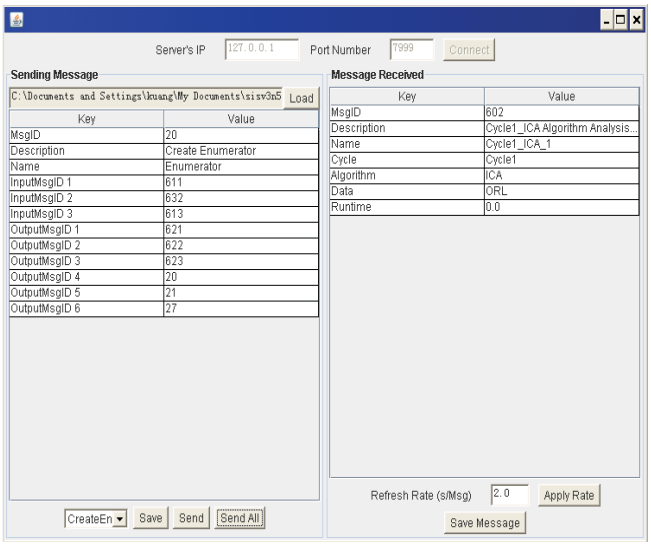


Figure 14. messages captured from the admistrator tool of SIS system.

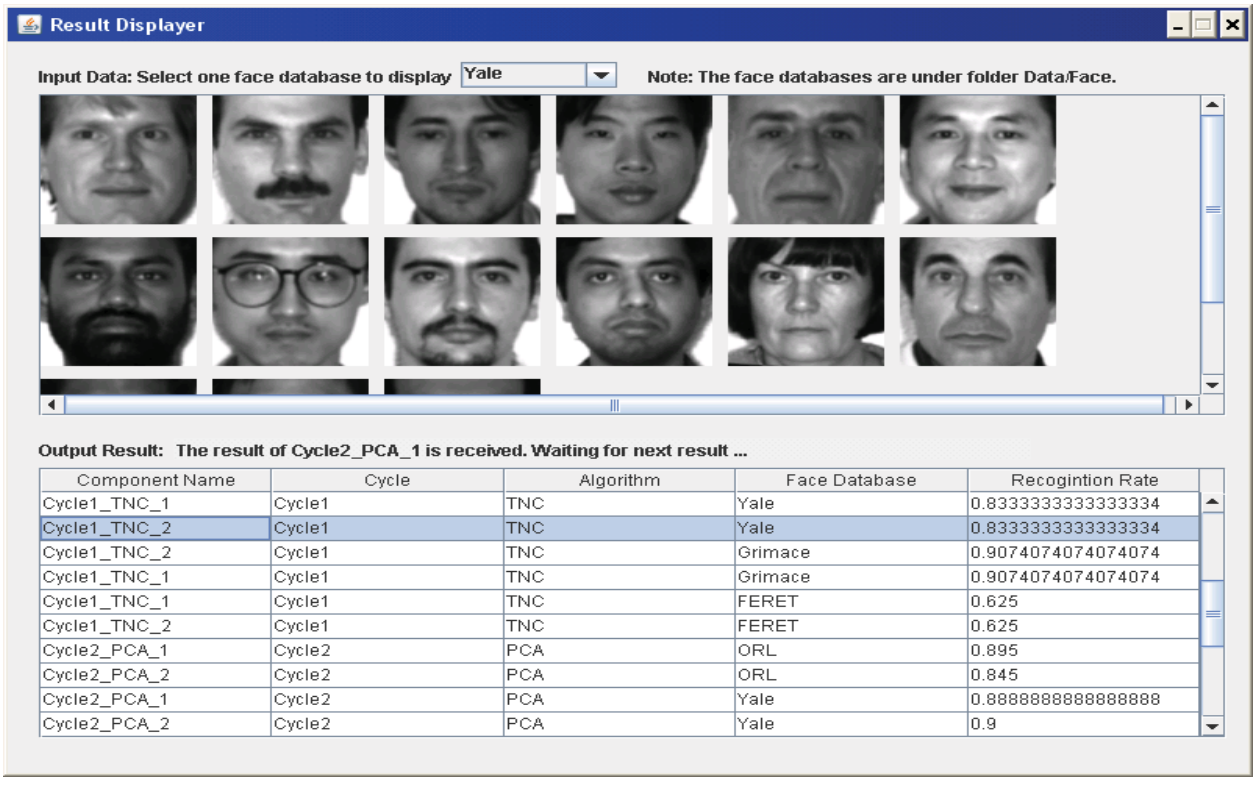


Figure 15. screenshot for experimental results in resultDisplaer componet

After running 50 cycles of all three algorithms by trying 50 appropriate combination groups of parameter values in our experiment, the accurate rates of algorithms for all groups of reorganized face databases are reported in Table II. The highest accurate rate algorithm for the variation of light, expression and angle is ICA, PCA and TNC respectively. If the variation is normal, the best algorithm is TNC. These best algorithms with image characteristics are shown in Table III and users are supposed to memorize them. But users don't need to remember the algorithm's best parameter sets as they can be automatically retrieved once in need. In fact, the more classification ways for original face database the users try, the more proficient they will be.

Then, we compare that trained human user with the three individual face recognition algorithms in our super component for all four original face databases. As we can see from Table IV, the trained human user greatly outperforms individual face recognition algorithm. For individual face database, Grimace achieves better accurate recognition rate as compared to ORL, Yale and FERET. It is because the face variation of each group in Grimace is very minor. However, it is hard to say which algorithm is best for all databases. The training human user utilizing all these three methods in SIS is superior to any individual one.

TABLE II. THE ACCURATE RATE OF ALGORITHMS IN REORGANIZED FACE DATABASES

Algorithms	**Reorganized Face Databases**			
	Light Variation	***Expression Variation***	***Angle Variation***	***Normal***
TNC	L_1=86.2%	E_1=87.3%	A_1=89.2%	N_1=95.6%
PCA	L_2=86.7%	E_2=91.6%	A_2=92.3%	N_2=92.4%
ICA	L_3=89.4%	E_3=88.5%	A_3=87.7%	N_3=93.3%

TABLE III. THE MATCHES OF ALGORITHMS WITH IMAGES CHARACTERISTICS

	Reorganized Face Databases			
	Light Variation	***Expression Variation***	***Angle Variation***	***Normal***
Best Algorithm	ICA	PCA	PCA	TNC

TABLE IV. PERFORMANCE OF THE ORIGINAL FACE DATABASES: THE ACCURATE RATE

Face Recognition Algorithms	**Original Face Databases**			
	FERET	***Yale***	***ORL***	***Grimace***
TNC	62.5%	83.3%	88%	96.7%
PCA	73.6%	88.9%	89.5%	98%
ICA	83.3%	82.2%	84%	96.7%
Trained human	87.4%	90.1%	92.3%	99.3%

VI. CONCLUSION AND FUTURE WORKS

We design and implement a new image analysis system using the general component-based SIS. It iteratively selects the best individual image analysis algorithm and propagates the learned knowledge over iterations. It leads a superior accuracy of image analysis to individual image analysis method. We test our system in face recognition. The experimental results show that our method is helpful to improve human's skill in face recognition. For the future work, we will apply our system in medical image analysis to help training medical novice and improving their diagnosis ability.

ACKNOWLEDGMENT

The authors are thankful to Dr. Libor Spacek from the University of Essex, AT&T Laboratories Cambridge, FERET program and Yale University for making the Grimace, ORL, FERET and Yale face-image database available online. The authors are also thankful to anonymous reviewers for their constructive comments.

REFERENCES

[1] Xuan Wang, Jiafeng Ma. " Digital image analysis and pattern recognition". Beijing: Science Press, 2011.242

[2] Shi-Kuo Chang, "A General Framework for Slow Intelligence Systems", International Journal of Software Engineering and Knowledge Engineering, Volume 20, Number 1, February 2010, 1-16.

[3] Shi-Kuo Chang, Yao Sun, Yingze Wang, Chia-Chun Shih and Ting-Chun Peng, "Design of Component-based Slow Intelligence Systems and Application to Social Influence Analysis", Proceedings of 2011 International Conference on Software Engineering and Knowledge Engineering, Miami, USA, July 7-9, 2011, 9-16.

[4] S Shi-Kuo Chang, Yingze Wang and Yao Sun, "Visual Specification of Component-based Slow Intelligence Systems", Proceedings of 2011 International Conference on Software Engineering and Knowledge Engineering, Miami, USA, July 7-9, 2011, 1-8.

[5] Ji Eun Kim, Yang Hu, Shi-Kuo Chang, Chia-Chun Shih and Ting-Chun Peng, "Design and Modeling of Topic/Trend Detection System By Applying Slow Intelligence System Principles", Proc. of DMS2011 Conference, Florence, Italy, Aug. 18-20, 2011, 3-9.

[6] Yingze Wang and Shi-Kuo Chang, "High Dimensional Feature Selection via a Slow Intelligence Approach", Proc. of DMS2011 Conference, Florence, Italy, Aug. 18-20, 2011, 10-15.

[7] Jiao Binliang and Chen Shuang, "Face recognition based on PCA", Computer Engineering and Applications, 2011,47(18), 201-203.

[8] Luo Lin, Zou Cairong and Yang Fengfan, "Modified algorithm of principal component analysis for face recognition", Journal of SoutheastUniversity (English Edition), 2006,l22(1), 26-30.

[9] Duan Zhi-juan, Zhou Xin-xin, Chen Fen-xiong and Sun Lin. "Recognition Based on Independent Component Analysis for Surface Detect of Cold Strip", Journal of Iron and Steel Research. 2011,23(10),63-66.

[10] Zhang Jian-ge and Liu Hong-bo, "Simulation of face recognition algorithm based on WTPCA and 3-neighbor classification", Computer Engineering and Applications, 2009,45(11),175-177.

[11] http://www.face-rec.org/databases/ , Last update: November 30, 2011.

[12] P. J. Phillips, H. Moon, P. J. Rauss, and S. Rizvi, "The FERET evaluation methodology for face recognition algorithms", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, October 2000, 1090-1104.

[13] http://cswww.essex.ac.uk/mv/allfaces/grimace.html, uploaded by Dr Libor Spacek, Feb 16, 2007

A Framework for Intrusion Detection System based on the Slow Intelligent Approach

Francesco Colace[1], Massimo De Santo[1], Salvatore Ferrandino[2], S.K. Chang[3]
[1] DIEII – Università degli Studi di Salerno, via Ponte Don Melillo
84084 Fisciano (Salerno), Italy
{fcolace, desanto}@unisa.it
[2] Ufficio Sistemi Tecnologici – Università degli Studi di Salerno, via Ponte Don Melillo
84084 Fisciano (Salerno), Italy
salfer@unisa.it
[3] Computer Science Department – University of Pittsburgh, Sennott Square, Pittsburgh, PA 15260 USA
chang@ cs.pitt.edu

Abstract – Computer network security is a fashionable and fast-moving field. In the last decade many methodologies and tools have been developed for improving the security of networks and their hosts, but the resources used to deal with the problem often do not yield results commensurate with costs. In the last period the adoption of Network Intrusion Prevention Systems promises to represent an effective line of defense against a variety of attacks that could compromise the security and proper functioning of an enterprise information system. This paper introduces a Network Intrusion Prevention System based on Ontological and Slow Intelligence approach. By the use of Ontology the proposed Network Intrusion Prevention System will analyze the input semantically while will improve over the time sharing knowledge among other similar systems or experts according to the Slow Intelligence approach. A first prototype of the environment has been developed and first experimental results have been showed.

Keywords: Network Security, Network Management, Ontology, Slow Intelligence Approach

I. Introduction

Network security is one of the hottest and challenging topics both in real operative contexts and in scientific literature. With the huge development and diffusion of web based services, the Network has become the most critical resource for organizations and companies. Social Networks, Search Engines, Web Applications, e-business, e-learning, information sharing communities and the huge variety of network based services have increased the online traffic and the number of network attacks. So the security of web applications has become increasingly important and a secure web environment has become a high priority the various communities. In literature there are several security assessments on this topic and in general they assess that at least the 90% of web application are vulnerable to some form of attack [1]. In this scenario an effective role is played by the Intrusion Detection and Prevention Systems (IDPS). IDPS are network security appliances that monitor network and/or system activities for malicious activity. The main functions of intrusion prevention systems are to identify malicious activity, log information about said activity, attempt to block/stop activity, and report activity [2]. Detecting or preventing intrusions are not trivial tasks in current networked environments since the open nature of a network allows the exploitation of its resources even by unauthorized users or in unauthorized ways. Although many IDPS systems have been proposed, their appropriate configuration and control for effective detection or prevention and efficient resources consumption have always been challenging [3]. A real problem in the use of IDPS is its production of a large number of false positives and more in general its capacity to manage only already known attacks. In fact this kind of systems cannot fruitfully use the past experience in the detection (or not) of misuses. The main limitation of these solutions is in their lack of diversity and correlation [4]. From the point of view of the diversity a real improvement is the introduction of heterogeneous sensors able to manage and process various data streams. The correlation can be improved by the use of semantic web techniques and methods [5]. In particular many researchers are investigating the use of ontology as support for IDPS: Undercoffer et al [6], for example, experimented the use of ontology as mean for the characterization and classification of attack or intrusion. Ontologies can escalate from attack symptoms to intrusion diagnosis and infer new knowledge. The previous considerations are the starting point for this paper that proposes a novel approach for designing an IDPS. In particular the main idea is the use of the Slow Intelligence Approach in order to develop a distributed and collaborative environment. Thanks to the Slow Intelligence Approach the proposed IDPS can improve continuously its knowledge base, expressed through the ontological formalism, thanks to the sharing of information with other similar peers that work in other contexts or with experts that can introduce new concepts. In this way the IDPS continuously learn and adapt its behavior along the time. The remainder of this paper is organized as follows: the next section presents related work in the area of IDPS and in the domain of using semantic web methodology in the IDPS. Section three and four present the proposed approach giving

more details on the Slow Intelligence Approach and Ontology. Section five presents the first results obtained using a first system prototype in a real scenario. Finally conclusion is in the end section.

II. Related Work

There are many scientific and commercial tools for intrusion detection and prevention. Usually these tools work according to well defined strategies or rulesets. One of the most famous is SNORT [7], an open source network intrusion prevention and detection system developed by Sourcefire. Combining the benefits of signature, protocol, and anomaly-based inspection Snort can detect intrusions. This system works according to a well-defined rule set that has to be updated in order to have the latest detection capabilities. Each Snort user may opt to manually download and updates rules files, however most Snort users automate the process using PulledPork, an open source Perl script. The effectiveness of this kind of approach is strictly related to the capacity of the system to be continuously updated. Snort adopts static rules and so lacks of semantic approaching the resolution of an intrusion: it adopts a syntactical approach. Using semantic web methodology in intrusion prevention and detection system is quite new. The first researches on the introduction of ontology as support for the computer network security tools has been developed in [5] and [8]. In [6] and [9] authors introduced an ontology specifying a model of computer attack. Their ontology is based upon an analysis of over 4.000 classes of computer intrusions and their corresponding attack strategies and it is categorized according to system component targeted, means of attack, consequence of attack and location of attacker. The main idea of these papers is in the statement that there are benefits of transitioning from taxonomies, that can be considered as a syntactical approach, to ontologies and ontology specification languages, which are able to simultaneously serve as recognition, reporting and correlation languages. Ontologies, unlike taxonomies, provide powerful constructs that include machine interpretable definitions of the concepts within a domain and the relations between them. Ontologies provide software systems with the ability to share a common understanding of the information at issue, in turn empowering software systems with a greater ability to reason over and analyze this information. In the last period many approaches based on ontology has been introduced. In [10] ontology has been used to extract semantic relations between computer attacks and intrusions. Every time the system detects an attack or anomalous condition, it tries to extract the semantic relationship among computer attacks and suspected situations in the network with proposed ontology. In this way ontology can be considered as a sort of a bridge between the attacks and their effects in the network. In this way the system can infer dynamically, thanks to the ontology formalism, how to react to an attack. Another approach is in [11] where ontology allows a semantic analysis of input, refining and extending the ontology over time. By the use of ontology, representing the various components of a malicious intrusion (kind of communication, number of the ports involved, …) the system can focus on specific portion of network packet where attack is possible: in this way the research space can be refined avoiding sequential search. Another interesting approach is in [4] where ontology has been used in order to drive the correlation process. The proposed approach relies on the use of ontologies to escalate from attack symptoms to intrusion diagnosis. In particular thanks to a decision engine can correlate the potential attack that can bring to such a symptom. The engine correlates the detected symptom with other incoming ones, combining the confidence of the sensors, which is formalized in the ontology, with the level of intensity of the symptoms, so producing more accurate alerts. As matter of fact, alerts are not generated as results of all detected symptoms, but only when the correlation between such symptoms indicates a really potential attach. This approach could be considered as the starting point for the IDPS proposed in this paper. The main idea is to use ontology as a semantic middleware for inferring actions to apply against the intrusion starting from the data streams that are in the network. In particular the system will improve its ability to recognize and solve problems sharing knowledge, represented by the use of ontology, and communicating with other peer that are working in different operative scenarios. In the next sections the system will be described in more detail.

III. Slow Intelligence Approach and Ontology

As previously said the main aim of this paper is the introduction of an Intrusion Prevention and Detection System based on the adoption of Slow Intelligence approach as inference engine and ontology for the representation of system knowledge base. First of all the concept of Slow Intelligence and the general framework for designing and specifying Slow Intelligence Systems (SIS) will be introduced. A Slow Intelligence System can be considered as a general-purpose system characterized by being able to improve performance over time through a process involving enumeration, propagation, adaptation, elimination and concentration [12]. A Slow Intelligence System continuously learns, searches for new solutions and propagates and shares its experience with other peers which work in different operative scenarios. A Slow Intelligence System differs from expert systems in that the learning is implicit and not always obvious. A Slow Intelligence System seems to be a slow learner because it analyzes the environmental changes and carefully and gradually absorbs that into its knowledge base while maintaining synergy with the environment. A slow intelligence system:

- solves problems by trying different solutions

- is context-aware to adapt to different situations and to propagate knowledge
- may not perform well in the short run but continuously learns to improve its performance over time

Slow Intelligence Systems typically exhibit the following operative characteristics:

- **Enumeration**: In problem solving, different solutions are enumerated until the appropriate solution or solutions can be found.
- **Propagation**: The system is aware of its environment and constantly exchanges information with the environment. Through this constant information exchange, one SIS may propagate information to other (logically or physically adjacent) SISs.
- **Adaptation**: Solutions are enumerated and adapted to the environment. Sometimes adapted solutions are mutations that transcend enumerated solutions of the past.
- **Elimination**: Unsuitable solutions are eliminated, so that only suitable solutions are further considered.
- **Concentration**: Among the suitable solutions left, resources are further concentrated to only one (or at most a few) of the suitable solutions.

There is one interesting and maybe unique characteristic for SIS:

- **Slow decision cycle(s) to complement quick decision cycle(s):** SIS possesses at least two decision cycles. The first one, defined as the quick decision cycle, provides an instantaneous response to the environment. The second one, defined as the slow decision cycle, tries to follow the gradual changes in the environment and analyze the information acquired by experts and past experiences. The two decision cycles enable the SIS to both cope with the environment and meet long-term goals. Sophisticated SIS may possess multiple slow decision cycles and multiple quick decision cycles. Most importantly, actions of slow decision cycle(s) may override actions of quick decision cycle(s), resulting in poorer performance in the short run but better performance in the long run.

The structure of SIS can be by the introduction of the basic building block and advanced building block. Figure 1 illustrates the Basic Building Block (BBB).

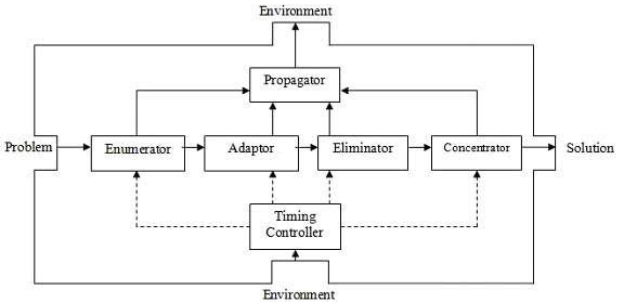


Figure 1 The basic building block BBB

Problem and solution are both functions of time, thus we can represent the time function for problem as $x(t)_{problem}$, and the time function for solution as $y(t)_{solution}$. The timing controller is also a time function timing-control(t). For the two-decision-cycle SIS, the basic building block BBB can be expressed as follows:

```
if timing-control(t) == 'slow'
then /* timing-control(t) is 'slow' */
       y(t)solution = gconcentrate (geliminate (gadapt
(genumerate(x(t)problem))))
else  /* timing-control(t) is not 'slow' */
       y(t)solution = fconcentrate (feliminate (fadapt
(fenumerate(x(t)problem))))
```

where $g_{enumerate}$, g_{adapt}, $g_{eliminate}$, and $g_{concentrate}$ are the transform functions for enumeration, adaptation, elimination and concentration respectively during slow decision cycles, and $f_{enumerate}$, f_{adapt}, $f_{eliminate}$, and $f_{concentrate}$ are the transform functions for enumeration, adaptation, elimination and concentration respectively during quick decision cycles. An Advanced Building Block can be a stand-alone system as shown in Figure 2. The major difference between an ABB and a BBB is the inclusion of a knowledge base, further improving the SIS's problem solving abilities.

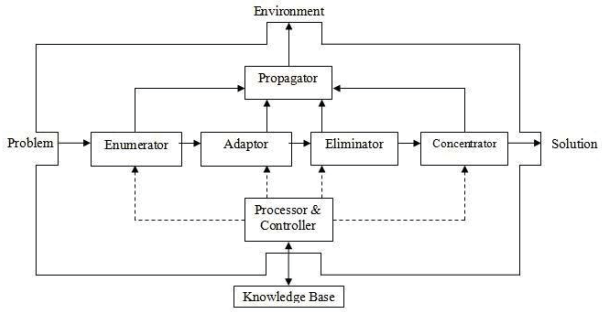


Figure 2 The advanced building block ABB

Just the introduction of the knowledge base justifies the introduction of ontologies. In fact by the use of ontology each SIS can share knowledge with the other peer or experts.
In the domain of Computer Science, in fact, ontology is a powerful tool for representing a knowledge base and establishing formal relationships among the concepts belonging it. Thanks to the ontological formalism it is easy to introduce some processes of automated reasoning or establish explicit formal vocabulary to be shared among applications. The term 'ontology' was

first used in the computer science field by Gruber who used the term to refer to an explicit specification of a conceptualization [13]. Also this kind of definition is still satisfactory for each field where ontology can be applied and so perhaps a good practical definition would be this: "an ontology is a method of representing items of knowledge (ideas, facts, things) in a way that defines the relationships and classification of concepts within a specified domain of knowledge" [14]. Following this point of view, ontologies are "content theories", since their principal contribution lies in identifying specific classes of objects and the relations that exist in some knowledge domains [15]. Commonly ontology is defined as $O = \{C, A, H, R_T\}$ where:

- C is the concept set. $c \in C$ expresses one concept and in each ontology there is ever a root concept marked as "Thing". In particular for each $c \in C$ there exist a descendant nodes set (C_{DN}) containing all its under layer concepts and an ancestry nodes set (C_{AN}) containing all upper layer concepts
- A is the concept attributes set. For $c \in C$ its attributes set is expressed as $A_C = \{a_1, \ldots, a_n\}$ where n expresses the number of attributes related to *c*
- H expresses the concept hierarchy set. The formalism (c_i,c_j) means that c_i is the sub-concept of c_j. In other words this set contains the *is_a* relations among the classes.
- R_T is the set of semantic relations type. $R_T = R_{TD} \bigcup R_{TU}$. R_{TD} means the set of predefined relation (*same_as, disjoint_with, equivalent*) while R_{TU} means the set of user defined relation type. The formalism (c_i,c_j, r) with $r \in R_T$ means that between c_i and c_j there is the r relation. The set $RelR_T(c_i,c_j)$ contains the relation r between c_i and c_j

In general ontologies can be classified into lightweight and heavyweight ontologies [16]. Lightweight ontologies include concepts, simple relationships between concepts (such as specialization "is_a") and properties that describes concepts. Heavyweight ontologies add axioms and constraints to lightweight ontologies. Axioms and constraints clarify the intended meaning of the terms gathered in the ontology. In the case of a Lightweight Ontology there are not R_T set and is defined as: $O_L = \{C, A, H\}$. The lightweight ontology is a very simple and basic representation of a knowledge domain and contains a general statement of a problem. Obviously, it could be complicated by the insertion of new non-hierarchical relationships that enrich its semantic expressivity. In this case, a more mathematically rigorous representation is needed to provide greater confidence that the real meaning behind terms coming from different systems is the same. Heavyweight ontologies are extensively axiomatized and thus represent ontological commitment explicitly. Axioms help to exclude terminological and conceptual ambiguities due to unintended interpretations: the heavyweight ontologies add axioms, well-formed formulas in a formal language, and constraints to lightweight ontologies in order to clarify the intended meaning of the terms gathered on the ontology. Therefore, a heavyweight ontology is defined by $O_H = \{C, A, H, R_T, R, A_X\}$ where A_X indicates the axioms that are in the ontology. Lightweight and heavyweight ontologies do not represent the good solution for a Slow Intelligent System. Lightweight ontologies, in fact, propose a simple view of the domain and do not allow a detailed description of the interactions among all the components in the domain. Using a lightweight ontology, users or systems can only share a very small and simple description of a domain that might not be useful for the full resolution of a problem. On the other hand, using a heavyweight ontology, users can create and share very complex domains, but it is very easy to have problems defining axioms as well as managing the ontology. To address these problems, a solution could be the introduction of a lightweight plus ontology defined as $O_{L+} = \{C, A, H, R_T, R\}$. By introducing non-hierarchical relations, a lightweight plus ontology is more complex and semantically richer than the lightweight ontology, but it is not complex as heavyweight ontology because there are no axioms to consider. The starting point of this approach is the idela that it is reasonable to think that between the lightweight and the heavyweight ontology there are a series of ontologies that can be defined as lightweight-plus ontologies. Some user requests can be supported by the use of one of these lightweight plus ontologies. Each of these ontologies enriches its semantic level by the introduction of new classes, relations, functions, formal axioms and instances. These enrichments allow a better configuration of the system and the satisfaction of user requests. The lightweight plus ontology will be the starting point for the definition of a Slow Intelligent Intrusion Prevention and Detection System.

IV A Slow Intelligence Intrusion Prevention and Detection System

As previously said the aim of this paper is the introduction of an Intelligent Intrusion Prevention and Detection System based on the Slow Intelligence approach.

First of all t he operative scenario has to be set. T he Slow Intelligence Intrusion Prevention and Detection System analyses all traffic incoming and out coming from a local area n etwork. In particular it can signal anomalous events to network administrator by a web based application and react, acco rding to well defined strategies, to th e intrusions. In order to co llect information about the traffic the SNORT framework has been adopted. In particular it is i mportant to underline that SNORT simply collects the traffic and send it, by the use o f our developed module, to the Slow Intelligence System Server that infers the action to adopt. In each h ost of the network there is a S low Intelligence Agent able to execute the action inferred by the Local Server. Each Local Server can interact with other local servers that control other Local Area Networks. In this scenario a fundamental role is played by ontologies. Each local ser ver works on own set of ontologies that can differ. In particular it is n ecessary to introduce and define the following ontologies:

- $O_{Attacks}$ = { $C_{Attacks}$, $A_{Attacks}$, $H_{SAttacks}$, $R_{TAttack}$, R_{Attack}}. This ontology aims to define the attacks that the system can manage

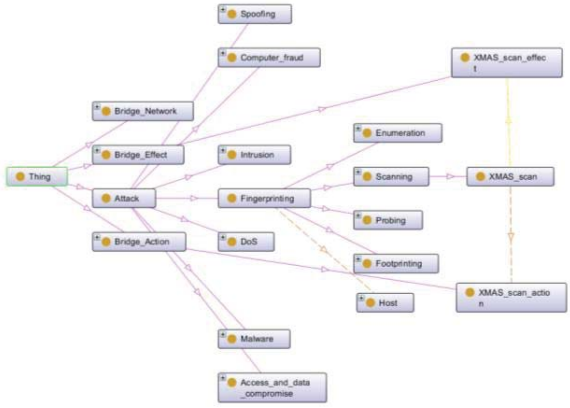


Figure 3 An example of the ontology of attacks in the case of the x-mas scan intrusion

- $O_{Effects}$ = {$C_{Effects}$, $A_{Effects}$, $H_{Effects}$, $R_{TEffects}$, $R_{effects}$}. This ontology describes the effects caused by the intrusions

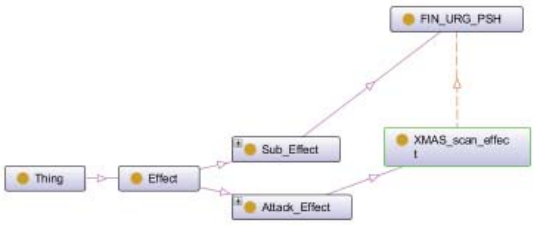


Figure 4 An example of the ontology of effects in the case of the x-mas scan intrusion

- $O_{Actions}$ = {$C_{Actions}$, $A_{Actions}$, $H_{Actions}$, $R_{TActions}$, $R_{Actions}$}. This ontology aims to identify t he actions to be taken in order to recover from fault situations

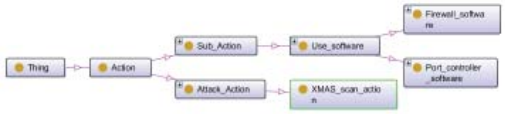


Figure 5 An example of the ontology of actions in the case of the x-mas scan intrusion

- $O_{Environment}$ = {$C_{Environment}$, $A_{Environment}$, $H_{Environment}$, $R_{TEnvironment}$, $R_{Environment}$}. This ontology describes the LAN and t he operative scenario where it works

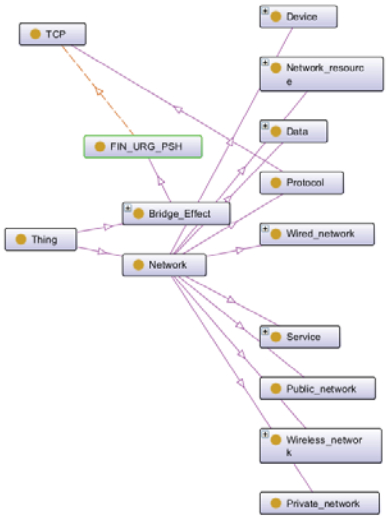


Figure 6 An example of the ontology of the environment in the case of the x-mas scan intrusion

In order to allow the communication among the various hosts and servers that are i n the various LAN the following messages have to be introduced:

- **M_{SlCH}({Action})** = th is message, sent by local server to the host agent, contains the actions that the host have to implement for the resolution of the highlighted fault.
- **M_{SlcSlc}({Effects})** = th is message, sent b y local server to t he other local servers, contains the symptoms that have been retrieved in the network data stream

The local serv er has to i mplement the following functions:

$O'_{Attacks}$ = f(Effects, $O_{Effects}$, $O_{Attacks}$) = this function aims to build the ontology of probable attacks from the analysis of the effects retrieved in the network.

$O''_{Attacks}$ = g(k(M_{SlcSlc}(Effects), $O'_{Attacks}$) = th is function aims to obtain the ontology of the actions that

could be adopted in order to detect and prevent the intrusion related to the anomalies detected in $\mathbf{O'_{Attacks}}$. The function $\mathbf{k(M_{SlcSlc}(Effects)}$ gives the ontology of attacks retrieved by the other local servers. As previously said each local server works on a different set of ontologies that can evolve over the time, and so can retrieve different actions for solving the intrusion

$\mathbf{O'_{Action} = h(O''_{Attack}, O_{Action}, O_{Environment})}$ = this function calculates the ontology of possible actions that the local server can apply or suggest to host in order to solve the intrusion problem

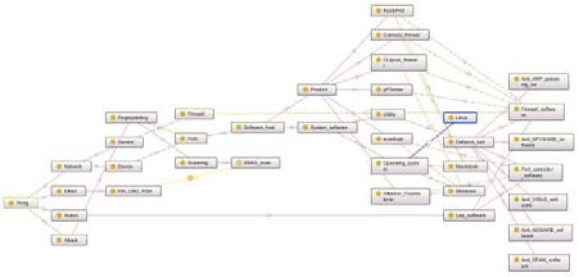


Figure 7 An example of the $\mathbf{O'_{Action}}$

These functions can be considered as the enumeration, propagation and adaptation phases of the Slow Intelligent Approach. At the end of these three phases the actions has to be determined in the elimination phase:

$\mathbf{\{Action\} = t(O'_{Action})}$ = this is the set of actions that the local server, or the host involved in the fault, must adopt in order to solve the intrusion. This set is obtained by the use of Bayesian network. In particular the methodology developed in [17] has been adopted. The ontology $\mathbf{O'_{Action}}$ is mapped in a Bayesian Network, which is trained by a dataset previously built, and the action with the maximum probability will be selected. If there two or more actions have the same probability one of them will be selected in a randomatic way by the use of Montecarlo method. This phase is the Elimination while the Concentration phase is the implementation by the host and/or the local server of the retrieved action. So the operational workflow can be so summarized:

- Step 1: an anomalous data stream is retrieved in the network by the local server
- Step 2: The local server tries to identify the problem through analysis of various ontologies.
- Step 3 If the local server can identify the problem it generates the solutions and the actions that have to be adopted.
- Step 3.1 The hosts and/or the local server get the actions and put them into practice
- Step 4 If the local server does not identify the problem sends the report to the other local servers.
- Step 5 The other local servers after receiving the message attempt to determine the possible actions and then send them, together to the part of ontologies needed, to the local server.
- Step 6 the local server selects the most effective actions and sends them to the various hosts involved in the attempt of intrusion.
- Step 6.1 The local server and the hosts get the actions and put them into practice
- Step 6.2 The local server's ontologies have to be updated with the information retrieved by the other local servers
- Step 6.3 If no action has been retrieved an alarm has to be sent to the LAN administrator

It is important to underline that local server ontologies are continuously updated: in fact if a local server is not able to manage an intrusion will ask to the other local servers that are working in other contexts. So it collects not only the actions to apply in order to manage the intrusion but also the parts of ontologies that need for recognize the intrusion. So in this way the next time the local server will be able to manage, without helps, this kind of intrusion. In the same way when in an ontology new concepts will be introduced, for example by an expert in order to manage new intrusions, they will be shared with the other local servers.

V Experiment setup and results

This section describes a first prototype of the IDPS. The proposed prototype has to satisfy and offer the services previously described according to a SIS approach. Figure 8 depicts the experimental scenario.

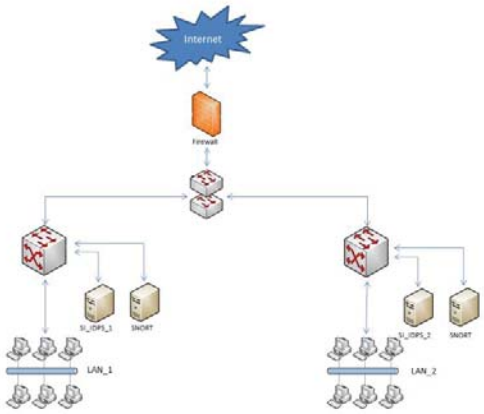


Figure 8 Experimental Scenario

There are two local servers that intercept the traffic related to two local area networks that are used as students' laboratories. Each local network is composed by 40 hosts that can surf on the net. In order to monitoring the packets that flow in the two networks for each of them two span ports has been created. The Slow Intelligence IDPS works on one them while an installation of Snort works on the second one. In this way a comparison between the two approaches can be conducted. The first step was the mapping of SNORT ruleset in the various ontologies. This task has been accomplished by the support of some experts and at the end of this phase $O_{Effects}$, $O_{Attacks}$, $O_{Actions}$, $O_{Environment}$ have been achieved. At the same time a ground truth has been created in order to train the Bayesian Network representative of the ontologies. In particular data from one month observation period have been collected by the use of SNORT on the two networks. At the end of this period data on symptoms of the attacks and relative resolution actions have been collected and used in order to train the various Bayesian networks. At the end of this phase the real experimentation started. At the beginning of the experimentation each local server had the 50% of the concepts and relationships of the various ontologies. The comparison between SNORT e and the Slow Intelligence IDPS has been conducted in terms of precision and recall. The two networks has been tested from the point of view of intrusion by the use of the BackTrack tool [18]. BackTrack is a Linux-based penetration testing arsenal that aids security professionals in the ability to perform assessments in a purely native environment dedicated to hacking. In particular for the first network 1500 intrusions has been attempted while for the second one the intrusions were 2500. The hosts of the two networks can surf on internet but their traffic has been continuously monitored by a firewall in order to verify that no external intrusions have been attempted. The obtained results are depicted in table 1.

	True Positive	False Positive	False Negative	Precision	Recall
SIIDPS#1	1378	78	122	0,946	0,918
SNORT#1	1289	201	211	0,865	0,859
SIIDPS#2	2298	143	202	0,941	0,919
SNORT#2	2167	278	333	0,886	0,866

Table 1 Obtained results

The obtained results are good: in particular the number of false positive and false negative is very low compared with the results showed by SNORT. Another parameter to monitor is the number of concepts that ontologies of the two systems have at end of the experimentation. As previously said each ontologies related to the various local servers at the beginning had only the 50% of the concepts and relationships of the full ontologies. In table 2 the final status of the ontologies has been depicted.

	Concepts	Relationship
$O_{Attacks}$	134	445
SIIDPS#1	128	401
SIIDPS#2	118	393
$O_{Effects}$	198	541
SIIDPS#1	176	428
SIIDPS#2	158	465
$O_{Actions}$	128	297
SIIDPS#1	107	238
SIIDPS#2	116	245

Table 2 SIIDPS Ontologies Final Status

Also in this case the results are good: the two systems share knowledge and at the end of the experimentation phase show the same knowledge base.

Conclusions

In this paper a novel approach for Intrusion Detection and Prevention System has been introduced. This approach is based on the use of the ontology formalism and of Slow Intelligence approach. It has been tested in an operative scenario and the first experimental results seems to be good. The future work aims to improve the system by the use of new approaches in the elimination phase, the introduction of more sophisticated artificial intelligence approach in the inference phase and in a more detailed experimentation also on standard datasets.

References

[1] WebCohort, "Only 10% of Web application are secured against common hacking techniques", http://www.imperva.com

[2] "NIST – Guide to Intrusion Detection and Prevention Systems (IDPS)". 2007-02. Retrieved

[3] S. Bellovin, R. Bush, "Configuration Management and Security", IEEE Journal on Selected Areas in Communication JSAC, vol. 27, no.3, 2009

[4] L. Coppolino, S. D'Antonio, M. Esposito, L. Romano, "Exploiting diversity and correlation to improve the performance of intrusion detection systems", proceedings of Network and Service Security. International Conference on, pp. 1-5, 2009

[5] V. Raskin, C. Helpenmann, K. Triezenberg, S. Nirenburg, "Ontology in information security: a useful theoretical foundation and methodological tool", New Security Paradigms Workshop, ACM Press, pp. 53-59, Cloudcroft, NM, 2001

[6] J. Undercoffer, A. Joshi, J. Pinkston, "Modeling Computer Attacks: an Ontology for Intrusion Detection", RAID 2003, LNCS 2820, pp 113-135, 2003

[7] SNORT http://www.snort.org/

[8] G. Denker, L. Kagal, T. Finin, M. Paolucci, K. Sycara, "Security for DAML Web Services:

Annotation and Matchmaking", The semantic Web (ISWC 2003), LNCS 2870, Springer, pp. 3 35-350, 2003
[9] J. Undercoffer, A. Joshi, T. Finin, John Pinkston, "A target centric ontology for intrusion detection: using DAML+OIL to classify intuitive behaviors", Knowledge Engineering Review, Cambridge University Press, pp. 23-29, 2004
[10] F. Abdoly, M. Kahani, "Ontology-based Distributed Intrusion Detection System", Proceedings of the 14th International CSI Computer Conference, IEEE, pp. 65-70, 2009
[11] A. Razzaq, A. Hur, H. Farooq Ahmed, N. Haider, "Ontology based Application Level Intrusion Detection System by Using Bayesian Filter", Computer, Control and Communication, 2009. IC4 2009. 2nd International Conference on, pp. 1-6, 2009
[12] Shi-Kuo Chang, "A General Framework for Slow Intelligence Systems", International Journal of Software Engineering and Knowledge Engineering, Volume 20, Number 1, February 2010, pp. 1-15.
[13] Gruber, T.R, Translation approach to portable ontology specification, Knowledge Acquisition, v. 5, pp. 199-220, 93
[14] Jepsen, T., Just What Is an Ontology, Anyway?, IT Professional, vol. 11, no. 5, pp. 22-27, Sep./Oct. 2009.
[15] Maedche A., Staab S., Ontology Learning for the Semantic Web, IEEE Intelligent Systems, vol. 16 no. 2, Mar/Apr 2001, Page(s): 72-79.
[16] Corcho, O., A Layered Declarative Approach to Ontology Translation with Knowledge Preservation, Volume 116 Frontiers in Artificial Intelligence and Applications, 2005.
[17] F. Colace, M. De Santo, M. Vento, P. Foggia, "A Bayesian Approach for Automatic Building Lightweight Ontologies for E-Learning Environment", ICEIS (2) 2005: 386-389
[18] http://www.backtrack-linux.org/

Cloud Services for Medical Image Processing

Application of Service Science and Slow Intelligence Systems

Kang-Lin Peng
Institute of Service Industries and Management
Minghsin University of Science and Technology
Institute of Computer Science and Engineering
National Chiao Tung University
Hsinchu, Tsiwan
kp@must.edu.tw

Shi-Kuo Chang
Department of Computer Science
University of Pittsburgh
Pittsburgh, U.S.A.
chang@cs.pitt.edu

Yu-Chee Tseng
College of Computer Science
National Chiao Tung Univeristy,
Hsinchu Taiwan
yctseng@cs.nctu.edu.tw

Thomas M. Deserno
Department of Medical Informatics
RWTH Aachen University,
Aachen, Germany
deserno@ieee.org

***Abstract*— Cloud services could be an effective solution for future medical image processing for computer-aided diagnosis (CAD). This research applies the concept of service science and Slow Intelligence System to enhance the value of medical image processing. A team from Germany, Taiwan and the US will cooperate to construct and improve the techniques and quality of cloud services for medical image processing for CAD. The goals are to design a system of cloud services for medical image processing, which could enhance the effectiveness of diagnosis and provide people with better medical care as well as to explore a concept of the Service Intelligence from this research.**

Keywords- cloud services; medical image processing; slow intelligence systems; service science; service intelligence

I. INTRODUCTION

Any image processing technique that can speed up the reading process tends to be viewed more positively than a method that does not change, or lengthens, interpretation time. Image processing techniques that can automate routine but manual and time consuming tasks, would benefit [1]. Cloud services could be a solution for future medical image processing (MIP) for CAD. As cloud computing grows to span more and more globally distributed data-centers, information service providers are involved actively in the development of "cloud services." However, cloud services for medical imaging processing are limited but essential to medical industries. Through the development of cloud concepts and techniques, this research is going to uncover essentials of cloud services for medical image processing. According to the concepts of Slow Intelligence Systems (SIS), this research proposes to integrate existed services and systems to plug-in of the medical imaging system ImageJ into the SIS. ImageJ is open source system developed by the National Institutes of Health (NIH), USA and has multiple image processing and analytic functions primarily used for medical and biology-related imaging. It is the most used software in medical image processing and is still being expanded by developers around the world. This research proposes applyiing the SIS as a primary structure in cloud services for medical image processing to actively support the application of CAD for improved medical services.

The purpose of this research is to allow instantaneous access to imaging data, advanced reformatting and viewing tools from any location across the globe, as well as powerful computer processing in a cloud environment. It eliminates costly dedicated storage servers and workstations, while providing interactive image viewing and advanced visualizations anytime, anywhere. The challenge to systematic service innovation in medical image processing is the interdisciplinary nature of the service, integrating technology, business, social, and client (demand) innovations [2]. Principles of service science, cloud computing and the SIS will help providing cloud services for medical image processing in medical routine and health care (Fig. 1).

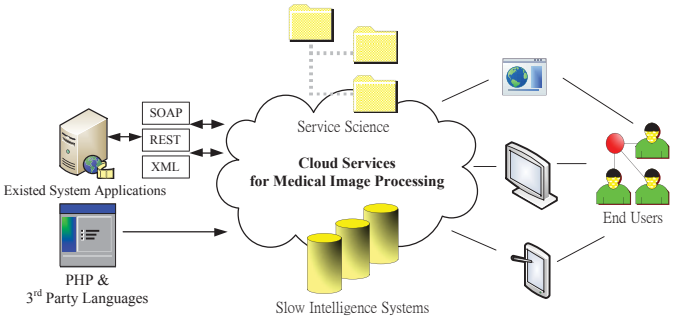


Figure 1. Cloud Services for Medical Image Processing

II. CLOUD SERVICES FOR MEDICAL IMAGE PROCESSING

A. Services Gap on Medical Image Processing

Traditional health care IT stores and manages image archives onsite using the internal hospital network that is protected by a firewall from the outside. However, the growing long-term cost of managing an onsite medical image repository has been a subject, which the health care industry struggles with [3]. In addition, services demanded by doctors for CAD systems require high scalability, accessibility, availability and disaster recoverability as well. What have been done through the development of medical image processing are the standard of the Digital Imaging and Communications in Medicine (DICOM), the Picture Archiving and Communication System (PACS), and the internet platform of open source ImageJ. Some techniques of medical image processing through the DICOM, the PACS and the NIH system have been improved for CAD. However, service gaps from the perspectives of cloud services for medical image processing have to be discussed because of emerging information technology and exis ted systems. How could we integrate old and new systems working on cloud services for medical image processing? This research aims to design an integrated cloud services system by combining techniques of medical image processing for CAD via the SIS. It is also expected that patients could use cloud services to understand, learn and care their own health through access to the results of medical image processing. Principles of service science, management, engineering and design (SSMED), techniques of cloud computing and application of the SIS are discussed based on actor network theory so as to closing the service gap (Fig.2).

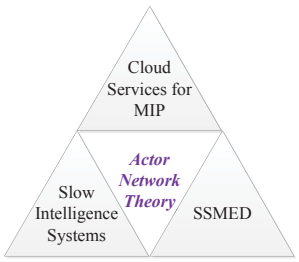


Figure 2. Research Framework

B. Actor Network Theory

Actor network theory tries to explain how material–semiotic networks come together to act as a whole; the clusters of actors involved in creating meaning are both human and nonhuman[4]. As a part of this it may look at explicit strategies for relating different entities together into a network so that they form an app arently coherent whole. These networks are potentially transient, existing in a constant making and re-making. This means that relations need to be repeatedly "performed" or the network will dissolve. Networks of relations are not intrinsically coherent, and may indeed contain conflicts. Social relations, in other words, are only ever in process, and must be performed continuously.

C. Service Science

Service science is the study of service systems, which aims to create a basis for systematic service innovation [5]. Service science is an abbreviation of service science, management, engineering and design (SSMED), which integrates diverse fields with an interdisciplinary approach aiming to study the service phenomena occurring in human society, and develop service systems toward a better society. Such an approach requires the collaboration of different disciplines, as well as of the government, academia, and enterprises for the purpose of service innovation. At the heart of service science is the transfer and sharing of resources within and among service systems. Four categories of resources have been noted and examined: (1) resources with rights, (2) resources as property, (3) physical entities, and (4) socially constructed entities [5]. While researchers and policy makers are still actively studying the security, privacy, and liability issues involving sensitive medical information in the cloud [3], this research aims at applying service science to cloud services for medical image processing.

D. Cloud Computing and Cloud Services

A cloud is a type of parallel and distributed system connected and virtualized computers that are dynamically provisioned and presented as one or more unified computing resources based on service level agreements established through negotiation between the service provider and consumer[6]. Cloud computing is a l arge-scale distributed computing paradigm that is d riven by economies of scale, in which a p ool of abstracted, virtualized, dynamically-scalable, managed computing power, storage, platforms, and services are delivered on demand to external customers over the internet [7] and hand held devices. Cloud computing is a pop ular topic on software and distributed computing based on Internet, which means users can access storages and applications from remote servers by web browsers or other fixed or mobile terminals. Because the constrained resources of fixed or mobile terminals, cloud computing will provide terminals with powerful complementation resources to acquire complicated services. For medical image processing systems, the cloud computing platform has to provide a long term offsite medical image archive solution [8]. New computational technologies allow users to "see" into and understand human's bodies with unprecedented depth and detail. As a result of these advances, biomedical computing and visualization help to produce exciting new biomedical scientific discoveries and clinical treatments [9].

From the beginning of the cloud concept till now, cloud services have been connected to concepts such as Software as a Service (Saas), Platform as a Service (Paas), Infrastructure as a Service (Iaas), Xen, Virtual machine (VM), Virtual Data Center (VDC), Virtual Cluster (VC), Microsoft Azure, Acadia-virtual computing environment coalition, etc. This research aims to explore these concepts apply to the SIS on the principle of SSMED for medical image processing so as to offer a better solution for CAD. Cloud services have advantages as a service platform such as flexibility in dynamic provisioning through virtualization. For clinicians who usually query and view images, an online portal provides a simple way to create and manage services [3]. The normative function of service systems is to connect people, technology and information through value propositions with the aim of co-creating value for the service systems participating in the exchange of resources within and across medical imaging systems [12].

E. Slow Intelligence Systems

The SIS technology is a novel technology for the design and improvement of complex information systems. Contrary to popular belief, not all intelligent systems have Quick Intelligence. There are a surprisingly large number of

intelligent systems, quasi-intelligent systems and sem i-intelligent systems that have Slow Intelligence. Such Slow Intelligence Systems are often neglected in mainstream research on intelligent systems, but they are really worthy of our attention and emulation. There are a large number of intelligent systems, quasi-intelligent systems and sem i-intelligent systems that are "slow". Distributed intelligence systems, multiple agents systems and em ergency management systems are m ostly slow intelligence systems that exhibit the characteristics of multiple decision cycles. The SIS are general purpose systems characterized by being able to improve performance over time through a process involving enumeration, propagation, adaptation, elimination and concentration as Fig.3 [10].

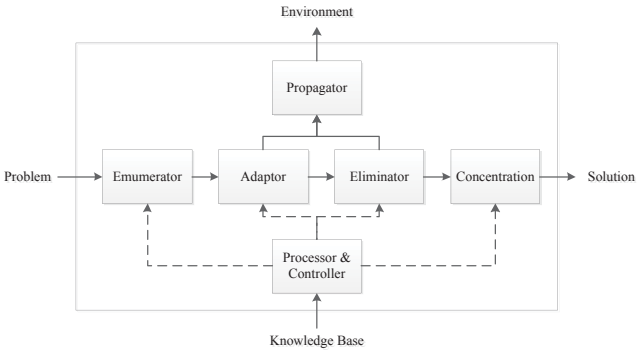


Figure 3. Slow Intelligence Systems

The concept underlying is akin to the Chinese philosophy of I-Ching (book of changes), which centered on the ideas of the dynamic balance of opposites, the evolution of events as a process, and acceptance of the inevitability of change. It is to say that complex changes can be systematic in simplified changes. Unification of changes can be done when all simplified changes merge together. Based on this concept, SIS emerge slow and quick intelligence systems together which lead to improvements in complex information systems and thus meet the dynamic reality of IT development with a m ore balanced approach. Hybrid Intelligence offers great opportunities we have to harness this data availability to build systems of immense potential[11]. (Fig. 4)

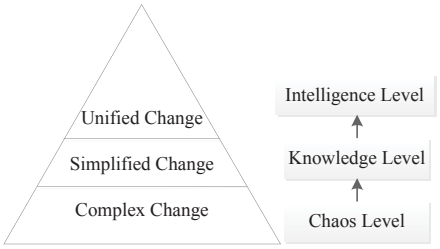


Figure 4. Levels of Change

III. METHODOLOGY

A. System Design

Health organizations seek to increase operational efficiency and improve care and services, cloud computing provides a way to sim plify information and communications technology (ICT) management, control costs, and better meet user ne eds. A trusted, comprehensive set of cloud services is essential to meet organizational unique needs now and in the future. Health organizations need flexible options that allow service providers to move what they want to the cloud, how and when they need to.This research examines the SIS of existing systems of cloud computing for medical image processing within the principle of SSMED. The approach taken is a thematic one, focusing on the mapping between the effectiveness of CAD and the interdisciplinary SSMED, in which an in tegrated platform of the SIS for medical image processing could be built. A cloud services system such as Microsoft Azure could be chosen to facilitate the concept of the research topic. Microsoft Azure's elastic, interoperable infrastructure and familiar tools can be used in the cloud (public and private), on-premises, or as a combination of both. At the same time, service providers can have the power of the cloud without compromise to security, privacy, and reliability.

B. Cases Application

We propose a collaboration between NCTU and NTU hospital Hsinchu branch to implement this project. In addition, in Taiwan, small and medium dental clinics provide dental care services for people. One problem is that not every dental clinic can afford expensive medical imaging equipment forapplications care, such as dental implant. In many cases, dentists need to send patients to radiology centers for the dental imaging. After dental clinics get the patients‘ dental images from the radiology center, patients return for dental care. This is all very time consuming for both dentists and patients to do the dental care in such cases. Our project could build a cloud services system for dental image processing for both dental clinics and rediology centers, and thus provide a solution for the problem mentioned above.

IV. EXPECTED RESULTS OF CLOUD SYSTEMS

The goal of service science is to increase the productivity of the service industry, promote innovation, and create greater validity and transparency when assessing the value of investments in services [12]. This project aims to apply the principle of service science, the techniques of cloud computing and the application of SIS to enhance the value of cloud services for medical image processing. With medical images stored, cached and serviced remotely in a cl oud, the system should ensure safe, long-term archiving, higher availability of image data and patient information, as well as serve as an image data clearing house for communication with offsite radiologists, clinicians, surgeons and others who need processed images or raw data for independent manipulation. The application also provides an image communication engine for patient electronic health record The cloud services architecture is designed to be simple and includes three main results:

A. The cloud up-loader-

It is a thin downloadable software module which could receive DICOM data from the local DICOM network and transfer it quickly and securely to the cloud services.

- Secure and efficient (compression/streaming) data transfer.
- No need for VPN (Upload using HTTPS protocol)

B. The cloud client

It is a thin downloadable software suite which can manage DICOM data locally or connect to remote sources to receive/send DICOM image data. By using ordinary PC, the cloud client should enable viewing and analysis with an

advanced multi-modality viewer and dedicated state of the clinical applications.

- Easy-to-use study list with relevant filters and DICOM functions.
- Intuitive image printing tool and 2D multi-monitor image pre-viewer.
- Advanced general medical image viewer for diagnostic quality review and analysis of CT, MRI, PET, ultrasound, and multi-modal resources of medical imaging.
- Advanced tools for medical image processing.
- Available also as stand-alone applications with easy PACS integration.

C. *The cloud services*

It is an online unlimited pool of managed repositories and processing engines for medical images. The cloud receives images from multiple sources (up-loaders), processes them if needed and archives them based on configurable rules. The cloud provides fast and secure image access to radiologists, referring physicians, and patients from any location across the globe.

- Provides service quality with a high level of image data and operational redundancy.
- Rule-based image data access for viewing and/or retrieving images.
- Auto batching and image pre-processing tasks.
- Unlimited storage capacity and computing power.
- Be installed on an enterprise server or a private data-center.

V. EXTECTED RESULTS OF SERVICE INTELLIGENCE

SIS continuously learns, searches for new solutions as well as propagates and shares its experience with other peers[10]. Service science is akin to SIS looking for new solutions for services on demand from the knowledge base of Management, Engineering, and Design, thus service intelligence can accumulate rapidly in an interdisciplinary knowledge, which we called it Service Intelligence.

Service intelligence can be service science term used to describe applications and technologies which are used to gather, provide access to as well as analyze data and information about services, to help design better service model and facilitate service innovation. In this research we apply SIS to cloud services for medical image processing that makes possible a seamless integration of SSMED into coherent service intelligence, thus enabling simplified service delivery and high service quality.

We will adapt SIS to an upcoming cloud services project, which is conducted at the southern Industrial Technology Research Institute in Taiwan, to promote the value co-creation capability of service intelligence for a better life.

ACKNOWLEDGMENT

We acknowledge the interdisciplinary research team from Taiwan, Germany and USA to develop and improve the techniques and quality of cloud services on medical image processing for CA D. Dr. Peng visited the Department of Medical Informatics at RWTH, Germany to d evelop the technology basis of medical image processing with funding of the DAAD/NSC summer institute program in 2011. Professor Deserno was the host of this program. Professor Chang invites Dr. Peng to the US for collaboration in applying SIS to this project. Professor Tseng will host collaboration between NCTU and NTU hospital Hsinchu branch which can sustain this project.

Dr. Peng of Mingshing University of Science and Technology (MUST) is a professional in the field of service science. Professor Chang is a professional in the field of intelligence systems. Professor Tseng of National Chiao Tung University (NCTU) is a professional in the field of cloud computing. Professor Deserno of RWTH Aachen University is a professional in the field of medical image processing. Our research team, with diverse expertise in service science, cloud computing, medical image processing and intelligence systems, offers the synergy that is critically relevant to this research topic, as well as being essential to a long term research project.

REFERENCES

[1] T. M. Deserno, *Biomedical Image Processing*. New York: Springer, 2011.

[2] J. Spohrer and P. P. Maglio, "The emergence of service science: Toward systematic service innovations to accelerate co-creation of value," *Production and Operations Magagement,* vol. 17, pp. 238-246, 2008.

[3] C. C. Teng, J. Mitchell, C. Walker, A. Swan, C. Davila, D. Howard, and T. Needham, "A medical image archive solution in the coud," *2010 IEEE,* pp. 431-434, 2010.

[4] Wikipedia. (2012). *Actor-network theory*. Available: http://en.wikipedia.org/wiki/Actor%E2%80%93network_theory

[5] P. P. Maglio and J. Spohrer, "Fundamentals of Service Science," *Journal of The Academy of Marketing Science,* vol. 36, pp. 18-20, 2008.

[6] R. Buyya, C. S. Yeo, and S. Venugopa, "Market-oriented cloud computing: Vision, hype, and reality for delivering it's services as computing utilities," presented at the The 10th IEEE International Conference on High Performance Computing and Communications, Los Alamitos, CA, USA, 2008.

[7] I. Foster, Y. Zhao, I. Raicu, and S. Lu, "Cloud computing and grid computing 360-degree compared," in *Grid Computing Environments Workshop*, Austin, TX, 2008, pp. 1-10.

[8] W. Zeng, Y. Zhao, and J. Zeng, "Cloud service and service selection algorithm research," in *2009 World Summit on Genetic and Evolutionary Computation*, Shanghai, China, 2009.

[9] C. R. Johnson and D. M. Weinstein, "Biomedical computing and visualization," in *ACSC 2006 Proceedings of the 29th Australasian Computer Science Conference* Darlinghurst, Australia, 2006.

[10] S.-K. Chang, "A general framework for Slow Intelligence Systems," *International Journal of Software Engineering and Knowledge Engineering,* vol. 20, pp. 1-15, 2010.

[11] A. Z. Spector, "Prodigious data, logic, processing, and usage," in *The 22nd International Conference on Software Engineering and Knowledge Engineering*, San Francisco Bay,CA, USA, 2010.

[12] T. Abe, "What is service science " in *FRI Research Report*. vol. 246, ed: Fujitsu Research Institute, 2005.

A Slow Intelligence Framework for Human Behavior Recognition Using Wearable Cameras

Wen-Hui Chen
Graduate Institute of Automation Technology
National Taipei University of Technology, Taipei, Taiwan
whchen@ntut.edu.tw

Shi-Kuo Chang
Department of Computer Science
University of Pittsburgh, Pittsburgh, PA 15260, USA
chang@cs.pitt.edu

***Abstract*—This study presents a slow intelligence framework (SIF) for automatically detecting and recognizing human behaviors from the image sequences captured by a wearable camera. The proposed framework comprises a typical process in a slow intelligence system (SIS), including propagation, enumeration, adaptation, elimination, and concentration. The quick decision cycle in the proposed SIF acts as a reflex agent to instantaneously respond on the basis of the current percept, while the slow decision cycle has its ability to update the beliefs by continuously learning from past experience. A potential application of the proposed framework is the development of an early detection system for people with dementia based on long-term observations of their daily activities. Other applications include pet homecare and senior person homecare.**

Keywords- Slow intelligence system, behavior recognition; Software agent.

I. INTRODUCTION

In video surveillance systems, recognition of human behaviors from image sequences has many important applications, including eldercare, security monitoring, and anomaly detection, etc. As a result, various algorithms for human behavior recognition have been proposed in the past. Most of the approaches to behavior recognition have relied on stochastic models because of its nature in high-level interpretation. For example, Oliver *et al*. presented a real-time computer vision approach for recognizing human behaviors in a surveillance system [1]. Liu *et al*. proposed a framework using an interaction embedded hidden Markov model for human behavior understanding in a nursing environment [2]. In [3], the authors utilized a model of stochastic context-free grammar combined with domain-specific information for recognizing multi-task activities involving multiple people from video. Vaswani *et al*. presented a model for abnormal activity detection using a continuous-state hidden Markov model [4]. A more complete survey of human activities recognition from video can be found in [5].

Most of the algorithms developed for human behavior recognition are based on the image sequences captured by a fixed camera installed at a specific location. However, a fixed camera has its blind spots so it is not flexible in practical applications. In order to complement this drawback, PTZ (pan/tilt/zoom) or wide view cameras are used to increase the field of view [6]-[7]. The limitation in using a PTZ camera exists for its slow motion mechanism, restricted rotation and tilt angles. As to the wide view camera such as omni-directional cameras, a correction is required to reconstruct the captured images to a normal perspective view for further analysis.

A wearable camera can be used as visual lifelogging devices. The application of wearable cameras in research has become popular in recent years [8]-[10]. The main features of using wearable cameras include the ease of data collection and being able to perceive information from the wearer's perspective. Due to constant movement, the quality of image sequences gathered from a wearable camera is poor as it can contain cluttered background, too much or less light, and motion blurs, which makes recognition tasks become difficult. How to handle the poor image quality is the main technical challenge of using wearable cameras.

In this study, an intelligence system for behavior recognition is proposed to analyze the content of the image data captured by a wearable camera in a long period of time based on a slow intelligence framework. A slow intelligence system is a new paradigm for building an intelligence system through a process involving enumeration, propagation, adaptation, elimination, and concentration [11].

The remainder of this paper is organized as follows. Section II describes the technical challenges of behavior recognition for elderly with dementia using wearable cameras. Section III provides a brief review of a slow intelligence system and its application to behavior recognition. Section IV describes application to pet homecare and senior person homecare. Conclusions and future work are finally drawn in section V.

II. BEHAVIOR RECOGNITION FOR ELDERLY WITH DEMENTIA

Recognizing human behaviors from wearable cameras has important applications in eldercare. When a senior at the age over sixty-five shows signs of common dementia symptoms such as changes in behaviors, it is recommended to seek a clinical diagnosis and assessment to determine if he or she develops dementia. Early detection of dementia is important as it can help slow down the progress of dementia and provide patients with timely treatment. However, for the elderly who live alone or their family members are all busy working, the

early-stage signs of symptoms are easily ignored or mistaken for signs of aging, missing the best time to receive treatment.

Therefore, the detection of behavior changes via machine recognition is beneficial for the diagnosis of early-stage dementia. This requires long-term observation of individual's daily activities. As a lifelogging device, a wearable camera can reach the goal of recording individual's daily activities in a noninvasive manner. A systematic approach to human behavior recognition is required for analyzing the recorded video to provide the elderly patients with better health care and timely treatment.

Although the recording video contains sufficient information for the detection of anomaly behaviors, the process of performing semantic analysis from the low-level features is complicated as the video can contain poor quality images due to constant movement of the recording device. The major technical challenges include the processing of cluttered background, motion blurs, too much or less light in images, to name a few. Fig. 1 illustrates the problems exist in handling images captured from wearable cameras.

(a) Cluttered background (b) Too much light

(c) Too little light (d) Motion blurs

Fig. 1. Poor quality images captured from wearable cameras. (a) Cluttered background (b) Too much light (c) Too little light (d) Motion blurs

In order to overcome the aforementioned problems, an intelligent system with image processing and advance machine learning techniques is required. The goal of this study is to provide a slow intelligence framework, which will be describe in the following session for the design of an intelligent system that can automatically classify human behaviors from recorded video data.

III. Slow Intelligence Systems for Behavior Recognition

A slow intelligence system (SIS) is a new paradigm for building an intelligent system through a process involving enumeration, propagation, adaptation, elimination, and concentration [11]. Fig. 2 shows a typical building block of a slow intelligence system.

As shown in Fig. 2, an SIS is characterized by some fundamental blocks. In addition, it possesses slow decision cycles to complement quick decision cycles in the inference process to find a solution. The quick decision cycle provides an instantaneous response, while the slow decision cycle tries to follow the gradual changes in the environment and analyze the information acquired by experts and past experiences. The two decision cycles enable the SIS to cope with the changes of the environment and achieve the long-term goal.

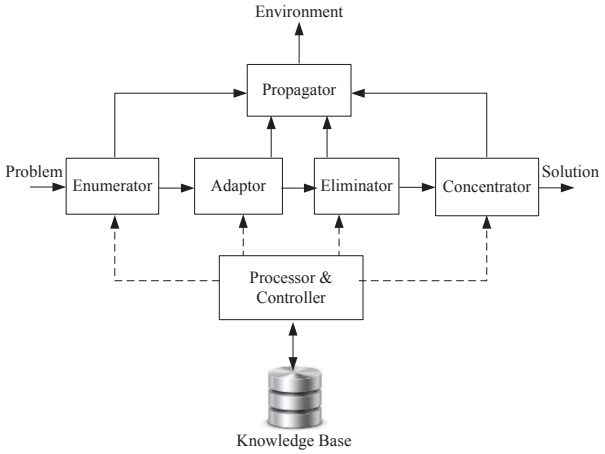


Fig. 2. A typical building block of a slow intelligence system

The proposed SIS framework for human behavior recognition is depicted in Fig. 3. Captured by a wearable camera, the video data is stored in an archive and processed via data processing unit to extract image features for dimensional reduction and further analysis. The proposed structure presents a hierarchical model for recognition. The data processing unit is a module for the operation of low-level features. The action-recognition modules, in a higher level, are implemented in the process of enumeration, propagation, adaptation, elimination, and concentration.

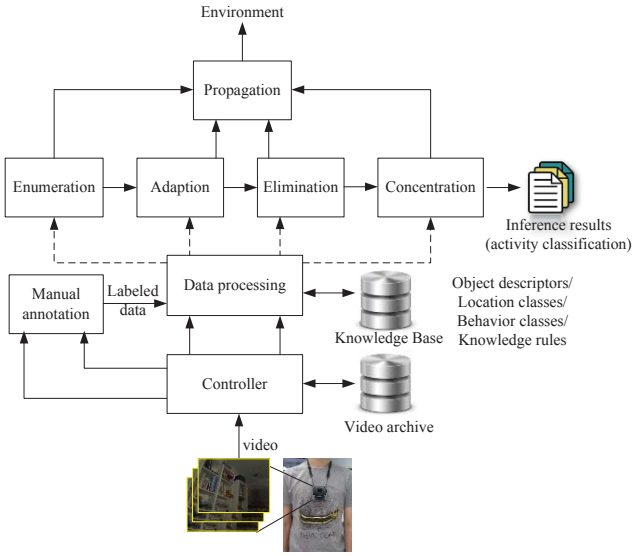


Fig. 3. The proposed slow intelligence framework for behavior recognition

As a slow decision cycle, the training or labeled data derived through manual annotation in the training process is used to adjust the model parameters of a classifier. The task of behavior recognition is performed through the trained classifier as the quick decision cycle.

The proposed SIS observes through input data and acts as an autonomous entity to issue a warning to caregivers or family members for providing timely support service to the elderly when the system detects the change of observed behaviors in the quick decision cycle. The controller in Fig. 3 is responsible for controlling the timing and data path to manual annotation, data processing, and data archive.

The recognition of human behaviors can be considered as the design of multi-class classifiers. There are two classifiers, location classifier and behavior classifier, described in the knowledge base for identifying both the location the individual situated and the behavior the individual performed. Location information provides useful prior knowledge for behavior recognition. For example, if we know that the elder is in the living room, the behavior of watching TV is more probable than that of preparing food.

Object descriptors, such as television, sofa, door handle, computer, etc., defined in the knowledge base are employed to help locate the possible space where the individual situated. In addition, the distance between two different areas are also taken into consideration for localization. In this study, the location classes were defined according to the layout of a residential house, including the living room, bedroom, dining room, kitchen, bathroom, and hallway.

The information of frequency, duration, and interactions with objects are three important indices used to infer human behaviors, which can be reached through the following processes.

1) Enumeration

All possible behaviors defined in the database are enumerated for evaluation. When dealing with behavior recognition tasks, we need to decide how many behaviors are considered, and define them in the knowledge base. Five daily activities (reading, watching TV, preparing food, washing hands, eating) and one abnormal behavior (wandering) were defined as behavior classes in this study.

2) Adaptation

Activities are enumerated in the enumeration stage and adapted in the process of changing to suit the observed inputs. The hidden Markov model (HMM) is well suitable for modeling time series data, and is one of the widely used approaches to behavior recognition. In this study, the HMM is adopted to design a behavior classifier. Six HMMs are constructed for evaluating the most probable activity coincides with the observed inputs. All possible activities will be ranked and adapted to the environment according to HMM inference results.

3) Elimination

As each activity is represented by feature sequences, the recognition of human behaviors from captured video can be considered as a pattern recognition problem. In this stage, unsuitable activities are eliminated to narrow down the search scope of the solution space so that more resources can be kept for the next stage.

4) Concentration

After ruling out unsuitable activities, resources for computations are more concentrated to the suitable solution. Hence, the inference result can be derived through evaluating the activities left. When the individual shows wandering behavior or makes repetitive movements at some point, this anomalous behavior will be detected.

5) Propagation

Acting upon the environment, the propagator in Fig. 3 plays an important role to communicate with other slow intelligence systems working towards achieving their goal.

IV. APPLICATION TO PET HOMECARE

A related application is pet homecare. An experimental system is shown in Fig. 4, where a house pet is monitored by a stationary camera. The pet owner who is away from home at work, can specify the image capture rate (such as one image per minute) and observe the pet's behavior on the web, for example, how frequent does the pet walk to the adjacent room on the left to drink water (or to pee).

With only a stationary camera, the adjacent room is occluded by the wall and therefore the pet's actual behavior must be guessed. If a wearable camera is added to the experimental system, the pet's behavior can be recognized with greater certainty. Using the slow intelligence approach, the experimental system can learn the pet's typical behavior patterns and detect behavioral anomaly when the pet's behavior deviates from the behavior patterns, for example, a sick pet may drink water more often (or less often).

The scenario is as follows. First, the generic behavior patterns are described by a visual grammar. Then the slow intelligence technique is applied to enumerate valid visual sentences and eliminate the unlikely ones. Behavioral anomaly is detected when a visual sentence not in the visual language is encountered. Needless to say, similar techniques may be applicable to senior person homecare.

Fig. 4. A house pet can be monitored by stationary and wearable cameras

V. CONCLUSIONS AND FUTURE WORK

In this paper, we described a framework of classifying behavior patterns from image sequences, captured from a wearable camera, based on a slow intelligence system. As an extension work of this study, more daily life activities and the behaviors associated with dementia symptoms should be included and tested. The recognition accuracy can be improved

by the fusion of audio data and other body worn sensors. A related application is pet homecare, where house pets are monitored by both stationary and wearable cameras. Pets' behavior patterns can be learned using slow intelligence approach. Similar techniques may be applicable to senior person homecare.

REFERENCES

[1] N. M. Oliver, B. Rosario, and A. P. Pentland, "A Bayesian computer vision system for modeling human interactions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no.8, pp. 831-843, August 2000.

[2] C. D. Liu, Y. N. Chung, and P. C. Chung, "An Interactoin-embedded HMM framework for human behavior understanding: with nursing environments as examples," *IEEE Trans. Information Technology in Biomedicine*, vol. 14, no. 5, pp. 1236-1246, September 2010.

[3] D. Moore and I. Essa, "Recognizing multitasked activities from video using stochastic context-free grammar,"

[4] N. Vaswani, A. K. Roy-Chowdhury, and R. Chellappa, "Shape activity: a continuous-state HMM for moving/deforming shapes with application to abnormal activity detection," *IEEE Trans. Image Processing*, vol. 14, no.10, pp.1603-1616, Oct. 2005.

[5] P. Turaga, R. Chellapa, V. S. Subrahmanian, and O. Udrea, "Machine recognition of human activities: a survey," *IEEE Trans. Circuit and Systems for Video Technology*, vol. 18, no. 11, pp. 1473-1488, Nov. 2008.

[6] N. Ukita, A. Kotera, and K. Kidode, "Behavior recognition with HMM and feature analysis using a wide-view camera and a PTZ camera," *International Conference on Machine Vision and Applications*, Yokohama, Japan, pp. 235-238, May 20-22, 2009.

[7] S. H. Kim, J. H. Hwang, and I. K. Jung, "Object tracking system with human extraction and recognition of behavior," *International conference on Management and Artificial Intelligence*, Bali, Indonesia, pp. 11-16, 2011.

[8] Megret *et al.*, "The IMMED project: wearable video monitoring of people with age dementia," *ACM Multimedia 2010*, Firenz, Italy.

[9] M. L. Lee and A. K. Dey, "Lifelogging memory appliance for people with edpisodic memory impairment," *in Proceedings of the 10th International Conference on Ubiquitous Computing*, Seoul, Korea, September 21 - 24, vol.344, pp. 44-53, 2008.

[10] S. E. Lindley *et al.*, "Narrative, memory and practice: tensions and choices in the use of a digital artefact," *in Proceedings of the 23rd BCS conference on Human Computer Interaction*, pp. 1-9, September 2009.

[11] S. K. Chang, "A General Framework for Slow Intelligence Systems," *International Journal of Software Engineering and Knowledge Engineering*, vol. 20, pp. 1-15, 2010.

Pages 33 to 38 are removed from the proceedings due to author’s nonpayment of the registration fee.

Development of A Real-Time Air Pollution Alert System Using Smart Phones

Nuri Yilmazer[1], Young Lee[1], Juan Mora[1], and William Webb[1]
[1]Electrical Engineering and Computer Science Dept.
Texas A&M University-Kingsville
Kingsville, Texas, USA
nuri.yilmazer@tamuk.edu

Kuo-Jen Liao[2]
[2]Environmental Engineering Department
Texas A&M University-Kingsville
Kingsville, Texas, USA
kuo-jen.liao@tamuk.edu

***Abstract*—This paper introduces a real-time air pollution alert system by utilizing location awareness capability of smart phones and systems. Two popular platforms, Android and Apple's iOS, are chosen to develop applications which can inform the public in a timely manner in the case of worsened air quality conditions and these Apps can be easily extended to be used for other natural environmental disasters such as hurricanes, and chemical spills.**

Keywords-Air Quality; Real-time Alert System; Location aware systems; Smart Phones; Andorid, Ipad; Iphone.

I. INTRODUCTION

Current studies have shown that close to half of the population in United States is exposed to unhealthful levels of air pollution [1, 2]. This is either in the form of ozone or short-term level of particles. Increased level of air pollution pose a high risk to the human health so it is very important to develop smart and reliable systems to inform and update the public in a timely manner. There are some existing alert systems for air quality alert systems such as through broadcasting or customized emails and these approaches may not be very effective if the users do not have the access to internet, TV, or radio. Especially, it is important to determine the people in the region of the contamination and convey the alert to these group of people in real-time. There is need to incorporate the location of each person and track it if they are in the proximity.

In this paper, the advantage of the location-aware capabilities of smart-phones is taken to develop a warning system. Current smart phone systems are capable of providing the exact location of users in real time through the Global Positioning System (GPS) chips built in or utilizing the multiple cellular base-stations to determine the geographical locations [3]. Time delays of the signals from at least three base stations to the cell phones are measured to determine it. However, in this study, we have used the data provided by built in GPS chips which has an accuracy and resolution on the order of 30 cm which is quite suitable for developing alert system.

Smart phone technologies are enjoying their blooming years and so far close to 300 million smart phones were sold globally. Smart-phones find their applications in many fields such as context-based advertising, traffic planning, route predictions, early warning systems, and pollution alert systems [4-6].

Proposed air pollution alert system readily identifies locations of smart-phone users and delivers timely and effective air pollution alerts to people in the forecasted pollution areas through smart-phones so that the public will be able to take more timely and effective actions to protect their health from air pollution.

Regional air quality models are used for air quality forecasting and they include the predicted meteorology and air pollutant emissions. In this study, US EPA's regional air quality modeling system, Models-3, was used for simulating air quality in the U.S. Air quality information should be easy to understand for the general public so US Environmental Protection Agency (EPA) has standardized the this information. Air Quality Index (AQI) describes how polluted the air and their specific location as well as associated health effects on the human health and environment. The Table I below shows the two of the major pollutants and their levels.

TABLE I. AQI values, categories, and pollutant concentration thresholds [7].

	Ozone	**$PM_{2.5}$**
Good	0 – 64	0 – 15
Moderate	65 – 84	15 – 40
Unhealthy for sensitive groups	85 – 104	40 – 65
Unhealthy	105 – 124	65 -150
Very unhealthy	125 – 404	150 – 250
Hazardous	404	250

This study has been supported by AT&T Technology and Environmental Award and Texas A&M University-Kingsville University Research Award grants.

II. APPLICATION DEVELOPMENT

A smart phone is a mobile phone equipped with computing capabilities and applications and in recent years, the demand and the popularity of smart phones are increasing tremendously. Apple's iOS and Google's Android were used as main platforms in this study, as they have gained the major market share of smart-phones. Objective C programming language has been used to develop the Apps and development tools are easily accessible to the developers. Additionally, the Android applications are developed using Java and tools available under an open source license.

The overall operation of the real-time air quality alert system is depicted in Figure 1 below. The data obtained from Air Quality Forecasting model is received by the Servers to process further. This model helps us to determine the air quality and based on the modeling results, levels of pollution. If the pollution level is higher than the normal acceptable values the warning messages as well as a regional map showing the affected region is distributed via these smart phone apps to the public in the region of pollution.

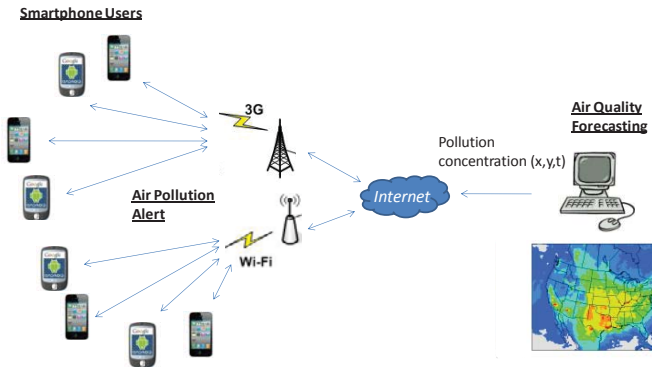


Figure 1. Integrated alert system using smart phones and air quality forecasting models

The jobs tasks implemented by the air quality alert system are;

- Web server:
  - store/receive temperature/air pollution data from Air Pollution Information Center.
  - responsible receiving smart phone users' location information
  - continuously checks to see if smart phone users are in air pollution areas, where concentrations or temperature are above certain values.
  - If the values exceed the normal values, then the alerts are sent to users in the form of text messages as well as regional map
- Smart Phone
  - to send smart phone user information
  - to receive air quality alert messages and map and display them

The air quality forecasting has been implemented on a high-performance computing cluster. The computing cluster has 40 nodes, each having 4 processors running at 3GHz and 8 GBs of DDR2 DRAM.

In this study, Forecasted air pollution data obtained from U.S. EPA's AirNOW air pollution system has been used to test the effectiveness of the apps [7]. This air quality alert system approach can readily be extended and modified for other environmental disaster alerts such as severe weather conditions and chemical spills.

The Apps send out the notifications to the users in the contaminated region. The Apps' settings can be easily customized by users such as the distance filter which allows users to adjust the range of the notification distance as well as sound or visual notification. It can be customized based on the preference of the users. The user in the contaminated area receives a map showing the contamination level and alert message includes the geographical location of the user, and the time/date the alert was sent. Different color coding has been used to denote air pollutant levels on an area map ranging from blue to red where the red color represents the high air pollutant levels. Additionally, Apps allow the users to store the location information and notification log sheet.

Both iOS and Android Apps has been successfully developed and tested with the actual data available. The Apps has been downloaded by users for one week of testing.

The snapshots taken from iOS are seen in Figure 2 and 3. As it can be seen, a warning message appears at the screen as well as the map indicating the level of the contamination.

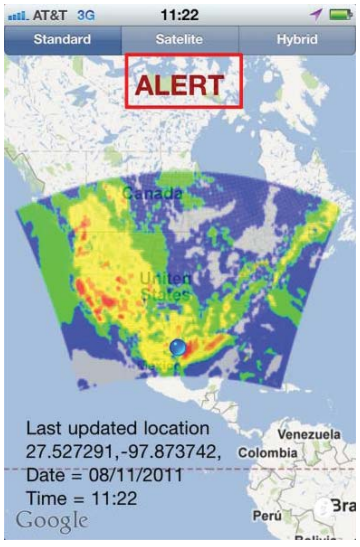


Figure 2. The iOS air pollution alert app developed

Figure 3. A snapshot of an air quality map on the smart phone screen

As it can be seen in Figure 3, the user can check the Ozone and PM2.5 levels on the map.

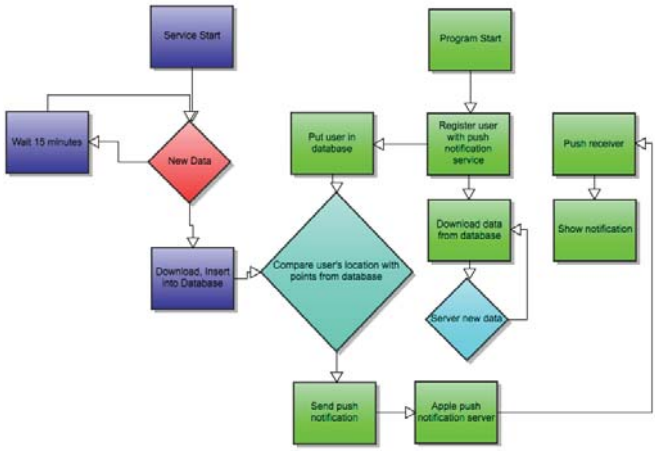


Figure 4. iOS flowchart

The programming flowchart and Class diagram of the iOS App can be seen in Figures 4 and 5 correspondingly.

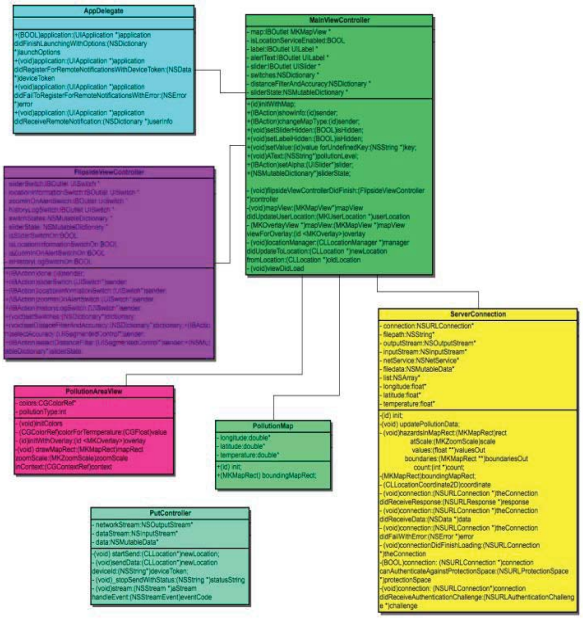


Figure 5. Class Diagram for iOS App

CONCLUSIONS

We have successfully shown that mobile devices with location awareness capability can be utilized to develo p a real-time air quality alert system. The alert system can deliver timely and effective air qu ality information to users who are in the contaminated region. Two popular platforms are cho sen to develop the Apps; Apples' iOS and Google's Anroid. As a future work, we plan to extend this system for other disastrous scenarios such as flooding.

REFERENCES

[1] American Lung Association: 2009, *State of the Air 2009*, p. 182.
[2] Dominici, F., Peng, R.D., Bell, M.L., Pham, L., McDermott, A., Zeger, S.L. and Samet, J.M.: 2006, 'Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases', *Jama-Journal of the American Medical Association* 295, 1127-1134.
[3] Panaghiotis Kikiras, Dimitris Drakoulis, 'An Integrated Approach for the Estimation of Mobile Subscriber Geolocation, *Journal Wireless Personal Communications*: Vol30 Issue 2-4, September 2004, 217–231.
[4] Peter Leijdekkers, Valerie Gay, 'Personal Heart Monitoring System Using Smart Phones To Detect Life Threatening Arrhythmias', Proceedings of the 19th IEEE Symposium on Computer-Based Medical Systems (CBMS'06), 2006.
[5] Y. Zhao. Mobile phone location determination and its impact on intelligent transportation systems. *IEEE Transactions on Intelligent Transportation Systems*, 1(1):55, 2000.
[6] P. Mohan, V. N. Padmanabhan, and R. Ramjee, Nericell: Rich Monitoring of Road and Traffic Condi-tions using Mobile Smartphones, In Proc. of ACM SenSys '08, Raleigh, NC, USA, Nov 2008.
[7] http://airnow.gov/

Sharing Digital and Physical Contents on and above Distributed Surfaces using Active Infrared Keying

Buntarou Shizuki, Yuichi Murata, and Jiro Tanaka
Department of Computer Science,
University of Tsukuba
1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8573, JAPAN
{shizuki,murata,jiro}@iplab.cs.tsukuba.ac.jp

***Abstract*— This paper presents SyncSurface, a system framework for distributed collaboration, which shares digital contents and physical contents on and above distributed surfaces. In such sharing, one important technique is to segment each frame, in a video stream from a camera placed above a surface, into its foreground (e.g., objects and users' bodies) and its background (i.e., displays). This paper describes a simple and low-cost segmentation technique called "active infrared keying" that uses a display that emits infrared light from its surface. We have implemented two systems using the framework to examine the feasibility of the technique. The results of our pilot studies indicate that the technique enables users to communicate with each other using physical miniatures on the surfaces, gestures for pointing out physical objects, and facial expressions.**

***Keywords*—tabletop, large display, whiteboard, CSCW, groupware, gestures, shared spaces, chroma keying, image processing, user interface.**

I. Introduction

During traditional co-located collaboration using surfaces (e.g., tables and whiteboards), physical contents, other than writing (e.g., documents and sketches), can also be shared and play important roles in collaboration. For example, miniatures of furniture are useful for discussing building plan; they are placed and moved by users on room maps on horizontal surfaces. Moreover, physical objects above ("in front of" in the case where vertical surfaces are used) surfaces are also important. For example, users hold miniatures in the air during discussions. Furthermore, users' hands, bodies, and faces are also physical objects above surfaces that play important roles in collaboration. That is, they use their hands and bodies to give gestures (e.g., deictic gestures such as pointing and tapping to point out some physical object on a surface), and their faces to express their emotion.

This also occurs in co-located collaboration using computer displays (e.g., co-writing technical documents and pair-programming). While users share digital contents on one or more computer displays, they also share physical objects including miniatures they hold in their hands, memoranda pasted on large displays, and themselves, providing gestures and expressing themselves.

In contrast, traditional CSCW systems support collaboration by sharing only digital contents such as still images, animations, GUI windows of applications, and cursors between distributed sites. Some CSCW systems and researchers have tried to improve distributed collaboration by sharing physical contents, including users' hands, bodies, and faces, as well as digital contents to provide rich expressiveness in communication. However, there is sill plenty to be explored, especially in the way to implement such sharing.

In this paper, we present an alternative framework, called SyncSurface (Fig. 1), for sharing digital and physical contents on and above distributed surfaces to support distributed collaboration. For such sharing, one important technique is to segment each frame, in a video stream from a camera placed above a surface, into its foreground (e.g., objects and users' bodies) and its background (i.e., displays). To this segmentation, we present a simple and low-cost technique called *active infrared keying*. The technique uses a display that emits infrared light from its surface for robust segmentation. We also describe two systems we implemented using the technique to examine its feasibility.

II. Related Work

Much research has investigated distributed collaboration by sharing physical contents on and above surfaces using cameras, aimed at the surfaces, to capture images of the contents, which are then sent to remote sites. An example

Figure 1. Communication using SyncSurface Table.

of such pioneering research is VideoDraw [1], [2]. It uses video cameras, each of which is aimed at a television set, to send physical contents including marks, which users draw with whiteboard markers directly on the surface, as well as the accompanying hand gestures on and above the surface. VideoWhiteboard [3] places cameras behind screens to capture an image of the marks that users draw on the front surface of the screens using whiteboard markers, and the shadow of the users. Then the system sends these data to the other site and projects them on the screen using a video projector. This enables the users to share physical contents: both real and remote marks, as well as shadows that contain the remote users' gestures and actions. ClearBoard-1 [4] uses a camera mounted above a surface (i.e., a half mirror with polarizing film) to obtain physical contents on and above the surface, including users facing the surface, and drawings made with whiteboard markers. The system sends the images from the camera to a remote surface and projects them onto the surface from its rear after mirror-reversing them. This allows eye contact with the remote user and makes visible each other's gestures, actions, and facial expressions as well as the drawing. Agora [5] and AgoraG [6] are tabletops that use multiple cameras at each site. The systems use cameras mounted above the surfaces to capture physical contents on and above the surfaces, and two cameras to capture users. The former images are sent to a remote site and projected onto the surface with a projector mounted above the surface. The latter images are projected onto vertical screens around the surfaces at the remote site, enabling users to use eye contact, actions, and gestures in their collaboration.

Although sharing a computer's desktop or GUI windows of applications (i.e., a kind of digital contents) among distributed sites using remote desktop software such as VNC, Apple Remote Desktop, and Microsoft Remote Desktop Connection, is the simplest and most common way for distributed collaboration, much research has been trying to share not only digital contents but also additional contents derived from physical contents by using cameras to give users awareness of other users. One such example is LIDS [7]. In LIDS, users can directly write on a horizontal display using a stylus, the position of which is tracked on the surface and is used to draw strokes in digital ink on the surface. At the same time, a camera placed behind the users captures images around the surface, including the users. The system converts the images into shadows (i.e., physical contents) and transmits them to other sites along with digital contents including the strokes. A projector placed behind the surface projects them from its rear. Thus, sharing the shadows allows the users to perform gestures, and provides the users with awareness of other users at remote sites. Distributed Tabletops [8] also shows shadows of users. The system uses cameras placed above the horizontal surfaces to capture the users' hands and other objects on and above the surfaces. A projector above each remote tabletop projects the users' shadows (i.e., physical contents) onto the the surface. VideoArms [9] extracts the users' hands and arms (i.e., physical contents) by detecting skin-colored pixels among images from cameras above the surfaces and sends them to remote surfaces, which are then combined with digital contents and presented to the remote users.

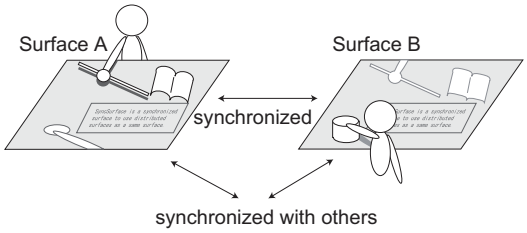


Figure 2. An implementation style of SyncSurface: table style.

Some researchers have explored sharing both digital and physical contents between distributed sites, in which cameras are also used to capture physical contents. DoubleDigitalDesk [10] uses two sets of DigitalDesk, which projects digital contents onto the desk and uses cameras above the desk to capture physical objects including paper documents on the desk, and recognizes the content, enabling the user to interact with them. By placing one set of DigitalDesk on each site, the system makes it possible to share physical contents on and above both desks as well as digital contents. ClearBoard-2 [11], [12] can also share both digital and physical contents on and above surfaces. It uses a camera mounted above a transparent digitizer with a half mirror with polarizing film to obtain physical contents on and above the surface, including users. The system uses the digitizer to capture strokes drawn on the surface. Then it sends the images from the camera and the strokes as well as a computer's desktop to a remote surface and projects them onto the surface from its rear.

Our proposed framework, SyncSurface, enables users to share both digital and physical contents on and above distributed surfaces to support distributed collaboration. In this sense, SyncSurface shares the same goals as DigitalDesk and ClearBoard-2. However, in contrast, our focus is on developing and examining a technique to achieve the sharing of both digital and all physical contents (e.g., books, miniatures, users' bodies and hands) on and above the surfaces in a low-cost and robust manner. To this end, we developed a technique to extract foreground objects from the background. We implemented two systems using this technique, and discuss its feasibility based on the results.

III. SyncSurface

SyncSurface is a system framework for distributed collaboration, which shares digital contents and physical contents on and above distributed surfaces. Fig. 2 illustrates distributed collaboration using SyncSurface. The system displays the same digital contents on all the surfaces located at multiple distributed sites. At the same time, it shares physical contents. Specifically, it captures physical objects on and above Surface A using cameras, then displays them on Surface B and other surfaces (if any). At the same time, the system also captures physical objects on and above Surface B, and displays

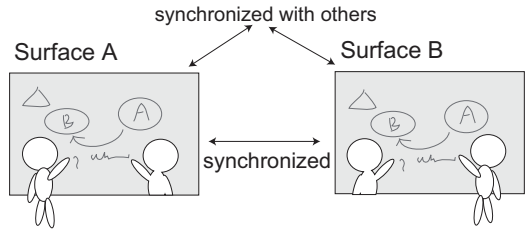


Figure 3. Another implementation style of SyncSurface: board style.

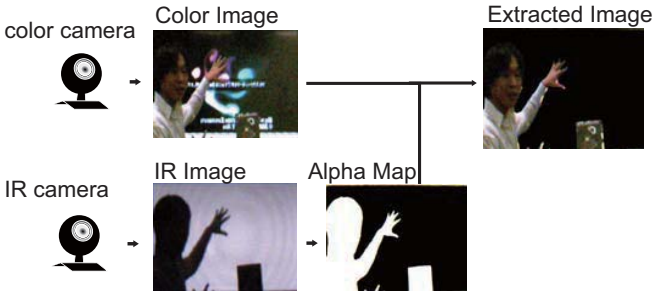


Figure 4. Image processing steps of active infrared keying.

them on Surface A and others. Because all of the physical objects on and above all of the surfaces are displayed on all of the surfaces, distributed users can collaborate using these objects and hand gestures almost in the same manner in which they can engage in co-located collaboration using one physical surface.

There are two possible implementation styles in SyncSurface: table style (Fig. 2) and board style (Fig. 3). Each has its own suitable applications. In table style, the users can place and move physical objects on the surface readily. Thus, this style suits designing, planning, and prototyping. For example, remote users can collaboratively examine the arrangement of the furniture in their new office by displaying a floor map of the office (i.e., digital contents) and then arranging various furniture models (i.e., physical contents) on the floor map. In contrast to table style, the users can use gestures that they use in face-to-face communication in board style, because this style presents the hands, bodies, and faces of the users in front of the surfaces. Thus, this style is suitable for distant presentation and remote lectures.

IV. Active Infrared Keying

To share physical contents on and above distributed surfaces, it is necessary to segment each frame, in a video stream from the camera placed above the surface, into its foreground (e.g., objects and users' hands, bodies, and faces) and its background. To this end, we developed *active infrared keying* (AIR keying). This segmentation technique uses a display emitting infrared light from its surface, achieving robust segmentation with a simple setup.

A. Focus

Techniques to segment regions of foreground objects within images from a camera have been heavily investigated and are widely used.

Many techniques use special backgrounds or utilize visual features of the foreground objects. For example, some techniques use a blue screen as the background, some extract regions where brightness is higher than a threshold, some use skin color-based detection, some extract regions where the shapes are the same as the target objects, some compare them with images captured beforehand, and others use polarizing films.

There are also keying techniques that use sensing technologies in addition to a camera. For example, [13] uses a thermographic camera to detect human regions. This technique is suited to segmentation of human regions while other kinds of objects are recognized as background. A depth-sensing camera can be used to recognize the regions of foreground objects. This method is attractive because it does not impose any visual precondition on the foreground or background objects. However, it is difficult to detect regions of thin objects (e.g., paper) placed on surfaces, because the depths of the regions are almost the same as those of the surfaces.

In contrast, our AIR keying takes a more active approach. The technique uses a display emitting infrared light from its surface as the background. Within the images from an infrared camera placed above the surface, it identifies regions where no infrared appears as a foreground region. Thus, it can extract regions of objects that no infrared light penetrates, and does not impose any visual precondition on foreground objects. While this technique does not detect transparent objects that infrared light penetrates, this is not a serious problem because most physical contents used in collaborations, such as paper, stationery, and miniatures, are opaque.

The approach of our technique is similar to that of IR Matte [14], which uses a retroreflector sheet as the background. Infrared light is provided from a ring of bright LEDs attached around the camera lens. In contrast to IR Matte, our contribution is to provide an alternative system, which is simple and low-cost; consequently it is easy to implement.

B. Process of AIR keying

AIR keying uses two cameras: a color camera and a camera with an IR pass filter (IR camera). The two cameras are placed in front of the surface, and a half mirror is situated between them so that the two cameras share the same optical axis (see Fig. 5, left).

The image processing steps of AIR keying proceeds as follows (Fig. 4):

1) The IR camera produces an IR Image, with a display emitting infrared light from its surface as the background. In this image, regions corresponding to foreground objects, including persons, are dark.
2) AIR keying uses alpha keying (as in [13]) to make edges of the foreground regions appear natural in the resulting image. To this end, AIR keying constructs an Alpha Map

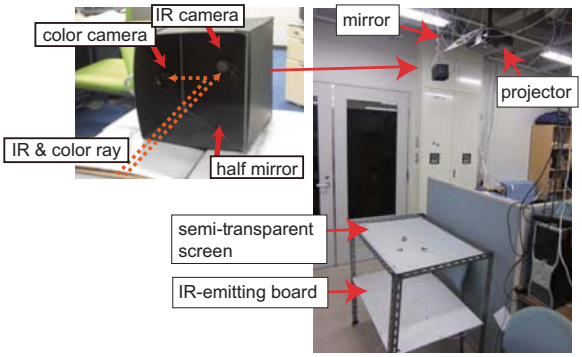


Figure 5. Left: a color camera, a camera with an IR pass filter, and a half mirror that is used so that the two cameras share the same optical axis. Right: our prototype of SyncSurface Table.

by applying a function with two thresholds to the IR Image.

3) AIR keying uses the IR Map to subtract the background from the Color Image. The result is the foreground image with alpha values (Extracted Image).

The system executes this process at each site, and sends the Extracted Image to other sites. Then, each surface obtains a synthesized image by collecting the Extracted Images from other sites and superimposing them on the digital contents. Finally, the surface displays the synthesized image.

C. Display that emits infrared light from its surface

An important component of AIR keying is a display that emits infrared light from its surface. There are two low-cost approaches to realize this:

- The first approach is to place an infrared projector behind a screen, in addition to a video projector. The video projector projects digital contents onto the screen from its front surface. In contrast, the infrared projector projects infrared light onto the screen from its rear. As a result, it is possible to obtain such a display. The merit of the first approach is that robust segmentation is achieved simply by using a powerful infrared projector.
- The second approach is to use an LCD display with back lighting that emits infrared light. The advantage of this approach is that the setup of the implementation is relatively smaller than the first approach.

V. Study Prototypes

We implemented two prototypes of SyncSurface to examine AIR keying and the communication using SyncSurface: a prototype of SyncSurface Table using the first approach and a prototype of SyncSurface Board using the second approach.

Fig. 5 shows our prototype of SyncSurface Table. The prototype has a video projector (and a mirror to reflect the video from the projector) mounted in the ceiling for projecting digital contents downward onto the semi-transparent screen (80 cm $\times$ 60 cm, roughly 40-inch). Below the screen, the prototype also has a IR-emitting board (same size as the screen) as an infrared projector. The IR-emitter board has an array of 3×6 infrared LEDs on the surface (Fig. 6). We coated the surface of the board with aluminum foil to prevent the board from absorbing the infrared light from the LEDs. The infrared light is captured with an IR camera above the surface, mounted in the ceiling. A color camera is also mounted to capture physical contents on and above the surface. The video projector can be scaled arbitrarily in terms of the size of the surface simply by using a larger board with more LEDs on its surface.

Figure 6. IR-emitting board (before being coated with aluminum foil) that we developed for our prototype of SyncSurface Table.

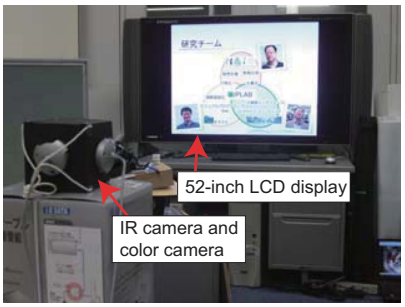


Figure 7. Prototype of SyncSurface Board.

Fig. 7 shows our prototype of the SyncSurface Board. We used a 52-inch LCD display in this prototype. The same cameras with a half mirror described in Fig. 5 were used to build this prototype.

In both implementations, before executing the process described in Section IV-B, a homography is applied to each frame in the video streams from the color camera and the IR camera. To this end, a calibration is necessary for each camera when the surface is set up.

VI. Pilot Studies

We conducted two pilot studies to examine AIR keying and the communication using SyncSurface.

A. Examining SyncSurface Table in remote communication

The goal of this pilot study was to examine whether users can communicate naturally with each other using SyncSurface Table. More specifically, we were interested in examining whether users can see both digital and physical contents on and above the remote surface naturally, use gestures (especially deictic gestures such as pointing and tapping) to physical contents on the remote surface, and use gestures to digital

Figure 8. Digital contents and physical contents are projected onto the users' hands (center) and onto the notebook (upper-left) that a user holds in his hand.

contents (i.e., observe the intention of gestures by remote users).

Setup

We built two separate sets of SyncSurface Table. We displayed a map of our university as the digital content and placed a miniature car on each SyncSurface Table. Two participants were involved. We asked one participant (host) to explain a route to our university to another participant (guest) (see Figs. 1 and 8). After the task, we interviewed both participants.

Results

Both participants said that they could see the place, color, and shape of the car of its counterpart. The host said that he could readily use the car to explain the route by moving it along the route on the map and to suggest a parking lot by placing the car on the map. At the same time, the guest said that he could understand the route and the parking lot.

Moreover, we observed that both participants frequently use deictic pointing gestures to indicate places on the map. The host used them for explanation. The guest used them for confirmation. This result suggests that users could effectively and naturally use gestures to digital contents on the SyncSurface.

However, we identified three problems in this study. First, when a participant picked up the car with his hand, the other participant could not see the car because the hand largely covered it. Second, as shown in Fig. 8, the system projected the digital contents and the physical contents of the remote site onto the participant ' s hands and onto the notebook that one participant held in his hand, because we used a front projection technique in this study. Third, the participants said that they sometimes felt frustration during communication, because one could not touch and move the physical contents of the remote site directly while he could see them.

B. Examining SyncSurface Board in remote presentation

There were two goals in this pilot study: to assess the second approach of AIR keying (i.e., using an LCD display as the display that emits infrared light from its surface) and to evaluate the efficacy of communication with physical objects visible only in front of a board.

Figure 9. Pilot study of SyncSurface Board.

Figure 10. Left: a color image whose background is subtracted using AIR keying (Extracted Image). Right: the result shown on SyncSurface Board, which is the color image synthesized with the digital content.

Setup

We built one SyncSurface Board (due to a lack of our hardware resources) and placed it in a remote room (Fig. 7). Then one author of this paper (presenter), standing before a prototype of SyncSurface Board, conducted a remote presentation, using slides as well as body gestures. We asked a participant to observe the remote presentation in another room, where we showed the synthesized images on a 50-inch plasma display (Fig. 9), almost the same size as the prototype of SyncSurface Board. After the presentation, we interviewed the participant.

Results

The participant said that he could see the gestures of the presenter during the presentation. Moreover, as shown in Fig. 10, the display clearly showed the actions and facial expressions of the presenter.

One problem revealed by this pilot study was the presence of noise in the images of physical contents.

VII. Discussion

In the first pilot study described in Section VI-A, SyncSurface Table successfully provided the users with shared digital and physical contents for communication. However, because the two sets of SyncSurface Table were connected to the same local network in this pilot study, it is necessary to further examine its effectiveness in various cases, especially in a case where there is severe network latency.

The first and third problems observed in the first pilot study (described above) are inherent to distributed communication and are still open for discussion. However, these problems are not serious, because users would quickly get used to avoiding such problems by communicating with each other.

We plan to address the second problem simply by using an Alpha Map (Fig. 4) as a mask to make the regions in the digital contents where physical contents exist black, before synthesizing the digital and physical contents. Such a system would still allow users to see and to indicate both local and remote physical content using pointing gestures.

The noise observed in the second pilot study was due to the half mirror we used. The half mirror was an acrylic panel, the reflection rate of which was low to capture the physical content using a color camera (in contrast, the infrared light penetrating the half mirror was powerful enough for keying). Thus, we will solve this problem by replacing the acrylic half mirror with one made of material that reflects infrared light well but allows visible light to penetrate.

Our pilot studies were conducted under controlled illumination, avoiding effects from external infrared light (e.g., a fluorescent light and ambient light, which also emit infrared light). In actual usage, however, such effects would need to be addressed. While the second approach requires an environment with little infrared light, the first approach can eliminate the effect, synchronizing the infrared projector with the IR camera. That is, the system turns the projector on and off repeatedly. Bright areas in IR images captured when the projector is turned off are infrared light from the environment. Thus, using bright areas found only in images captured when the projector is turned on produces the Alpha Map, where the effect of infrared light from the environment is eliminated.

VIII. Conclusions

This paper presents SyncSurface, a system framework for sharing digital and physical contents on and above distributed surfaces to support distributed collaboration. The framework utilizes *active infrared keying*, a simple and low-cost technique for extracting physical objects from camera images, by using a display that emits infrared light from its surface. The system proved useful in two different pilot studies with two different implementations (SyncSurface Table and SyncSurface Board). In the future, we will expand our system in a way that addresses these problems found in two pilot studies and use the refined system in further user studies.

References

[1] J. C. Tang and S. L. Minneman, "VideoDraw: a video interface for collaborative drawing," in *Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people*, pp. 313–320, ACM, Apr. 1990.

[2] J. C. Tang and S. L. Minneman, "VideoDraw: a video interface for collaborative drawing," *ACM Transactions on Information Systems*, vol. 9, pp. 170–184, Apr. 1991.

[3] J. C. Tang and S. L. Minneman, "VideoWhiteboard: video shadows to support remote collaboration," in *Proceedings of the SIGCHI conference on Human factors in computing systems: Reaching through technology*, pp. 315–322, ACM, April/May 1991.

[4] H. Ishii and M. Kobayashi, "ClearBoard: a seamless medium for shared drawing and conversation with eye contact," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, CHI '92, pp. 525–532, ACM, May 1992.

[5] H. Kuzuoka, J. Yamashita, K. Yamazaki, and A. Yamazaki, "Agora: a remote collaboration system that enables mutual monitoring," in *CHI '99 extended abstracts on Human factors in computing systems*, pp. 190–191, ACM, May 1999.

[6] J. Yamashita, H. Kuzuoka, P. Luff, C. Heath, and K. Yamazaki, "Supporting paper mediated video conference with Agora system," in *Proceedings of the 2002 ACM on Computer supported cooperative work video program*, p. 2, ACM, Nov. 2002.

[7] M. Apperley, L. McLeod, M. Masoodian, L. Paine, M. Phillips, B. Rogers, and K. Thomson, "Use of video shadow for small group interaction awareness on a large interactive display surface," in *Proceedings of the Fourth Australasian User Interface Conference (AUIC2003)*, pp. 81–90, Australian Computer Society, Inc., Feb. 2003.

[8] P. Tuddenham and P. Robinson, "Distributed Tabletops: Supporting remote and mixed-presence tabletop collaboration," in *Proceedings of the Second Annual IEEE International Workshop on Horizontal Interactive Human-Computer Systems (TABLETOP 2007)*, pp. 19–26, Oct. 2007.

[9] A. Tang, C. Neustaedter, and S. Greenberg, "VideoArms: Embodiments for mixed presence groupware," in *Proceedings of the 20th BCS HCI Group Annual Conference (HCI 2006)*, pp. 85–102, Springer, Sept. 2006.

[10] P. Wellner, "Interacting with paper on the DigitalDesk," *Communications of the ACM*, vol. 36, pp. 87–96, July 1993.

[11] H. Ishii, M. Kobayashi, and J. Grudin, "Integration of interpersonal space and shared workspace: ClearBoard design and experiments," *ACM Transactions on Information Systems*, vol. 11, pp. 349–375, Oct. 1993.

[12] H. Ishii, M. Kobayashi, and K. Arita, "Iterative design of seamless collaboration media," *Communications of the ACM*, vol. 37, pp. 83–97, Aug. 1994.

[13] K. Yasuda, T. Naemura, and H. Harashima, "Thermo-Key: Human region segmentation from video," *IEEE Computer Graphics and Applications*, vol. 24, pp. 26–30, January/February 2004.

[14] H. Mitsumine, Y. Yamanouchi, T. Fukaya, H. Okubo, and S. Inoue, "A novel image compositing method considering studio lighting – IR Matte –," *Proceedings of the IEICE General Conference*, vol. 2008, p. 103, Mar. 2008. (in Japanese).

Fuzzy Color Space Segmentation to Identify the Same Dominant Colors as Users

José C. Amante Manuel J. Fonseca
Department of Computer Science and Engineering
INESC-ID/IST/TU Lisbon, Portugal
jcxa79@gmail.com, mjf@inesc-id.pt

Abstract

In this paper we propose a solution for the identification of the dominant colors in pictures taking into account the way people perceive them. To that end we developed an algorithm based on a reduced palette of 12 colors and on a non-uniform segmentation of the HSV color space using fuzzy membership functions. Experimental results from an evaluation with users, showed that our algorithm based on the fuzzy histogram is able to identify the same three dominant colors as users in 50% of the cases, while the classic histogram is only able to identify 43%.

1. Introduction

In the last years, due to the massive volume of images collected everyday, efficient, effective and easy to use systems for exploration, organization and retrieval of images are needed. In general, color is one of the most visually distinguishable characteristics, causing that the majority of the existing solutions take advantage of that and use the dominant colors (DCs) to describe the image content.

However, most solutions for dominant color identification use an algorithm-centered approach, rather than looking at how people perceive the DCs. In general, these solutions are only concerned with some system-centered measures, and do not consider the DCs that the users see on the images. Additionally, most of the existing works consider too many colors as possible DCs, making the enumeration and specification of colors by the users an almost impossible task, when they want to explore or retrieve images.

In this paper we present a new method for the identification of the dominant colors in images, which was built around the following three principles: i) a palette of only 12 colors, that all cultures are able to enumerate; ii) a non-uniform segmentation of the HSV (Hue, Saturation and Value) color space to better reflect the "size" of each color; iii) the use of fuzzy membership functions [3] to convey the ambiguity that is present in some regions of the color space.

With this approach we want to foster the creation of retrieval and exploration systems able to identify the real DCs that people see when looking at a picture, and where users could easily enumerate the DCs that they want to find in the images, using for instance textual queries, natural language or sketches of color blobs.

We performed an experimental evaluation with users, where we compared the DCs identified by users with the DCs identified by our algorithm and by the classic histogram, using two sets of pictures. Results showed that our algorithm was able to identify the same three DCs as the users in 50% of the times.

In the remainder of this paper we analyze some solutions for the identification of the DCs, present our fuzzy segmentation of the HSV color space and describe our algorithm for the DCs identification. To conclude the paper, we present the main results from the evaluation with users and provide some conclusions.

2. Related Work

Along the years several solutions for the identification of the DCs in images have been developed. Here we describe some of them, ranging from histogram, segmentation, correlogram and statistical based approaches.

One of the first segmentation based methods for the identification of DCs was developed by Smith et al. [16]. The authors used the HSV color space and a palette of 18 different hues, to identify a maximum of five dominant colors per region. Their method was based on the uniform segmentation of the color space and in the application of a color median filter.

The MPEG-7 standard [14] also has a descriptor for the dominant colors, called Dominant Color Descriptor (DCD), which extracts the distribution of dominant colors in images, using clustering and color quantization. The descriptor contains the dominant colors, their percentages, variances and their spatial coherency.

Yang et al. developed a new algorithm [18] based on MPEG-7, which uses the modified Generalized Lloyd Algo-

rithm (GLA) to obtain a small number of representative colors and their percentages. Authors divided the RGB color space in eight equal regions, and selected the centroid of each block as its quantized color by averaging the color distribution. Their algorithm identifies four to five dominant colors.

Liu et al. [13] developed a region-based image retrieval system with high-level concepts obtained from colored region features. The authors, first segmented each image into homogeneous regions and then defined a DC for each region and assigned a semantic name.

In [8] Huang et al. introduced an algorithm for extracting DCs in the CIELab color space, using an Ant Colony Clustering algorithm. The authors tried to address the two known issues of the classic GLA quantization schemes: avoid the clustering to get into local optimality and the sensibility to initial clustering centers.

Younes et al. proposed a fuzzy approach to segment the HLS color space, using a palette of nine colors [19]. Their main goal was to create a solution where users could retrieve images according to their DCs, expressed through fuzzy linguistic expressions.

In [5] Chamorro-Martínez et al. presented an approach where the HSI color space is divided in three zones: chromatic, semi-chromatic (gray) and achromatic (black). The chromatic zones is then divided using fuzzy membership functions for the Hue, Saturation and Intensity. In combination with the fuzzy logic, authors used also a clustering algorithm to identify the DCs.

Kiranyaz et al. addressed the DC extraction as a dynamic clustering problem [10]. The authors used techniques based on Particle Swarm Optimization to find the optimal number of DCs in the HSV color space. Their algorithm looked for the best solution between a range of dimensions related with the number of dominant colors. The algorithm is very computationally demanding and there is no guarantee that the solution will not converge to a local minimum, due to over- or under-clustering.

In [11] Lézoray et al. introduced an unsupervised morphological clustering on 2D histograms. These 2D histograms are pair-wise associations of the RGB color space, namely Red-Green, Red-Blue and Green-Blue. After the independent clustering on each of these three pair-wise histograms, a region merging is performed with the help of a Minimum Spanning Tree based algorithm. The DCs correspond to the maximums (centroids) in the 2D histograms.

Bhoyar et al. proposed another system for color histogram computation, based on fuzzy logic [4]. Authors used a palette of 11 colors and fuzzy if-then rules to decide the color of a pixel from its RGB components. A color histogram with 11 bins is then computed for each image, and used as the color descriptor.

In [9] Kiranyaz et al. presented a color descriptor which represents DCs both in spatial and color domains. The DCs are extracted using a quad-tree decomposition to partition the image, and recursively subdivide each cluster into four quadrants. The authors used a clustering method that iterates until reaching a maximum number of clusters, which correspond to the DCs.

Almost all of the previously described works, do not consider the perception that users have of the DCs in pictures. None of them performed tests with users to verify which DCs and how many DCs people see in common pictures. Moreover, the segmentation of the color space is often made with empirical rules and using the RGB color space, which is not the best to separate the chromatic, from the concentration or the brightness. Finally, in general the methods for DCs identification only take into account the Precision and Recall measures or the Averaged Normalized Modified Retrieval Rate (ANMRR), putting all the evaluation focus on the system and forgetting the users.

In our algorithm, we took a different approach by putting the users in the center of the development cycle, to create a solution that will identify the DCs perceived by humans.

3. Fuzzy Segmentation of the Color Space

It is relatively unanimous among various authors, like for instance [19, 5, 12, 1, 4, 20, 10], that the use of fuzzy logic applies well to the segmentation of the HSV color space, because the Hue circle has a subtle continuous variation of colors. For the case of far apart pure hues, such as red (HSV: 0,1,1) or cyan (HSV: 180,1,1), we almost never have doubts, but for any two pure consecutive hues, like for instance yellow and green, it is difficult for a person to see the boundaries. With the fuzzy approach we can have a variable grade from two adjacent hues and gently combine these hues with a proportional weight. For example, we are able to say that a pixel is 0.6 green and 0.4 yellow.

In this section we describe our proposal for the segmentation of the HSV color space using fuzzy logic and a palette of 12 colors. Contrary to other authors [5, 13, 15, 18] that first divide the color space and then assign names to the resulting "slices" of color, we first selected the color palette, and only after we divided the color space.

3.1. The 12 Colors Palette

As we have seen in the related work section, there is no consensus in the number of colors to use in the quantization process or in the segmentation of the color space. Some authors use few colors to get a short processing time, but with the disadvantage of diverging from the original colors, while others use more colors, resembling the original colors found in the pictures, but achieving more colors than the human eye can distinguish.

In the last decade various authors [6, 1, 4] adopted the set of 11 colors identified by Berlin et al. [2] in 1969 (red, brown, orange, yellow, green, blue, purple, pink, black, gray and white). This set, which contains the colors that can be named by all cultures, was coined by Chang et al. [6] in 2000 as the "Just Not the Same" colors (JNS), because any two colors from this set are perceived as JNS.

In our work we decided to use an extended version of this set, introduced by Ware in the scope of an application for nominal information coding [17]. This palette has 12 colors, and is composed by the JNS colors plus the cyan. We opted for this palette because it is reacher than the JNS, but still does not contain too many colors, and especially, because it contains colors whose names people can easily enumerate, enabling them to specify colors using various modalities, such as speech, writing or sketches. This same set of 12 colors is also used by Google Images.

3.2. The Achromatic Colors

Our approach performs the segmentation of the HSV color space in two steps. First, we consider the achromatic colors and only after we isolate the chromatic ones.

Taking into consideration that the HSV color space has two particularities: i) the Hues of colors with small Values or small Saturations are not representative; ii) the Saturation of colors with low levels of Value is not representative; we identified three zones in the color space for the achromatic colors (black, gray and white), based on the brightness degree of each point. While Chamorro-Martínez et al. [5] opted only for two zones (black and gray), we decided to include a third zone for white, because when we have a high Value (for certain Saturation) from the user point of view it is white. We confirmed this by doing some informal tests with users.

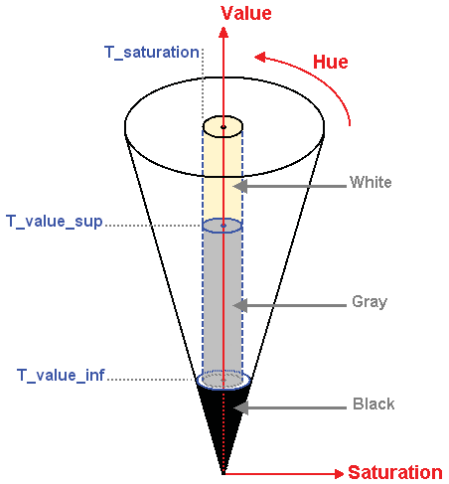


Figure 1. The three achromatic colors (black, gray and white) of the HSV color space.

Figure 1, represents the HSV color space with the three achromatic zones. In summary, a color in HSV is black if its Value component is less than T_value_inf. A color is gray if its Value component is delimited between T_value_inf and T_value_sup, and at the same time its Saturation component is less than $T\_saturation$. A color is white if its Value is greater than T_value_sup and its Saturation is less than $T\_saturation$. The values for these thresholds were defined based on the intervals defined by Chamorro-Martínez et al. [5], namely for $T\_value\_inf = 0.19$ we considered the intervals `Dark` and `Very Low Illuminated`; for $T\_value\_sup = 0.81$ we considered the intervals `Very High Illuminated` and `Bright`; and finally for $T\_saturation = 0.14$ we considered the interval `Very Low Saturated`.

3.3. The Chromatic Fuzzy Membership Functions

While most authors [5, 15] segmented the Hue dimension in equal intervals and then assigned a name to each, in our approach we first identified the colors and only after we defined the size of the intervals for each color.

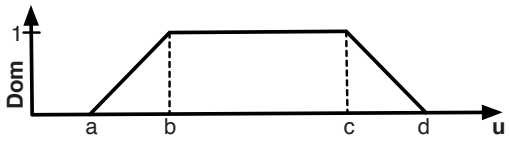


Figure 2. Definition of a fuzzy set with a trapezoidal-shaped membership function.

We use trapezoidal-shaped membership functions (see Figure 2), because according to Benavente et al. [1] there are no worth gain comparatively to Gaussian-shaped functions. To define the four values for the Hue of the 9 chromatic colors (see Figure 3), we applied the following rules:

1. The central hue of each trapezoidal membership function is the one defined by the Color Module of the last CSS3 specification, published by W3C (http://www.w3.org/TR/2011/REC-css3-color-20110607).
2. The kernel[1] size of each membership function is 10° for brown and orange, 20° for red, yellow, purple and pink, and 40° for green, cyan and blue;
3. Two membership functions always cross at the 0.5 degree of membership (dom);

[1]The kernel of a fuzzy membership function is where its degree of membership is equal to 1.

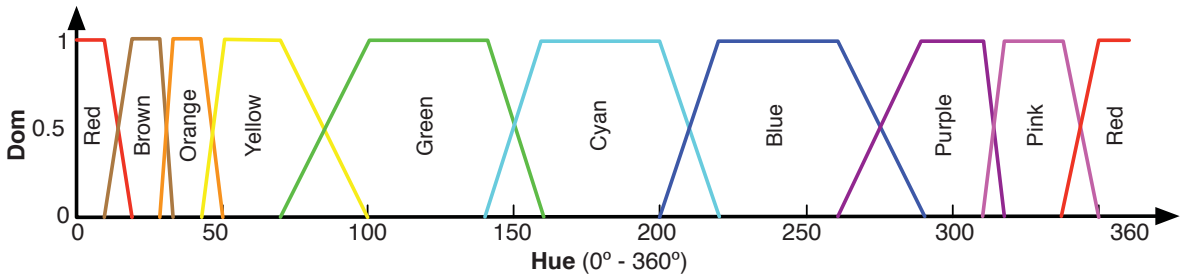


Figure 3. Membership functions for the Hue of the 9 chromatic colors.

4. The right slope of one trapezoidal-shaped membership function is equal to the left slope of the adjacent membership function (and vice versa);

5. Three different membership functions never overlap;

We also defined membership functions for the Saturation and Value components of the HSV color space. Similarly to Younes et al. [19] we defined three membership functions for Saturation and three for Value: low (< 0.25), mid (> 0.25 & < 0.75) and high (> 0.75).

These membership functions are used only to separate brown and pink from their neighbors in the color space, because they are non pure-hue colors (the Hue component alone is not enough to completely define them). Brown, for instance has a Hue of 24°, a Saturation of 0.86 and a Value of 0.54. Based on this, to completely define brown, orange, pink and purple, we considered not only their Hues but also their Saturation and/or Value components.

A orange Hue with mid or low value (dark orange) is considered brown, and a low saturated bright orange (beige) is also considered brown. On the other hand, a brown Hue with high Saturation and high Value (saturated bright brown) is considered orange.

When a color has red or brown Hue and has mid or low Saturation and high Value (skin color), then it is considered pink. It is also considered pink when we have purple Hue with high Value (bright purple). Finally, a dark pink (pink Hue with mid or low Value) is considered purple.

4. Fuzzy Histogram for DC Identification

A 'normal' histogram is created by dividing the color space into a set of bins and then by counting the number of pixels of the image that belong to each bin. For a space of N colors, the histogram is an N-dimensional vector $(h_1, ..., h_N)$, where each element h_i represents the portion of pixels of color i in the image.

In our solution, for each pixel of the image we evaluated the degree of membership for its Hue and assign it to the correspondent bin of the fuzzy histogram, according to the following rules:

1. If the pixel color has a hue with a single dom we add one to the correspondent bin;

2. When a color has a hue with two degrees of membership, then we evaluate the dom for each hue, and add it to the respective bin.

For example, a pixel with a Hue of 75 will have a dom of 0.83 for yellow and a dom of 0.17 for green. It means that we will add 0.83 to the yellow bin of the histogram and 0.17 to the green bin. The achromatic colors (black, gray and white) are always processed before the chromatic ones. A pixel classified as achromatic does not enter into the chromatic tests.

After processing the complete image, each of the 12 bins of the fuzzy histogram will have the sum of the doms for their Hues. Finally, we normalize the values (0-1) of the bins to make the comparison easier.

To identify the dominant colors, we order the bins according to its value and select the biggest ones as the dominant colors. All bins with values below 0.1 are not consider for being a dominant color.

5. Experimental Evaluation

To address the problem of identifying the dominant colors in images from the human point of view, we performed a study with users to identify the **name** of the colors that users know, **how many** dominant colors do users perceive on typical images, like for instance photos, and **what** dominant colors do they identify on each image.

The collected results were used to validate the 12 colors palette and to evaluate and compare our fuzzy histogram with the classic histogram.

5.1. Procedure of the Study

The study was done using an on-line survey composed of four parts. The first part was used to collect information about the users, such as the age, sex and if they wear glasses or contact lenses. The second part was used to test the color

blindness of the subjects. To that end we used three pictures from the Ishihara test [7].

In the third part we presented five images to the users (one at a time) and asked them to write freely the name of the dominant colors that they perceive in the pictures, using their mother language. Finally, in the fourth part we showed 20 images to the users (one at a time) complemented with a palette of 12 colors and a five levels Likert scale (with 1 meaning *Completely Disagree* and 5 meaning *Completely Agree* that this color is dominant), for users to express their level of agreement with each of the 12 colors, relatively to the corresponding picture.

Intentionally we did not provide a definition of "dominant color" nor mentioned the "correct" number of dominant colors to identify, to avoid affecting the choices of the inquired subjects.

The 25 pictures were chosen from the Stock.xchng (www.sxc.hu) website, taking into account that they should be representative of what a normal user will have in his/her personal photo collection and that they should be filter-free. Pictures main topics go from close portraits, to wide nature landscapes, sport scenes, Venice gondolas or even a crowded street of New York. The 25 images were randomly assigned to the third and fourth part of the survey.

5.2. Results from the Study

We collected 39 valid surveys, 14 from females and 25 from males, with half of them belonging to the 18–29 age group. Half of the participants (19) wear glasses or contact lenses and none of them had color blindness problems. We did not encounter any relation between age, sex or the use of glasses/contact lenses and the given answers.

For the first set of five pictures, participants enumerated on average 1.94 ($SD = 0.36$) dominant colors per image, and the most frequent number of DCs written per image (the mode) was one. In the second set of 20 pictures the average number of DCs identified per image was 3.39 ($SD = 0.49$), and the most frequent number of DCs selected per image (the mode) was three.

From the obtained results and also from some informal conversations with users, after they answered the survey, we can conclude that when users have the color palette available or know the name and number of colors that they can use, they enumerate more DCs per image.

We also counted the names of the colors written by the users, for the group of five pictures, that belonged to the 12 colors palette, and we observed that 94.6% of the names written by the users belonged to it. From this, we can conclude that the color palette that we adopted is appropriated for the identification of dominant colors, when users have to write or mention the name of a color.

5.3. Comparative Evaluation

To evaluate our algorithm for DCs identification based on the fuzzy segmentation of the color space, we decided to compare it with the classic histogram, which is a mere count of pixels, collected on each respective bin. Although the majority of the algorithms that compute the classic histogram to identify the DCs, do a uniform segmentation of the color space, we used the same palette of 12 colors and our non-uniform segmentation of the HSV color space, to make the comparison fair. However, in this case we used a strict non-fuzzy segmentation of the color space, with the limits for each Hue in the values where it has a dom of 0.5.

The DCs identified by our algorithm using fuzzy histograms and the classic histogram were compared with the DCs enumerated by the users in the study. To perform the comparison, we considered the three most dominant colors from each image, and we analyzed the set of five images and the set of 20 images separately.

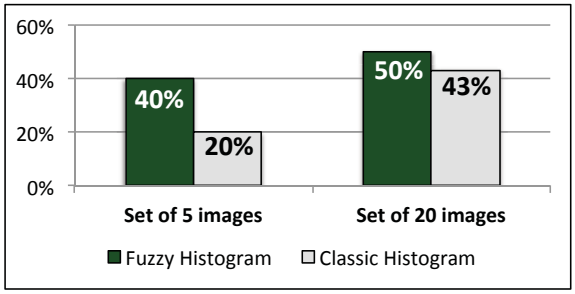


Figure 4. Percentage of three DCs identified by the algorithms equal to those enumerated by the users.

From Figure 4 we can see that in both cases, the fuzzy histogram identifies more colors similar to those enumerated by users than the classic histogram.

In the case of the set of five images, we have for the fuzzy histogram an average of 40.0% ($SD = 14.9$) of the DCs equal to those identified by the users, while for the classic histogram we have only 20% ($SD = 18.3$). The pairwise t-test shows that the average percentage of equal DCs for each image is significantly bigger on our algorithm than on the classic histogram, with a value of $p < 0.07$.

In the case of the set of 20 images, we have for the fuzzy histogram an average of 50.0% ($SD = 22.9$) of the DCs equal to those identified by the users, while for the classic histogram we have 43.3% ($SD = 21.9$). The pairwise t-test shows that the average percentage of equal DCs for each image is significantly bigger on our algorithm than on the classic histogram, with a value of $p < 0.04$.

Although the classic histogram had inherited a great advantage from our approach - the non-uniform segmentation

of the color space - from the experimental results we can state, with at least 97% certainty, that our approach is better than the classic histogram, in identifying the three DCs that users perceive from images.

6. Conclusions

We presented in this paper a new algorithm for the identification of the DCs in images, based on fuzzy logic and on a non-uniform segmentation of the HSV color space. Moreover, and contrary to other approaches, we first defined the color palette and only after we segmented the color space. We selected a palette of 12 colors, which includes only colors that can be enumerated by all cultures.

To identify the DCs in an image, we first create a fuzzy histogram by adding the degrees of membership associated to the color of each pixel in the image. The colors with the highest values are considered the DCs.

To evaluate our solution we followed a different path from other researchers, by performing tests with users. With these, we wanted to know how many and what DCs users saw on the images, and verify if our algorithm was identifying the same DCs as users.

Experimental results revealed that the 12 colors palette was a correct choice, since almost 95% of the colors enumerated by the users belonged to it. We also found that the majority of the users identified only three DCs on each picture, contrary to some authors who claimed that common users can distinguish up to 256 or 512 different DCs.

Finally, the comparison of our fuzzy histogram with the classic histogram, showed that our algorithm for the identification of the dominant colors presents results more similar to those produced by users.

Acknowledgments

This work was supported by FCT through the PIDDAC Program funds (INESC-ID multiannual funding) and the Crush project, PTDC/EIA-EIA/108077/2008.

References

[1] R. Benavente, M. Vanrell, and R. Baldrich. Parametric fuzzy sets for automatic color naming. *Journal of The Optical Society of America*, 25(10):2582–2593, 2008.

[2] B. Berlin and P. Kay. *Basic color terms : their universality and evolution*. University of California Press, 1969.

[3] J. C. Bezdek and S. K. Pal. *Fuzzy Models for Pattern Recognition*. IEEE Press, 1992.

[4] K. K. Bhoyar and O. G. Kakde. Image retrieval using fuzzy and neuro-fuzzy approaches with fuzzy color semantics. In *Proc. of the Int. Conf. on Digital Image Processing (ICDIP'09)*, pages 39–44. IEEE Computer Society, 2009.

[5] J. Chamorro-Martínez, J. M. Medina, C. D. Barranco, E. Galán-Perales, and J. M. Soto-Hidalgo. Retrieving images in fuzzy object-relational databases using dominant color descriptors. *Fuzzy Sets & Syst.*, 158(3):312–324, 2007.

[6] E. Y. Chang, B. Li, and C. Li. Toward perception-based image retrieval. In *Proc. of the IEEE Workshop on Content-based Access of Image and Video Libraries (CBAIVL'00)*, pages 101–105. IEEE Computer Society, 2000.

[7] L. H. Hardy, G. Rand, and M. C. Rittler. Tests for the detection and analysis of color-blindness. *Journal of The Optical Society of America*, 35(4):268–271, April 1945.

[8] X. Huang, S. Zhang, G. Wang, and H. Wang. A new image retrieval method based on optimal color matching. In *Proc. of the Int. Conf. on Image Processing, Computer Vision & Pattern Recognition (IPCV'06)*, pages 276–281, 2006.

[9] S. Kiranyaz, M. Birinci, and M. Gabbouj. Perceptual color descriptor based on spatial distribution: A top-down approach. *Journal of Image and Vision Computing*, 28(8):1309–1326, August 2010.

[10] S. Kiranyaz, S. Uhlmann, and M. Gabbouj. Dominant color extraction based on dynamic clustering by multi-dimensional particle swarm optimization. In *Int. Workshop on Content-Based Multimedia Indexing (CBMI'09)*, pages 181–188, June 2009.

[11] O. Lézoray and C. Charrier. Color image segmentation using morphological clustering and fusion with automatic scale selection. *Pattern Recognition Letters*, 30(4):397–406, 2009.

[12] Q. Li, Z. Shi, and S. Luo. Image retrieval based on fuzzy color semantics. In *IEEE Int. Conf. on Fuzzy Systems (FUZZ-IEEE'07)*, pages 1 –5, July 2007.

[13] Y. Liu, D. Zhang, G. Lu, and W.-y. Ma. Region-based image retrieval with perceptual colors. In *Proc. of Pacific-Rim Multimedia Conference (PCM'04)*, pages 931–938, 2004.

[14] F. C. D. MPEG-7. Text of iso/iec 15 938-3 multimedia content description interface - part 3: Visual, February 2001. URL: http://mpeg.chiariglione.org/standards/mpeg-7/mpeg-7.htm.

[15] M. Saeed and H. Nezamabadi-Pour. Fuzzy color quantization and its application in content-based image retrieval. In *Proc. of the WSEAS Int. Conf. on Circuits, Systems, Signal and Telecommunications (CISST'08)*, pages 60–66, 2008.

[16] J. R. Smith and S.-F. Chang. Single color extraction and image query. In *Proc. of the Int. Conf. on Image Processing (ICIP'95)*, page 3528. IEEE Computer Society, 1995.

[17] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers Inc., 2004.

[18] N.-C. Yang, W.-H. Chang, C.-M. Kuo, and T.-H. Li. A fast mpeg-7 dominant color extraction with new similarity measure for image retrieval. *Journal of Visual Communication and Image Representation*, 19(2):92–105, 2008.

[19] A. A. Younes, I. Truck, and H. Akdag. Image retrieval using fuzzy representation of colors. *Journal of Soft Computing*, 11(3):287–298, October 2006.

[20] X. Zhu, J. Zhao, J. Yuan, and H. Xu. A fuzzy quantization approach to image retrieval based on color and texture. In *Proc. of the Int. Conf. on Ubiquitous Information Management and Communication*, pages 141–149. ACM, 2009.

Insights on the development of visual tools for analysis of pollution data

Paolo Buono and Maria Francesca Costabile
Dipartimento di Informatica
Università degli Studi di Bari Aldo Moro
Via Orabona, 4 - 70125 Bari, Italy
{buono, costabile}@di.uniba.it

***Abstract*—Developing visual tools that support data analysis in a specific application domain requires a careful investigation in order to understand needs and expectations of people who will use such tools. The domain experts addressed in this paper are chemists specialized in environmental data analysis. Their main activity is to detect and monitor chemical compounds in the air through many devices in order to detect anomalies or prevent risks. One of the main problems that chemists face is the analysis of the huge amount of data produced by devices. They perform explorative data analysis and are willing to use software tools that can help them to get insights from data. This paper reports the experience in working with chemists to identify interactive visual tools that can be useful for their purposes. It provides insights on the difficulty of creating systems that users find really useful for their work, even when users participate in the design team. Because of the complexity of the considered problem and the fact that people are unable to make explicit all their needs and requirements, the identification of proper tools resulted very challenging.**

Keywords—Environmental data; data analysis; user-centred design; participatory design.

I. INTRODUCTION

Today, one of the problems in data analysis is the quantity of data that have to be analyzed in different application domains and for many purposes. Tools and techniques have been studied and developed in order to assist people in this heavy task. Visual analytics is an emerging interdisciplinary research field that includes, among others, Data Mining and Information Visualization techniques. The purpose is to make sense of very large and complex datasets by combining "automated data analysis with interactive visualizations for an effective understanding, reasoning and decision making on the basis of very large and complex datasets" [14]. This research field aims at utilizing the strengths of automatic methods as well as the innate human ability to visually perceive patterns and trends to help people analyzing large complex data.

Many visual analytics tools have already been developed, e.g. Tableau [22], Spotfire [21], Jigsaw [11]. They all claim to be general purposes, i.e. they support users in the analysis process in several context. The experience reported in this paper shows the difficulty to adapt such tools in order to fulfill the requirements of the experts of the application domains in which these tools might be applied, who usually are not IT professionals. Indeed, domain experts are used to their working and reasoning strategies and they are often reluctant to change their habits, and they do not want to be constrained by the technology. The experience reported in this paper shows that, when they are using a technological tool that does not reflect their working habits, after a certain time they do not use it anymore.

We adopted user-centered design and participatory design paradigms in order to design tools that chemists could find adequate to their needs and pleasant to use. User-centered design requires that end users' needs, profiles, tasks, context of use are deeply analyzed and that system is designed and developed by iterating a design-implementation-evaluation cycle, which includes end users in prototype evaluation [10]. Participatory design goes even further in user involvement, since it requires the participation of end users in the design process [23]. The rationale is that users are experts of the work domain thus a system can be effective if these experts are allowed to participate to its design, giving indications on their needs and expectations. However, even by adopting those paradigms, another problem emerged, determined by the fact that end users are unable to make explicit all their needs and expectations during the requirement analysis, even when they are involved in the design team, as suggested by participatory design. As shown in the paper, users really understand how the system works and are able to provide useful feedbacks only when they use the system, or a prototype, in their working practices.

The domain experts addressed in this paper are chemists specialized in environmental data analysis. Everyday, they analyze huge amount of data produced by different devices located on a territory, in order to monitor air pollution. By involving the chemists in a multidisciplinary team, we have analyzed several visual analytics tools that may support the

chemists' work. In this paper we report the difficulties encountered in designing tools that satisfy chemists needs and expectations so that they can successfully use them in their working practices.

Next section illustrates the data analysis problem faced by chemists. Section III discusses the attempts made to adapt existing and renowned tools to the chemists' problem, and the reasons they resulted not suitable for them. Section IV illustrates the adopted platform to support the analysis process and the plugins developed to adapt it to the chemists' needs. Section V concludes the paper providing some lesson learned from this experience.

II. THE ANALYSIS OF POLLUTION DATA

Today the air quality is one of the major problems considered by governments. There are different types of pollution that experts analyze. This paper refers to the work of chemists who detect the concentration in the air of specific chemical compounds by analyzing the smell perceived through electronic nose instrumentations, i.e. machines designed to detect and discriminate among complex odors using a sensor array [17]. This array consists of sensors that are treated with a variety of odor-sensitive biological or chemical materials. An odor stimulus generates a characteristic pattern (or smellprint) from the sensor array. Smellprints from known odors are used to construct a database and train a pattern recognition system so that unknown odors can subsequently be classified and identified. Thus, an electronic nose includes hardware components (different devices collecting odors) to collect and transport odors to the sensor array, and electronic circuitry to digitize and store the sensor responses for signal processing. Early prototypes of electronic noses date back to 70's, but only today, thanks to advances in microfabrication techniques, there are detector arrays with enhanced sensitivity and selectivity with characteristics comparable to those of humans.

The chemists addressed in this paper are researchers that are investigating about the causes of unpleasant smells in the air in order to prevent air pollution. They are not people that routinely check pollution data in order to communicate to government about critical situations for human safety, so that proper actions can be decided. Instead, they are searching for reasons of pollution and possible solutions to improve air quality. Their work is not restricted to odor detection, because multiple factors can generate unpleasant smells. In particular, it is important to correlate smell data with some weather parameters coming from weather stations. Weather information is very important to better understand behaviors of the observed chemical compounds because they are detected in the air. The most relevant weather parameters are wind speed and direction.

Although different companies are now producing electronic noses these systems are not yet totally reliable. Researchers want to check the responses of the electronic system with those provided by human beings. To this aim, they have selected sixteen housewives, resident in different locations of the territory, who are asked to note in a form every time they feel a bad smell in the air, annotating the time, the duration and the intensity of the smell. The ladies transmit such data via phone to a toll free number. Figure 1 shows the different ways of collecting data about air pollution, due, for example, to the emissions of a landfill. Besides the already mentioned data collected by the electronic nose, the housewives and detectors of weather parameters, data are also collected by PID (PhotoIonization Detector) and OPC (Optical Particle Counter), two more devices to collect information about pollution.

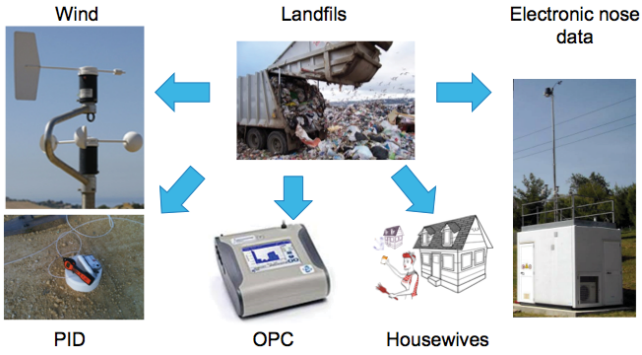


Figure 1. Sources of analyzed data

Different devices of the electronic nose, located in different stations on the territory, monitor chemical compounds in the air with different techniques, time granularity, and unit measures. Data are stored in logs files, which can be transferred via Internet. Chemists collect such log files, integrate and/or compare them with data coming from other sources (weather, housewives), apply transformations and analyze the resulting data. Before the beginning of this work, for their analysis, they used tools, such as MATLAB and Microsoft Excel, in order to make computations, and/or produce visualizations that may help them to understand and to explain what is going on, and to perform predictions.

III. PROBLEMS IN USING VISUALIZATION TOOLS

There are many visualization tools in literature with different purposes. Several authors have also provided some classifications. For example, Shneiderman has classified visualizations tools according to tasks that users have to accomplish and according to data types [20]. From the description provided in the previous section, it is evident that the data considered in this paper are time dependent. All considered data (coming from the electronic nose devices, weather parameters, data reported by the selected housewives) are associated to the time when they have been collected. Thus, they fall in the category of time series. In fact, a time series is a sequence of data points, measured typically at successive time instants spaced at uniform time intervals. Sometimes the constraint of uniformity in collecting time series data is

relaxed, but not in the domain presented in this work, so this is not addressed in this paper.

There is a wide literature in time series and in time dependent analysis. A survey is provided by Aigner et al. [1], who analyze a wide spectrum of key techniques in visualizing time oriented data, classifying each of the about one hundred identified techniques according to different dimensions; one refers to the main tasks users want to perform for analyzing time series data, namely: select, explore, reconfigure, encode, abstract/elaborate, filter, connect, undo/redo, and configuration change. These tasks match very closely the needs observed during the chemists' requirements analysis.

We setup a multidisciplinary team, which included the domain experts, in order to analyze different tools, looking for those that might better fit the experts' needs. The team considered various tools. One that was deeply discussed is LiveRAC [15]. One of it's main characteristics is that it displays on the screen many time series using a focus + context technique combined with semantic zoom [18]. Both are useful techniques that allow focusing on specific time series, keeping the context in which they are visualized, which is represented with much less details. However, chemists were not enthusiast about this kind of interface, since they found the screen too cluttered and, in their opinion, too many details were hidden by the visualization. Moreover, they found the side-by-side comparison of time series not useful for them because they are used to compare overlapping graphs.

Based on their comments, TimeSearcher was analyzed [5], [6]. TimeSearcher provides users with interactive visualization of multivariate time series in a single screen. For example, the screen in Figure 2 shows 8 panels, each displaying time series of a variable (e.g. temperature) whose values are collected at different locations. Panels are vertically juxtaposed for an easy comparison among variables. Data are aligned according to the time, and a vertical blue line (on the left in each panel of Figure 2) indicates this alignment. TimeSearcher supports explorative analysis of time series, to this aim it allows user to search for interesting patterns in the series and to filter according to some specific data values and time intervals. A detailed demo was shown to the domain experts, during which the features of TimeSearcher were discussed. The experts appreciated the tool and, in particular, the possibility to compare different time series in the same panel in order to easily identify correlations. For example, as shown in Figure 2, the interface of TimeSearcher has 8 panels; each panel visualizes several time series of the same variable, e.g. temperature, recorded at several monitoring stations.

The chemists decided to adopt this tool, since they were convinced that it was much better than spreadsheets in permitting visual interaction with data, thus supporting a more effective analysis. In order to manage the different formats of the data collected from the heterogeneous sources, a software module has been developed to convert the data in the format required by TimeSearcher. This module requires a human's intervention to solve some ambiguities in the data. This process required time, but this was not considered a problem since data conversion did not occur so frequently.

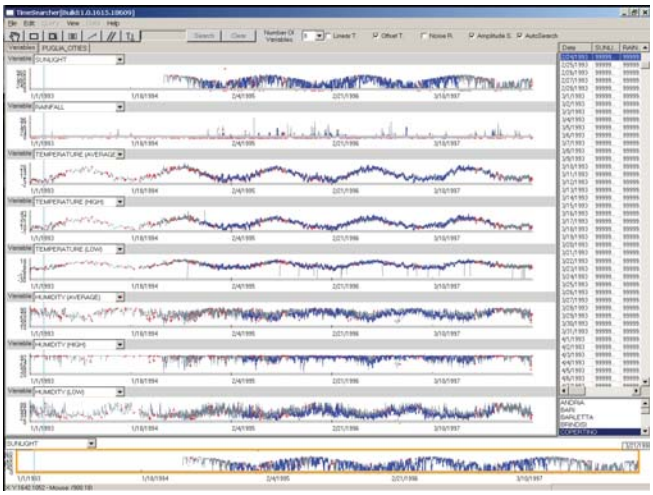


Figure 2. A screenshot of TimeSearcher with 8 panels, each displaying time series of a variable, whose values are collected at different locations.

Chemists used TimeSearcher, appreciating most of its features. However, pretty soon some of them complained about TimeSearcher, because they primarily considered a different type of analysis, i.e. they concentrated on data of a single monitoring station with more than 12 devices, each producing data about a specific variable. Such chemists wanted to plot values of these variables in the same panel. TimeSearcher does not permit this. It is worth noticing that this requirement never became explicit during the discussions with chemists and their trials with demos of TimeSearcher usage. It actually emerged only once they worked with the system during their daily activities. This finding has been already highlighted in [24] and experienced also in [3]: "end users provide the most valuable feedback about the possible problems once they get to work with the new system in real settings".

Another drawback emerged, this time due to a new requirement generated from the evolution of the chemists' activities. In fact, these chemists are experts who carry out research work related to the air quality. The data they work with are not collected routinely, as done in centers devoted to the air pollution control. However, at a certain point they started to collect data much more frequently than before, and this implied to perform much more often the data conversion process to adapt data to TimeSearcher. As we said, this process needs the involvement of a researcher. Thus, chemists very soon got annoyed of repeating the data conversion of new data so many times. As a consequence, they stopped using TimeSearcher and went back to their original work with MS Excel and MATLAB that, they also said, allowed them to

perform many other computations required from their recent research directions.

These new requirements shifted the interest of the participatory team to the data pre-processing phase, also because they started collecting data from other devices, namely PID and OPC (see Figure 1), which were not considered before. As it will be shown in the next session, data preprocessing is one of the main phases of the overall data analysis process because, especially when data come from different sources and are in different formats, as in our case, data have to be manipulated to get them in a form usable by the visualization tools, and check their quality.

IV. The Analysis Process with Knime

Our activity focused on the preprocessing phase due to the new requirements of the chemists, who often had to analyze new data. Preprocessing includes all the necessary steps to prepare data for the analysis. Some authors now refer to these steps as data wrangling, which is considered one of the big issues for a meaningful data analysis. Data wrangling is defined as "a process of iterative data exploration and transformation that enables analysis" [12]. In fact, an analyst has to first diagnose the data to make sure that they are usable for the analysis questions, to see the efforts required to put them into an appropriate format for the analysis tools, to check their quality, i.e. if there are missing data, inconsistent values, unresolved duplicates, etc. In other words the data have to be transformed and cleaned into a usable state. Data wrangling is actually the process to make data useful.

Numerous techniques for cleaning and integrating data have been proposed within the database community (for example see [9],[13],[19]). However, many of them focus on specific data quality problems, are not interactive, or are not accessible to a general audience. One of the current challenges in this field is to enrich data processing technologies with innovative visual interfaces that support data diagnostics and transformations [12]. Going back to the pollution data analysis, we realized that we needed tools for a more accurate data preprocessing and that we had to shift our attention to the overall analysis process.

A. *The Knime platform*

We looked at platforms that support this analysis process, since several are now available. However, most of them concentrate on some phases of the process. We found that Knime (Konstanz Information Miner) provides a working environment that supports users in their activities by providing several tools useful for the phases of data preprocessing, analysis, and exploration [7]. Developed at the University of Konstanz, it is open-source and based on Eclipse.

In Knime, the user creates the workflow for his/her overall data analysis process. Figure 3 shows an example. The user has composed the workflow by selecting, from a menu showing all available tools, those necessary for his/her analysis. Each node represents a step of the process; in other words, a node is a plugin that performs a specific activity. As shown in Figure 3, the node File Reader, on the left of the figure, loads data in the environment. Then data are transferred to the K-Means node, which creates clusters of data. Note that the little blue square on the right of a node means that the node is able to display data. The white arrows represent communication ports; if the arrow points toward the node it is an input port, otherwise it is an output port. K-Means node is connected to a Color Manager node, which associate colors to the clusters created by the previous node. The Color Manager node is connected to two other nodes that visualize data using different techniques, according to colors set in Color Manager. In Figure 3 the user is visualizing the menu associated to the Scatter Plot node and is about to press the "View: Scatterplot" menu item. In the example of Figure 3 all nodes are marked with a green light (at the bottom of the node icon) because the analysis is finished without errors. Each node can be in one of the following three states: to be configured (red light), ready to run (yellow light), node processed (green light).

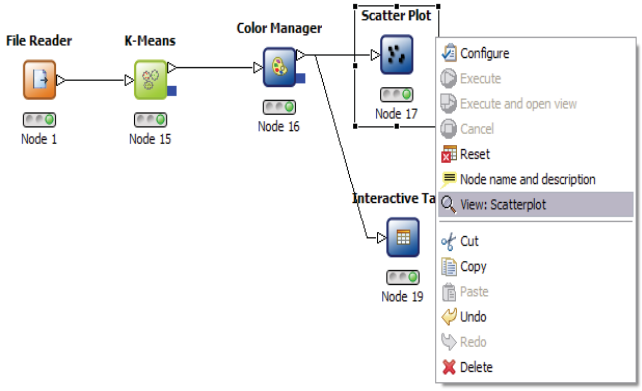


Figure 3. A Knime workflow with five nodes. File reader loads a data file and sends data to the K-Means clustering algorithm. The Color Manager node associates colors to clusters and the output is visualized through both a scatter plot and an interactive table.

Working with Knime, our participatory team soon realized that, even if Knime includes many tools, it was necessary to develop other plugins that could process the specific data of the chemists' problem. In order to create as soon as possible a running prototype with which chemists could work, we developed three plugins. These are:

1) Multifile reader, which loads many files containing data produced by a single data source (like those represented in Figure 1) and merges them according to the time of sampling;

2) Nose-wind converter, which merges electronic nose data with weather parameters;

3) PID/OPC converter, which merges PID and OPC data with weather parameters.

B. *Preliminary formative test*

A user test was performed as part of our formative evaluation. This refers to evaluation activities that are carried out during the design and development cycle of a software product in order to get feedback that it is very useful in "forming" the product, so that it can really fit users' needs and requirements. Sometimes only inspection methods are used in formative evaluation, since tests with users are more resource demanding. However, our experience is a further indicator that formative evaluation requires that users test the prototypes, especially in cases, like ours, in which the system is complex and its target users are professional people, who are very demanding in terms of functionality the system provides and expect a good mapping of the interaction strategies available in the system with their habits and ways of solving their problems.

The user test was conducted using the thinking aloud method, which is well known and appreciated for providing very useful results without requiring many resources [16]. Four domain experts participated. They had experience in the use of their traditional analysis tools, i.e. MS Excel and MATLAB. They were asked to perform some preprocessing activities with Knime so that they had to use the developed plugins. Specifically, the tasks covered activities ranging from the data acquisition to the data visualization.

The test took place in the participants' office. The four participants were first together in a session during which one of the computer scientists did a short presentation of the platform (see Figure 4). Then each participant performed the test in her/his office, sitting at a desk with two evaluators: the facilitator and the observer. A laptop was used for the test, equipped with a screen recorder, which recorded the interaction performed by the user, and a webcam to observe the user's reactions. In accordance with the thinking aloud method, the participant was asked to "think aloud", in order to get the most information on the strategies employed by the participant to accomplish the tasks and better understand any possible difficulty.

Figure 4. Training phase of the test

In Figure 5 a test session is shown. The role of the facilitator, sitting next to the participant, is to assist her in the event of difficulties and encourage to move forward in testing, but without interfering with the task execution, and without providing or requesting information that can influence participant's behavior. The observer was sitting next to the facilitator and his role was to observe the test, also annotating in a table the time taken for each task. After the test, each participant was interviewed in order to get qualitative feedback.

Figure 5. A participant performing a task

To the purpose of this paper, it is useless to report the details of the test results. It is enough to say that the chemists easily understood how to work with Knime and appreciated its utility for their work, especially because it allowed them to better understand the overall process and to easily repeat the preprocessing activities. They also complained about some usability problems of Knime that, however, will not compromise the use of the platform because they can be solved by providing some initial training. In this case, this is not so bad since the chemists will use the system very frequently and they will become soon expert users.

V. CONCLUSIONS

This paper presents our experience of working with chemists in a multidisciplinary team for developing visual tools that support the analysis of pollution data. The experience indicated some problems that have to be solved for creating visual analytics tools that are useful and usable for their intended users. A first problem is that it is very difficult to adapt visual tools that are claimed to be general in contexts that take into account very complex and very specific problems, like the one addressed in our case study. Another problem is that data preprocessing is extremely important in visualization systems and need special care [12]. The two problems refer to the development of any interactive system, and not only to visual analytics systems: 1) it is extremely hard for the users to communicate their requirements, even if a participatory design approach is adopted; 2) the only way to evaluate if a software system is useful and usable for its users is to let them to put their hand on prototypes.

Our suggestion to overcome these problems is to perform evaluations with methods involving end users, observing them when they are working with running prototypes of the system, as the user test described in this paper. We plan to go even further, by requiring longer sessions of use of the prototype by testers while working in their daily activities. This is in line with other findings already provided in literature [3],[24].

Future work could consider the design of user interfaces that support the analysis process on mobile devices. Indeed, it could be useful to perform some analysis steps at the locations where the devices are installed. This poses a great challenge in data visualization due to the limited size of the screen, which requires the adoption of specific visualization techniques, such as overview + details and semantic zoom (see, for example, [2], [4], [8]).

REFERENCES

[1] W. Aigner, S. Miksch, H. Schumann, and C. Tominski. Visualization of Time-Oriented Data. Springer-Verlag New York, 2011.

[2] C. Ardito, P. Buono, Costabile M., R. Lanzilotti (2006). Two Different Interfaces to Visualize Patient Histories on a PDA. Proc. of MobileHCI 2006, Espoo, Finland. September 12-15, 2006. New York: ACM Press (United States), pp. 37–40.

[3] C. Ardito, P. Buono, M. F. Costabile, R. Lanzilotti, A. Piccinno, A. L. Simeone. Analysis of the UCD process of a web-based system. Proc. of DMS 2010. Oak Brook, Illinois, USA, 14-16 October, 2010, Skokie, Illinois: Knowledge Systems Institute, pp. 180–185.

[4] B. B. Bederson, A. Clamage, M. P. Czerwinski, and G. G. Robertson. DateLens: A fisheye calendar interface for PDAs. ACM Trans. Computer-Human Interaction, vol 11, no. 1, 2004, pp. 90–119.

[5] P. Buono, A. Aris, C. Plaisant, A. Khella, B. Shneiderman. Interactive Pattern Search in Time Series. Visualization and Data Analysis. Proc. of VDA 2005, 16-20 January 2005, San Jose, CA USA, SPIE, Washington DC, pp. 175–186.

[6] P. Buono, C. Plaisant, A. Simeone, A. Aris, G. Shmueli, and W. Jank. Similarity-based forecasting with simultaneous previews: A river plot interface for time series forecasting. Proc. of IV 2007, Washington, DC, USA, 2007. IEEE Computer Society, pp. 191–196.

[7] N. Cebron and M. R. Berthold. Adaptive active classification of cell assay images. Knowledge Discovery in Databases: PKDD 2006, volume 4213, Springer Berlin / Heidelberg, 2006, pp. 79–90.

[8] L. Chittaro. Visualizing Information on Mobile Devices. Computer vol. 39, no. 3, 2006, pp. 40–45.

[9] A.K. Elmagarmid, P.G. Ipeirotis, and V.S. Verykios. Duplicate record detection: A survey. IEEE Trans Knowl Data Eng 2007; 19(1), pp. 1–16.

[10] International Organization for Standardization, ISO 13407: Human-Centered Design Process for Interactive Systems, 1999.

[11] Jigsaw. http://www.cc.gatech.edu/gvu/ii/jigsaw/. Last accessed: April 2012.

[12] S. Kandel, J. Heer, C. Plaisant, J. Kennedy, F. vanHam, N. H. Riche, C. Weaver, B. Lee, D. Brodbeck, and P. Buono. Research directions in data wrangling: Visualizations and transformations for usable and credible data. Information Visualization Journal, vol. 10, no. 4. 2011. SAGE, pp. 271–288.

[13] H. Kang, L. Getoor, B. Shneiderman, M. Bilgic, and L. Licamele. Interactive entity resolution in relational data: A visual analytic tool and its evaluation. IEEE Trans Visual Comput Graph 2008; 14, pp. 999–1014.

[14] D.A. Keim, J. Kohlhammer, G. Ellis, F. Mansmann. (Eds), Mastering the Information Age - Solving Problems with Visual Analytics. Eurographics Association. 2005.

[15] P. McLachlan, T. Munzner, E. Koutsofios, and S. North. Liverac: in - teractive visual exploration of system management time-series data. Proc. of CHI 2008, New York, NY, USA, 2008. ACM, pp. 1483–1492.

[16] J. Nielsen. Usability Engineering, Academic Press, San Diego, 1993.

[17] T.C. Pearce, S.S. Schiffman, H.T. Nagle, J. W. Gardner. (Eds) Handbook of Machine Olfaction: Electronic Nose Technology. Weinheim: Wiley-VCH Verlag. 2003.

[18] R. Rao, and S. K.Card, The table lens: merging graphical and symbolic representations in an interactive focus+context visualization for tabular information. Proc. of CHI '94. ACM, New York, NY, USA. 1994, p. 222.

[19] G.G. Robertson, M. P. Czerwinski, and J. E. Churchill. Visualization of mappings between schemas. Proc. of CHI 2005, pp. 431–439.

[20] B. Shneiderman. The eyes have it: A task by data type taxonomy for information visualizations. Proc. of Visual Languages (Boulder, CO, Sept. 3–6). IEEE Computer Science Press, Los Alamitos, CA, 1996, pp. 336–343.

[21] Spotfire. http://spotfire.tibco.com/. Last accessed: April 2012.

[22] Tableau. http://www.tableausoftware.com/. Last accessed: April 2012.

[23] D. Schuler, A. Namioka, Participatory Design: Principles and Practices, Lawrence Erlbaum Associates, Inc., 1993.

[24] J.L.Wagner, G. Piccoli, Moving beyond user participation to achieve successful IS design. Communications of the ACM, vol. 50, no. 12, 2007, pp. 51–55.

Ontological Filtering for Sentiment Analysis

Francesco Colace[1], Massimo De Santo[1], Paolo Napoletano[1], C. Becchi[2], S.K. Chang[3]

[1] Department of Electronic and Information Engineering, University of Salerno, 84084 Fisciano, Italy
{fcolace, desanto, pnapoletano}@unisa.it
[2] 77Agency, Dzirnavu 57A - 1, 1010, Riga, Latvija
c.becchi@77agency.com
[3] Computer Science Department – University of Pittsburgh, Sennott Square, Pittsburgh, PA 15260 USA
chang@ cs.pitt.edu

Abstract

The rapid growth of the internet has increased the number of online reviews, opinions and sentiments toward products, services or topics. People appreciate this opportunity so that e-Commerce websites provide services for users to publish their reviews. Social networks, blogs and websites enable, thanks to the reviews, a social structure that provides benefits for the users and the firms that hosts electronic markets. Therefore, this huge quantity of information can confuse users and does not produce useful knowledge. In such a context, in fact, who says what and how they say it, matters. In this scenario a valuable contribute can be given by the sentiment analysis that is one of the hottest current research area. This paper presents a novel approach to the sentiment analysis which is based on the ontological filtering approach. The proposed approach shows how to automatically mine, from a corpus of documents, positive and negative sentiments .Experimental evaluations, on real dataset, show that the proposed approach is effective and furnishes interesting results.

1. Introduction

During a decision making process people have to consider many things but there is a moment when usually think: "what other people think". Before the wide diffusion of the internet and web2.0 services people share opinions and recommendation by the use of traditional approaches: asking to friends, talking with experts and reading documents. The internet and the web have made it possible to find out about the opinions and experiences of those in the vast pool of people that are neither our personal acquaintances nor well known professional critics. The interest that individual users show in online opinions about products and services and the potential influence such opinions wield, is something that vendors of these items are paying more and more attention to. Companies are interested in what users are saying about their products as politicians are interested in how different news media are portraying them. Therefore there is a lot of information on the web that have to be properly utilized in order to provide vendors highly valuable network intelligence and social intelligence to facilitate the improvement of their business. In this scenario a very interesting approach is the sentiment analysis. More in general sentiment analysis is the computational study of opinions, sentiments and emotions expressed in text [1]. Sentiment classification is part of the opinion mining and refers to the identification of opinions and arguments in a text. Its main aim is the identification of the agreement or disagreement statements that deal with positive, negative or neutral in comments or reviews. There are many approaches to the sentiment analysis. A very broad overview of the existing work was presented in [2]. The authors describe in a very detailed way the main techniques and approaches for an opinion oriented information retrieval. Early work in this area was focused on determining the semantic orientation of documents. In particular some approaches attempt to learn a positive-negative classifier at the document level. In [3] three machine learning approaches (Naïve Bayes, Maximum Entropy and Support Vector Machines) has been adopted to label the polarity of a movie reviews datasets. A promising approach has been developed in [4] where a novel methodology has been obtained by the combination of rule based classification, supervised learning and machine learning. In literature some approaches are based on a computational approach to inferring the sentiment orientation of social media content and estimate sentiment orientations of a collection of documents as a text classification problem [5]. More in general according to these approaches sentiment related information can be encoded lexically within the actual words of the sentence syntactically and morphologically through changes in attitudinal shades of word meaning using suffixes [6]. This approach has been investigated in [7] where a lexicon for sentiment analysis has been obtained. Another interesting approach is in [8] where a probabilistic approach to sentiment mining has been adopted. In particular this approach adopt a probabilistic model called Sentiment Probabilistic Latent Semantic Analysis (S-PLSA) in which a review, and more in general a document, cab be considered as being generated under the influence of a number of hidden sentiment factors [9]. The S-PLSA is an extension of the PLSA where it is assumed that there are a set of hidden semantic factors or aspects in the documents that are related to documents and words

under a probabilistic framework. In [10] an approach combining the ontological formalism and a machine learning technique has been proposed. In particular the proposed system starting from a sentence uses domain ontology to extract the related concepts and attributes and then by the use of the Support Vector Machine (SVM) classifier for labelling it positive, negative or neutral. The last two papers are the starting point for the approach proposed in this paper. We propose a methodology for the sentiment analysis based on the combined use of probabilistic techniques (Latent Dirichlet Allocation and ontological formalism. In this way the document can be automatically filtered according their sentiment value.
The rest of the paper is organized as follows. Section 2 provides a brief introduction to the ontological filtering. The section 3 discusses the Latent Dirichlet Allocation and its use in the ontology building process. The section 4 shows the proposed sentiment analysis approach based on the use of the ontological filtering and the section 5 discusses the experimental results. The conclusions section concludes this paper.

2. Ontological Filtering

As previously said this section introduces the concept of ontological filtering. In order to introduce this concept an example can be showed. A user aims to make a search in a web repository that refers to a well-defined domain. Usually the user expresses his query by the use of keywords. In this way the search is conducted using the keywords and counting their occurrences (mutual or not) in the documents. This is a simple syntactic search. The ontological filter, besides, works in the following way: the user still search for by the use of keywords, but this time the keywords can be considered attributes of concepts that are in ontology. In fact, we can assume that each keyword can be associated to various concepts. For example, in ontology related to the sport the keyword "ball" could be associated to various kinds of sports. In this way, by the use of keywords, we can navigate the domain through the topics that are in. In other words by the use of the keywords we can obtain a lightweight ontology which can be considered a sort of sieve for the searching of documents. In fact, in this way the system can give to the user only the documents that belong to certain concepts that are in this new ontology. In this case also documents that do not contain keywords but are related to the concepts that contain ones will be selected. This approach can be easily generalized to each kind of problem. For example in the case of a video surveillance system, each component can be part of some topics and can be identified by some keywords. In this way the request of the user can be easily mapped as a set of topics belonging to ontology and an answer to a user request. The ontology filtering approach introduces a new layer the interaction between users and data represented by the ontology. A more formal definition of the Ontology Filtering can be developed in the following way. First of all, ontology representing the domain of the knowledge involved in the problem has to be defined. This ontology can be obtained both by the support of experts both by the use of automatic methodologies able to infer the description of the domain by the analysis of related data, as we will show in the section 3 of this paper. In this way the $O = \{C, A, H, R_T, R\}$ (where C is the set of concepts, A the set of the attributes, H the set of the hierarchical relationships, R_T the set of non-hierarchical relationships and R the set of semantic relationships) can represent the domain of interest and in particular the A attributes set contains both the set of keywords and the "items". The user introduces its query by the use of keywords $K_U=\{k_{ui}\}$. In this way the following strategy could be followed in order to build the ontology:

Step 1: $\forall k_{ui} \in K_U$ and $\in A_{Ci}$ add in the set C' of the Ontology $O'_{L+} = \{C', A', H', R'_T, R'\}$ the concept C_i belonging to an ontology $O_{L+} = \{C, A, H, R_T, R\}$. The set A_{Ci} contains the attributes related to the concept C_i and is a subset of A.

Step 2: $\forall C_i \in C'$ add in the set C' of the Ontology $O'_{L+} = \{C', A', H', R'_T, R'\}$ all concepts C_j that share relations of kind H, R_T and R with the nodes C_i. At the same time these relations have to be inserted in the sets H', R'_T and R' and all the attributes of the various C_j

At the end of these two steps the $O' = \{C', A', H', R'_T, R'\}$ is the ontology that expresses the domain according to the requests of the user. In this way the ontological filter can be developed: in fact all the "objects" that are in a complete domain can be filtered by the use of the ontology O'_{L+} obtaining only the objects that user needs. This approach is able to answer the request of the user determining in an automatic way all the "items" that need for the resolution. Thanks to the relations, besides, that are among the various concepts that are in the ontology a use order can be inferred. The problem is how to infer in an automatic way the ontology of a domain. In the next section an approach for the unsupervised ontology learning by the use of the Latent Dirichlet Allocation will be introduced.

3. Ontology Extraction by the use of LDA

In statistics, latent Dirichlet allocation (LDA) is a generative model that allows sets of observations to be explained by unobserved groups that explain why some parts of the data are similar. For example, if observations are words collected into documents, it posits that each document is a mixture of a small number of topics and that each word's creation is attributable to one of the document's topics. LDA is an example of a topic model and was first presented as a graphical model for topic discovery in [10]. In general this approach is based on the concept of query expansion in a vector of features. The vector of features

needed to expand the query is obtained as a result of an interactive process between the user and system. The user initially performs retrieval by inputting a query into the system and later identifying a small set $\mathcal{D}$ of relevant documents from the hit list of documents returned by the system, which is considered as the training set (the relevance feedback). Existing query expansion techniques principally use the relevance feedback of both relevant and irrelevant documents. Usually they obtain the term selection through a scoring function which assigns a weight to each term depending on its occurrence in both relevant and irrelevant documents. In contrast, in this paper we do not consider irrelevant documents and the vector of features extraction is performed through a method based on a supervised Term Clustering technique. Precisely, the vector of features, that we call mixed Graph of Terms, can be automatically extracted from a set of documents $\mathcal{D}$ using a method for term extraction based on a supervised Term Clustering technique [11] weighted by the Latent Dirichlet Allocation [10] implemented as the Probabilistic Topic Model [12]. The graph is composed of a directed and an undirected sub-graph (or level). We have the lowest level, namely the word level, that is obtained by grouping terms with a high degree of pairwise semantic relatedness; so there are several groups (clusters), each of them represented as a cloud of words connected to their respective centroids (directed edges), alternatively called concepts. In addition, we have the second level, namely the conceptual level, obtained by inferring semantic relatedness between centroids, and so between concepts (undirected edges). The final result is a mixed graph of terms (m$\mathcal{GT}$) as depicted in fig. 1. The general idea of this technique is supported by previous works [13] that have confirmed the potential of supervised clustering methods for term extraction, even in the case of query expansion [14], [15].

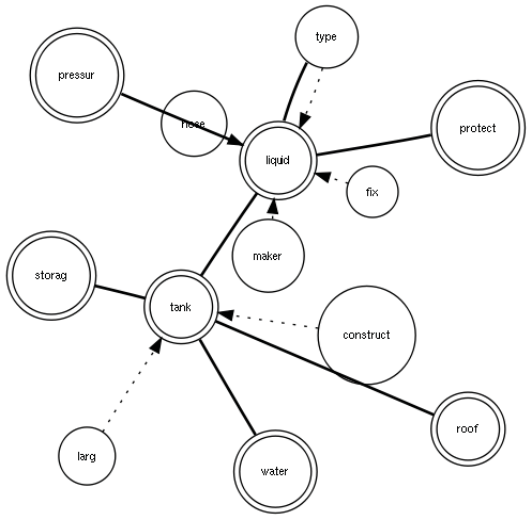


Figure 1 An Example of mGT

3.1. Extracting a mixed Graph of Terms

A mixed *Graph of Terms* (m$\mathcal{GT}$) is a hierarchical structure composed of two levels of information represented through a directed and an undirected sub-graph: the conceptual and word levels. We consider extracting it from a corpus $\mathcal{D}=\{\mathbf{w}_1,\mathbf{w}_2,\cdots,\mathbf{w}_M\}$ of M documents (that we call the training set), where each document is, following the Vector Space Model [16], a vector of feature weights $\mathbf{w}_j=(w_{1j},\ldots,w_{|\mathcal{T}|j})$, where $\mathcal{T}=\{t_1,\cdots,t_{|\mathcal{T}|}\}$ is the set of features that occur at least once in at least one document of $\mathcal{D}$, and $0\le w_{kj}\le 1$ represents how much the feature t_k contributes to the semantics of document $\mathbf{w}_j$. We choose to identify features with words, that is the bags of words assumption, and in this case $t_k=v_k$, where v_k is one of the words of a vocabulary $\mathcal{T}$. The word level is composed of a set of words v_s that specify through a directed weighted edge the concept c_i (see fig. 1(b), tab. 1 and fig. 2(a)), or better the centroid of such a set (group or cluster), which is, therefore, still lexically denoted as a word. The weight γ_{is} can measure how far a word is related to a concept, or how much we need such a word to specify that concept, and it can be considered as a probability: $\gamma_{is}=P(c_i\mid v_s)$. The resulting structure is a sub-graph rooted on c_i. Alternatively, the conceptual level is composed of a set of interconnected, through undirected weighted edges, concepts c_i (see fig. 1), so forming a sub-graph of pairs of centroids. The weight ω_{ij} can be considered as the degree of semantic correlation between the two concepts and it can be considered as a probability: $\omega_{ij}=P(c_i,c_j)$.

3.1.1. Graph drawing

A m$\mathcal{GT}$ is well determined through the learning of the weights, the Relation Learning, and through the learning of the three parameters, the Parameter Learning, that is $\Theta=(H,\Gamma,\Omega)$ which specifies the shape of the graph. In more details, we have:

1. H: the number of concepts (namely the number of clusters) of the corpus $\mathcal{D}$;
2. γ_i: the threshold that establishes for each concept the number of edges of the directed sub-graph, and so the number of concept/word pairs of the corpus $\mathcal{D}$. An edge between the word s and the concept i can be saved if $\gamma_{is}\ge\gamma_i$. We consider, to simplify the formulation, $\gamma_i=\Gamma$, $\forall i$;
3. Ω: the threshold that establishes the number of edges of the undirected sub-graph, and so the number of concept/concept pairs of the corpus $\mathcal{D}$. An edge between the concept i and concept j can be saved if $\omega_{ij}\ge\Omega$.

3.1.2. Relations Learning

Due to the fact that each concept is lexically represented by a word of the vocabulary, then we have

that $\square_{is} = P(c_i \mid v_s) = P(v_i \mid v_s)$, and $\square_{ij} = P(c_i, c_j) = P(v_i, v_j)$. As a result, we can obtain each possible relation by computing the joint probability $P(v_i, v_j)\ \forall i, j \in \mathcal{T}$, which can be considered as a word association problem and so can be solved through a smoothed version of the generative model introduced in [10] called Latent Dirichlet allocation, which makes use of Gibbs sampling [12].

3.1.3. Parameters Learning

Given a corpus $\mathcal{D}$, once each $\square_{ij}$ and $\square_{is}$ is known $\square i, j, s$, letting the parameters assume a different set of values $\square_t$, we can observe a different graph $m\mathcal{GT}_t$, where *t* is representative of different parameter values.

A way of proving that a $m\mathcal{GT}$ is the best possible for that set of documents is to demonstrate that it produces the maximum score attainable for each of the documents when the same graph is used as a knowledge base for querying in a set containing just those documents which have fed the $m\mathcal{GT}$ builder.

Each graph $m\mathcal{GT}_t$ can be represented, following again the *Vector Space Model* [16], as a vector of feature weights, that we call $\mathbf{q}_t$ and is defined as $\mathbf{q}_t = (w'_{1t}, \ldots, w'_{|\mathcal{T}_p|t})$, where $|\mathcal{T}_p|_t$ represents the total number of pairs. We have that each feature $t_k = (v_i, v_j)$, which is not the simple bags of words assumption, and w_{kj} being the weight calculated thanks to the *tf-idf* model applied to the pairs represented through t_k, and with the addition of the *boost* b_k which is the semantic relatedness between the words of each pair, at both the conceptual and the word level, namely $\square_{ij}$ and $\square_{is}$. You will recall that both $\square_{ij}$ and $\square_{is}$ are real values (probabilities) of the interval [0,1], and so to distinguish the relevance between the three cases, the traditional case ($b_k = 1$), the concept/word pair and the concept/concept pair, we have distributed such values with a wider interval. Specifically:

1. $b_k = 1$ being the lowest level of relatedness;
2. $b_k \square [\square_{min}, \square_{max}]$ with $\square_{min} \square 1$ and $(\square_{max} \square \square_{min}) = 1$;
3. $b_k \square [\square_{min}, \square_{max}]$ with $\square_{min} > \square_{max}$ and $(\square_{max} \square \square_{min}) = 1$

.

In the experiments we have chosen $\square_{min} = 1$ and $\square_{min} = 3$ (see table 1).

Conceptual Level		
Concept *i*	**Concept *j***	**Relation Factor ($\square_{ij}$)**
tank	roof	4,0
tank	water	3,37246
tank	liquid	3,13853
...	...	...
liquid	type	3,43828
liquid	pressur	3,07028
...	...	...
Word Level		
Concept *i*	**Word *s***	**Relation Factor ($\square_{is}$)**
tank	Larg	2,0
tank	construct	1,6
...	...	...
liquid	type	1,21123
liquid	maker	1,11673
liquid	hose	1,06024
liquid	Fix	1
...	...	...

Table 1. An example of a *mGT* for the topic *Storage Tank*.

At this point, a document $\mathbf{w}_j$ can be viewed as a vector of weights in the space $|\mathcal{T}_p|_t$, and so the general formula of each weight is:

$$w'_{kj} = \frac{\text{tf-idf}(t_k, \mathbf{w}_j) \cdot b_k}{\sqrt{\sum_{s=1}^{|\mathcal{T}_p|_t} (\text{tf-idf}(t_s, \mathbf{w}_j) \cdot b_k)^2}} \qquad (1)$$

The score for each graph at time *t*, namely $\mathbf{S}_t$, can be computed following the cosine similarity model in the space $|\mathcal{T}_p|_t$, and so we have

$$\mathbf{S}_t(\mathbf{q}_t, \mathbf{w}_j) = \frac{\sum_{k=1}^{|\mathcal{T}_p|_t} w'_{kj} w'_{kt}}{\sqrt{\sum_{k=1}^{|\mathcal{T}_p|_t} w'^2_{kj}} \sqrt{\sum_{k=1}^{|\mathcal{T}_p|_t} w'^2_{kt}}} \qquad (2)$$

Finally for the graph at time *t* we have a score for each document, $\mathbf{S}_t = \{\mathcal{S}(\mathbf{q}_t, \mathbf{w}_1), \cdots, \mathcal{S}(\mathbf{q}_t, \mathbf{w}_M)\}_t$.

As a result, to compute the optimum set of parameters $\square_t$ we can maximize the *Fitness* ($\mathcal{F}$),

$$\Lambda^* = \underset{\Lambda_t}{\text{argmax}}\{\mathcal{F}(\Lambda_t)\},$$

where $\mathcal{F}(\Lambda_t) = E_m\left[\mathcal{S}(\mathbf{q}_t, \mathbf{w}_m)\right] - \sigma_m\left[\mathcal{S}(\mathbf{q}_t, \mathbf{w}_m)\right]$. E_m is the mean value of all elements of $\mathbf{S}_t$ and $\square_m$ is the standard deviation. Since the space of possible solutions could grow exponentially, we have limited the space to $|\mathcal{T}_p|_t < 150$, $\square t$. Furthermore, we have reduced the remaining space of possible solutions by applying a clustering method, that is the *K-means* algorithm, to all $\square_{ij}$ and $\square_{is}$ values, so that the optimum solution can be exactly obtained after the exploration of the entire space. This reduction allows us to compute a $m\mathcal{GT}$ from a repository composed of a few documents in a reasonable time (e.g. for 3 documents it takes about 3 seconds with a Mac OS X based computer and a 2.66 GHz Intel Core i7 CPU and a 8GB RAM). It is important to make clear that the mixed *Graph of Terms* cannot be considered as a co-occurrence matrix. In fact, the core of the graph is the probability $P(v_i, v_j)$, which we regard as a word association problem, which in the topic model is considered as a problem of prediction: given that a cue is presented, which new words might occur next in that context? It means that the model does not take into account the fact that two words occur in the same document, but that they occur in the same document

when a specific topic (and so a context) is assigned to that document [12]. Furthermore, in the field of statistical learning, a similar structure has been introduced, with the name Hierarchical Mixture of Experts [17]. Such a structure is employed as a method for supervised learning and it is considered as a variant of the well-known tree-based methods. The similarity between such a structure and the proposed graph can be obtained by considering the "experts" as "concepts". Nevertheless, the mixed Graph of terms is not a tree structure, and more importantly is not rigid but is dynamically built depending on the optimization stage. Moreover, the Hierarchical Mixture of Experts does not consider relations between experts which are, on the other hand, largely employed in the mixed Graph of Terms. Notwithstanding this, we will explore further connections between these two structures in future works.

4. A Sentiment Analysis Approach based on the use of the ontological filtering

In this section the proposed approach will be introduced and described in details. As previously said the methodology aims to combine the LDA and the ontological filtering approaches in order to realize an automatic sentiment detector able to label the mood of a document or a collection of them. The proposed approach is composed by the following steps:

- the first step is composed by the building process of ontologies representing a set of documents labeled according their sentiment. In other words by the use of the LDA approach the ontologies representing the positive, negative and neutral sentiment documents. The process is depicted in figure 2. These ontologies, that will name mood ontologies, contain the concepts, the words and their combination representing positive/negative and neutral sentiments. Thanks to the LDA approach it is possible to achieve representative ontologies by the use of few documents.

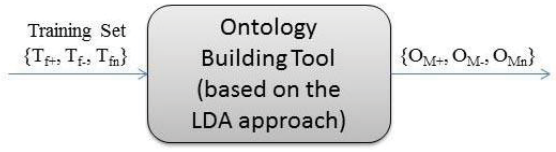


Figure 2 The Ontology Mood Building tool

- the second step is based on the introduction of a tool that uses the obtained ontologies as ontological filters. In particular this tool accepts as input a set of documents (or a single document) and labels it according the detected sentiment (positive/negative/neutral) by the use of the ontological filtering (figure 3).

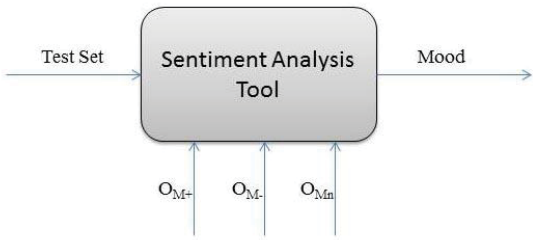


Figure 3 The Sentiment Analysis Tool

In particular the sentiment is detected according to the following formula:

$Mood_i$= 1 if $f(D_i, O_+)$> $|f(D_i, O_-)|$ and $f(D_i, O_+)$> $f(D_i, O_N)$

$Mood_i$= -1 if $|f(D_i, O_-)|$> $f(D_i, O_+)$ and $|f(D_i, O_-)|$> $f(D_i, O_N)$

$Mood_i$= 0 in the other cases

where $Mood_i$ is the sentiment associated to the document D_i. The function f detects the sentiment of the document in this way:

1. set f = 0
2. for each word that is in the document that matches a concept in the ontology increase of 2 the value
3. for each word that is in the document and in the ontology increase of 1 the value
4. for each word that is not in the ontology a research in a lexical dictionary, as Wordnet [18] for English language or IWN [19] for Italian language, has to be conducted. If there is a link between the word belonging to the document and a word or a concept increase of 1 the value of the function f

The overall mood is obtained in the following way.

$$\text{Mood} = \sum_{i=1}^{k} Mood_i$$

If Mood is >0 the overall mood of the documents' set is positive else if it is <0 the mood is negative. If the mood is 0 the overall sentiment is neutral. In order to assess the degree of positivity and negativity the mood can be normalized on the overall number of documents.

$$SD = |Mood|/\#Documents$$

SD can assume values in the range [0, 1]. Values close to 1 mean high degree of positive or negative sentiments while values close to 0 mean very low sentiment in the documents.

5. Experimental Results

To evaluate the proposed sentiment classification approach an experimental campaign on a real dataset has been conducted. The dataset has been organized

collecting post on the official blog of Samsung on the product Samsung Galaxy Tab. The dataset has been composed by 2.000 posts in Italian language and has been collected in three months (November 2011 – January 2012). The posts usually are a series of comments on the product and its evolution. An example of post is the following (that is one of the longest): *"Un regalo meraviglioso...uno stupendo GALAXY TAB :-D ed è una vera figata.....lo consiglio a tutti: semplice nell'uso, ottimo nelle prestazioni e mi permette di fare tutto quello che voglio senza le assurde limitazioni che mette la concorrenza... ;-)"*. In this case the user is delighted by his phone. First of all the full dataset has been labeled by three experts splitting the posts in positive, negative and neutral. At this point a training test of 100 positive posts, a training set of 100 negative posts and a training set of 100 neutral posts have been obtained selecting casually posts from positive, negative and neutral datasets. In this way a test set of 459 positive posts, a training set of 384 negative posts and a training set of 857 neutral posts have been obtained. In order to test the approach two experimental campaigns has been conducted. The first aimed to classify the documents according their sentiment. By the use of the three training set the mood ontologies has been built (figure 4).

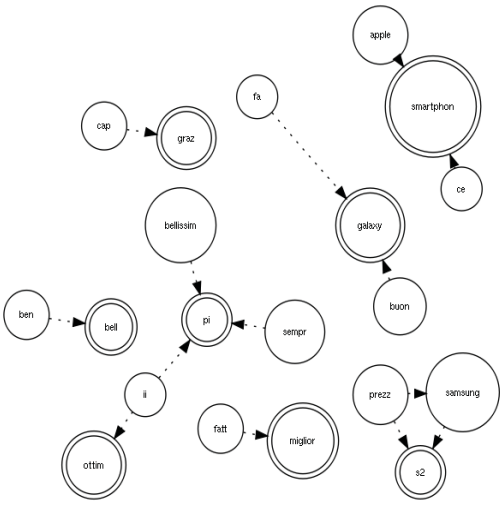


Figure 4 An example of positive sentiment ontology

We used them in order to label the various documents. In particular the full test set (positive test set, negative test set and neutral test set) has been labeled by the use of the various ontological filters. Introducing the concept of Precision defined as:

Precision = True Positive / (True Positive + False Positive)

And Recall defined as

Recall = True Positive / (True Positive + False Negative)

the obtained results are depicted in table 2

	Positive	Negative	Neutral	Total
Positive Test Set	397	19	43	459
Negative Test Set	23	332	29	384
Neutral Test Set	130	121	606	857
Precision	0,77			
Recall	0,81			

Table 2 Obtained Results for the case of document sentiment classification

The second experimental campaign has aimed to evaluate the skill of the system in the sentiment classification of a set of documents. This feature of the system can be very interesting for the evaluation of a stream of posts as for example Twitter. So we created in a random way 8 datasets of 300 posts organized as described in table 3.

	Positive Posts	Negative Posts	Neutral Posts	Total
DataSet_1	270	20	10	300
DataSet_2	225	50	25	300
DataSet_3	180	80	40	300
DataSet_4	100	100	100	300
DataSet_5	50	50	200	300
DataSet_6	80	180	40	300
DataSet_7	50	225	25	300
DataSet_8	20	270	10	300

Table 3 Configuration of the various datasets

The various datasets has been evaluated by the sentiment classifier based on ontological filtering with the following results (table 3).

	Sentiment Trend	SD
DataSet_1	Positive	0,83
DataSet_2	Positive	0,65
DataSet_3	Positive	0,57
DataSet_4	Positive	0,52
DataSet_5	Negative	0,54
DataSet_6	Negative	0,61
DataSet_7	Negative	0,69
DataSet_8	Negative	0,78

Table 4 Obtained Results for the datasets sentiment classification

Also in this case the system offers good results

6. Conclusions

This paper proposes the combination of the LDA approach and ontological filtering for the enhancement of the sentiment classification. Thanks to the LDA approach an ontology representing a set of document with a well-defined sentiment can be obtained. These ontologies can be adopted as ontological filtering and so detect the sentiment in a document or in a set of

documents. The first results on real datasets are quite interesting. Further development of this approach is still ongoing and a more detailed experimental campaign on standard datasets will be arranged.

References

[1] Bing Liu, "Sentiment Analysis and Subjectivity", Handbook of Natural Language Processing, Ed. Indurkhya and Damerau, 2010

[2] Bo Pang, Lee L., "Opinion mining and sentiment analysis", Found. Trend Inf. Retr., 2(1-2):1-135

[3] Bo Pang, Lee L., Vaithyanathan, "Thumbs up? Sentiment Classification Using Machine Learning Techniques", Proc. ACL-02 Conf. Empirical Methods in Natural Language Processing (EMNLP), 2002

[4] Prabowo R., Thelwall M., "Sentiment Analysis: a Combined Approach", Journal of Informetrics, Volume 3, Issue 2, April 2009, Pages 143–157

[5] Colbaugh, R., Glass K., "Estimating Sentiment Orientation in Social Media for Intelligence Monitoring and Analysis", Intelligence and Security Informatics (ISI), 2010 IEEE International Conference on, pp. 135-137, 2010

[6] Esuli A., Sebastiani F., "SentiWordNet: a Publicly Available Lexical Resource for Opinion Mining", Proc. Fifth Int'l Conf. Language Resources and Evaluation, pp. 417-422

[7] Neviarouskaya A., Prendinger H., Ishizuka M., "SentiFul: a Lexicon for Sentiment Analysis", IEEE Transactions on Affective Computing, 2-1, 2011

[8] Xiaohui Yu, Yang Liu, Xiangji Huang, Aijun An, "Mining Online Reviews for Predicting Sales Performance: a Case Study in the Movie Domain", IEEE Transactions on Knowledge and Data Engineering, 24-4, 2012

[9] Hofmann T., "Probabilistic Latent Semantic Analysis", Proc. Uncertainty in Artificial Intelligence (IJAI), 1999

[10] Blei D. M., Ng, A. Y., Jordan M., "Latent Dirichlet allocation", Journal of Machine Learning Research 3 (4–5): pp. 993–1022, 2003

[11] F. Sebastiani, "Machine learning in automated text categorization," ACM Comput. Surv., vol. 34, pp. 1–47, March 2002.

[12] T.L.Griffiths,M.Steyvers,andJ.B.Tenenbaum,"Topics in semantic representation," Psychological Review, vol. 114, no. 2, pp. 211–244, 2007.

[13] S. Noam and T. Naftali, "The power of word clusters for text classification," in In 23rd European Colloquium on Information Retrieval Research, 2001.

[14] G. Cao, J.-Y. Nie, J. Gao, and S. Robertson, "Selecting good expansion terms for pseudo-relevance feedback," in Proceedings of the 31st annual international ACM SIGIR. 2008, pp. 243–250. New York, NY,

[15] C.-J. Lee, Y.-C. Lin, R.-C. Chen, and P.-J. Cheng, "Selecting effective terms for query formulation," in Information Retrieval Technology, ser. Lecture Notes in Computer Science, Springer Berlin / Heidelberg, 2009, vol. 5839, pp. 168–180.

[16] P. R. Christopher D. Manning and H. Schtze, Introduction to Information Retrieval. Cambridge University, 2008.

[17] T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning, Springer, 2009

[18] http://wordnet.princeton.edu/

[19]http://www.ilc.cnr.it/viewpage.php/sez=ricerca/id=834/vers=ita

Efficient Computation of Object Boundary Intersection and Error Tolerance in VRCC-3D+

Nathan Eloe
nwe5g8@mst.edu

Jennifer Leopold
leopoldj@mst.edu

Chaman Sabharwal
chaman@mst.edu

Zhaozheng Yin
yinz@mst.edu

Missouri University of Science and Technology

Abstract

Computational geometry is a field that is relevant to computer graphics rendering, computational physical simulation, and countless other problem domains involving the use of image data. Efficiently determining the intersection of the boundaries, interiors, and exteriors of two objects can mean the difference between a realistic and relevant simulation, and a slow program that produces results that do not keep pace with user manipulation of the object. However, the speed of these calculations is not the only area of concern. Taking into consideration the finite unit of resolution in a computer display (the pixel) and error in the floating-point representation of numbers, it may be the case that the perceived correctness of these computations does not necessarily correspond to the accuracy with which the calculations are carried out. In this paper, we examine two of the most well-known methods of determining such intersections, as well as various programming language libraries available to perform these calculations. These existing approaches are considered with respect to limitations in human perception, display resolution, and floating point error. We also propose a new method which lends itself to exploiting the inherently parallel nature of these calculations.

1. Introduction

VRCC-3D+ [11, 12, 6, 5, 7] is a mathematical model that describes relations between three dimensional objects in space in terms of the connectivity and obscuration between each pair of objects in a scene. The system facilitates knowledge discovery by determining possible intermediate configurations of the objects from one state to another. Implementing the VRCC-3D+ mathematical model resulted in computational inefficiency (unacceptable delays between initial model loading and the calculation of the spatial relations between all pairs of objects). Our investigation of a more efficient method to perform these computations focused on determination of the intersections between the interiors, exteriors, and boundaries of the objects.

2. VRCC-3D+Intersections

The VRCC-3D+ system distinguishes eight types of 3D connectivity relations using eight intersections that involve the boundary, interior, and/or exterior of one object with those of another (see Table 1).

Connectivity in 3D is only one part of each (composite) VRCC-3D+ relation. The other part of each VRCC-3D+ relation characterizes the obscuration between the two objects of interest, which can be none, partial, or complete (as well as converse relations for each). Obscuration in VRCC-3D+ is determined by the overlap of the projections into the image plane and the object depth. The techniques (Triangle Triangle intersection, AABB Bounding Boxes) and considerations (error tolerance and speed) discussed in this paper will apply to the calculations of both the connectivity and obscuration parts of the VRCC-3D+ relations.

Definition of the spatial relations using Table 1 results in an over-determined system [11]. A symmetry argument also applies. The intersection of the interior of object A and the boundary of object B is the same as the intersection of the boundary of object B and the interior of object A: IntBnd(A,B)=BndInt(B,A). Thus, an efficient algorithm for IntBnd will provide an efficient algorithm for BndInt, and vice versa. This also holds for the intersections of boundaries and exteriors.

These four intersection predicates are insufficient to uniquely distinguish all relationships. For example, DC and EC are indistinguishable using these four predicates, and require either the BndBnd or the IntInt predicate. Computing the intersection of the boundaries allows us to determine the 3D relationships of two objects while avoiding unnecessary computation to determine the interior or exterior. The remaining predicates that

Table 1. Definition of 3D Spatial Relationships Using Intersection Predicates. Φ denotes empty intersection set, $\neg\Phi$ denotes non-empty intersection set.

	IntInt	IntBnd	IntExt	BndInt	BndBnd	BndExt	ExtInt	ExtBnd
DC	Φ	Φ	$\neg\Phi$	Φ	Φ	$\neg\Phi$	$\neg\Phi$	$\neg\Phi$
EC	Φ	Φ	$\neg\Phi$	Φ	$\neg\Phi$	$\neg\Phi$	$\neg\Phi$	$\neg\Phi$
EQ	$\neg\Phi$	Φ	Φ	Φ	$\neg\Phi$	Φ	Φ	Φ
NTPP	$\neg\Phi$	Φ	Φ	$\neg\Phi$	Φ	Φ	$\neg\Phi$	$\neg\Phi$
NTPPc	$\neg\Phi$	$\neg\Phi$	$\neg\Phi$	Φ	Φ	$\neg\Phi$	Φ	Φ
PO	$\neg\Phi$	$\neg\Phi$	$\neg\Phi$	$\neg\Phi$	$\neg\Phi$	$\neg\Phi$	$\neg\Phi$	$\neg\Phi$
TPP	$\neg\Phi$	Φ	Φ	$\neg\Phi$	$\neg\Phi$	Φ	$\neg\Phi$	$\neg\Phi$
TPPc	$\neg\Phi$	$\neg\Phi$	$\neg\Phi$	Φ	$\neg\Phi$	$\neg\Phi$	Φ	Φ

do not involve the boundaries are unnecessary and inefficient for our implementation.

3. Computation of Boundary Intersection

A naive method of computing the boundary intersection predicate is to use an algorithm such as that presented in [9] to determine pairwise intersection between the triangular faces of objects A and B: an $O(f_a f_b)$ algorithm with f_a and f_b as the number of faces in objects A and B. A more sophisticated method, AABB trees, uses nested axis-aligned bounding boxes to decrease the number of faces that require the intersection calculation. The use of trees presents an opportunity to reduce the complexity by reducing the number of calculations. CGAL [2] is a computational geometry toolkit that uses AABB trees to implement intersection and distance algorithms. Table 2 shows the runtimes of the triangle intersection and the CGAL methods on two spheres (see Figure 1), each of which has approximately 2000 faces. All timing was done on an AMD Bulldozer processor running at 3.1Ghz. Results were obtained using a C module and the CGAL bindings for Python and averaged 100 runs.

Table 2. Intersection Test Runtimes

Implementation	Average Runtime (s)
C Triangle Intersection	12.0349209094
CGAL AABB	0.186076021194

4. Human Perception and Floating Point Error

There is no ambiguity in Figure 1 regarding the intersection of boundaries. However, in a more complicated case such as that shown in Figure 2, more rigorous calculations are required. Upon cursory examination of the rendering of the airplane, it appears that all of the wings are attached to the fuselage. However, according to the

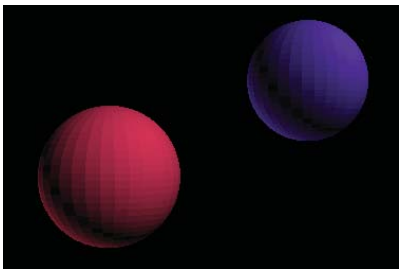


Figure 1. Two disconnected objects (DC).

computational intersection of boundaries, the wings and the fuselage do not intersect: they cannot be considered externally connected.

While the calculations are being carried out accurately, they do not reflect the cognitive perception of the image. This could be attributed to several factors such as errors in the representation of floating point numbers in both the viewer display and model generation software, or the minimum distance that can be represented in rendering to a monitor with finite resolution.

To remedy this, we must allow some tolerance: the shortest distance between the two planes must be less than some small value, ϵ. Implementing this in the direct triangle intersection test is trivial. However, using the CGAL AABB implementation requires a significant change in the algorithm itself: instead of determining whether the triangles intersect, we need to find the minimum distance between the two closest faces. One way to do this is to determine all of the segments that make up the faces of the object and find the closest distance to a face in the other object.

Table 3 shows the timings for these modified epsilon tolerance tests. Both algorithms also were used to determine the intersection of a high resolution depiction of a nut and bolt (Figure 3).

5. ϵ Determination

Once a tolerance test was developed for the intersection, it was necessary to determine an appropriate value

Table 3. Intersection Runtimes using ϵ Tolerance

	Airplane	Spheres	Nut and Bolt
Obj1 Name / Face Count	Wings_4 / 18	A / 1984	Helix01 / 43872
Obj2 Name / Face Count	Prop_4 / 144	B / 1984	Helix02 / 27432
C Triangle Intersection Runtime(s)	0.015720	12.017024	388.967400
CGAL AABB Runtime (s)	0.0110691	0.23603410	86.336132

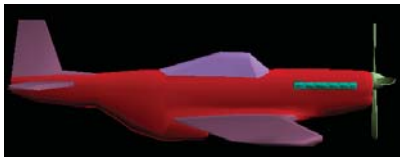


Figure 2. Model of an airplane.

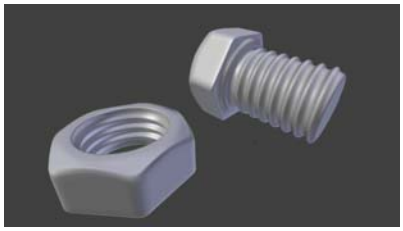


Figure 3. Nut and Bolt model.

of epsilon. Trial and error yielded an ϵ of approximately 10^{-4}. Appropriate values of epsilon will change based on the precision with which the values are stored in the model and the finite nature of the pixels that comprise the image. Of these two factors, we can both calculate and control the effect of the latter at the viewer level. Knowing the distance in world units from the virtual camera to the viewing plane (f in world units), the horizontal width in pixels of the scene (h in world units, w in pixel count), and the field of view angle (θ) allows the calculation of the physical size of the pixel at the image plane (see Figure 4). We can determine this value as follows. Let "px" be pixel units and "units" be world units:

$$h = f * tan\left(\frac{\theta}{2}\right)$$

We calculate the size in world units per pixel as

$$\frac{h}{\frac{w}{2}}\frac{units}{px} = \frac{2h}{w}\frac{units}{px}$$

The size of one pixel on the image plane is then

$$\epsilon_f = \frac{2h}{w}\frac{units}{px} * 1px = \frac{2h}{w}units = \frac{2f}{w}tan\left(\frac{\theta}{2}\right)units$$

The default visualizer parameters yield a value of approximately $1.4 * 10^{-3}$, which is a greater tolerance than the $1 * 10^{-4}$ that was found through manual experimentation. This value of ϵ gives the perceived results for

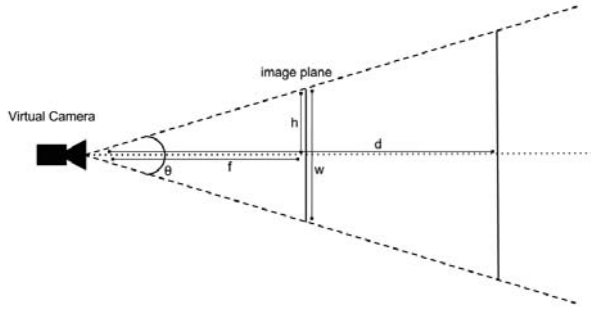


Figure 4. Camera and image plane.

Figure 2, and is only dependent upon the position of the objects and the resolution of the viewer window.

The minimum visible distance is directly proportional to the object distance from the camera. This calculation can be broken into two cases: the two objects have equal depth from the image plane, and the two objects have differing depths. The first case is trivial: given two objects of depth d from the camera, we use similar triangles (as depicted in Figure 4) to obtain the smallest distinguishable distance, ϵ_w, at that depth:

$$\frac{\epsilon_f}{f} = \frac{\epsilon_d}{d}$$

Since

$$\begin{aligned}\frac{\epsilon_f}{f} &= \frac{\frac{2f}{w}tan(\frac{\theta}{2})}{f} \\ &= \frac{2}{w}tan(\frac{\theta}{2})\end{aligned}$$

we can calculate the epsilon value at a depth d from the camera as

$$\epsilon_d = \frac{2d}{w}tan(\frac{\theta}{2})$$

This result can be reinterpreted to generalize it to all points regardless of distance to the camera. The above states that due to the finite nature of the pixel, we can define a sphere around any point with radius $\frac{\epsilon_w}{2}$ in which we cannot determine the actual position of the point (see Figure 5). This concept of a *probability cloud* is frequently used in Quantum Physics to approximately describe the locations of electrons in given states. We say that two points $P1$ and $P2$ are indistinguishable if their probability clouds overlap; that is, if $|P1 - P2|$ is less than or equal to the sum of the radii of their

probability clouds. If ϵ_1 and ϵ_2 are the respective epsilon values computed for these points $P1$ and $P2$, the resulting tolerance is:

$$\epsilon = \frac{\epsilon_1}{2} + \frac{\epsilon_2}{2} = \frac{\epsilon_1 + \epsilon_2}{2}$$

In Figure 5, the points $C1$ and $C2$ are considered to be within the error tolerance because the probability clouds overlap. The same holds for points $C2$ and $C3$. $C1$ and $C3$ are not within the error tolerance even though $C3$ is at the same depth as $C2$.

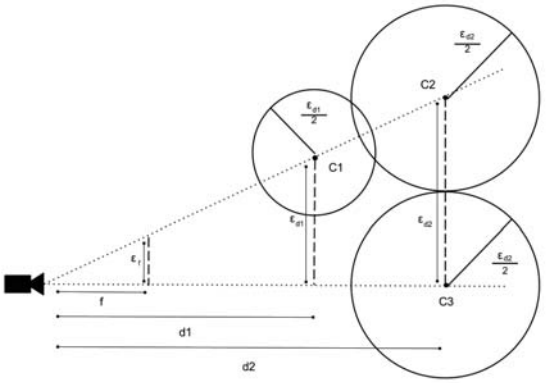


Figure 5. Using a probability cloud to dynamically determine ϵ tolerance.

6. Future work

This problem space is inherently parallel: the intersection of any pair of faces is independent of the intersection of any other pair of faces. The large number of similar calculations on a large, but static, set of data should be enough to overcome the primary downfall of using general purpose GPU computing technologies (e.g., OpenCL [3], CUDA [4] by NVIDIA, or Stream [1]) by AMD: the time it takes to transfer the data to the graphics card [8]. For the intersection of triangles method to be as fast as AABB trees, we require a speedup of 50X (e.g. $\frac{12.02}{0.236}$ for the spheres), which is well within the plausible speedups reported in [10].

A heavily parallelized AABB algorithm may also work, but would not necessarily be as cost efficient due to the recursive nature of trees. Exploration of an OpenCL implementation of these algorithms may even allow these computations to be done efficiently on mobile devices with sufficiently new integrated graphics.

7. Summary

Being able to efficiently calculate the intersections between the boundaries, interiors, and exteriors of 3D objects introduces new ways to program simulations, collisions, and model transformations over time. Exploiting the ability of VRCC-3D+ to identify impossible states and eliminating some calculations would lead to more efficient collision detection in game and simulation engines, and modeling software.

In this paper we have explored two methods of calculating these intersections and discussed methods of dynamically determining the error tolerance. These strategies will help to make spatial reasoning applications such as VRCC-3D+ more practical for knowledge validation and discovery in real-time.

References

[1] Ati stream technology. http://www.amd.com/stream.

[2] CGAL, Computational Geometry Algorithms Library. http://www.cgal.org.

[3] Chronos Group: OpenCL. http://www.khronos.org/opencl/.

[4] Nvidia technologies: Cuda. http://developer.nvidia.com/category/zone/cuda-zone.

[5] J. Albath, J. Leopold, and C. Sabharwal. Visualization of Spatio-Temporal Reasoning Over 3D Images. In *The 2010 International Workshop on Visual Languages and Computing (in conjunction with the 16th International Conference on Distributed Multimedia Systems (DMS'10))*, pages 277–282. KSI, Oct. 2010.

[6] J. Albath, J. Leopold, C. Sabharwal, and A. Maglia. RCC-3D: Qualitative Spatial Reasoning in 3D. In *The 23rd International Conference on Computer Applications in Industry and Engineering (CAINE 2010)*, pages 74–79. CAINE, Nov. 2010.

[7] J. Albath, J. Leopold, C. Sabharwal, and K. Perry. Efficient Reasoning with RCC-3D. In *The 4th International Conference on Knowledge Science, Engineering, and Management (KSEM 2010)*, pages 470–481. KSEM, Sept. 2010.

[8] D. B. Kirk and W.-m. W. Hwu. *Programming Massively Parallel Processors: A Hands-on Approach*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1st edition, 2010.

[9] T. Möller. A Fast Triangle-Triangle Intersection Test. *journal of graphics, gpu, and game tools*, 2(2):25–30, 1997.

[10] A. J. R. Ruiz and L. M. Ortega. Geometric algorithms on cuda. In *GRAPP*, pages 107–112, 2008.

[11] C. Sabharwal and J. Leopold. Reducing 9-Intersection to 4-Intersection for Identifying Relations in Region Connection Calculus. In *The 24th International Conference on Computer Applications in Industry and Engineering (CAINE 2011)*, pages 118–123. CAINE, Nov. 2011.

[12] C. Sabharwal, J. Leopold, and N. Eloe. A More Expressive 3D Region Connection Calculus. In *The 2011 International Workshop on Visual Languages and Computing (in conjunction with the 17th International Conference on Distributed Multimedia Systems (DMS'11))*, pages 307–311. KSI, Aug. 2011.

Determining Minimal Transitions Between VRCC-3D+ Relations

Adaleigh Martin, Ian Kottman, Robyn Littleton, Jennifer Leopold, Chaman Sabharwal, Nathan Eloe
Department of Computer Science - Missouri University of Science and Technology - Rolla, MO, USA
{admkbd, itk526, jllbd4, leopoldj, chaman, nwe5g8}@mst.edu

ABSTRACT

3D models are used in many problem domains. Typically the spatial relations between objects are approximated as necessary for human recognition of an image. For some applications, this may be insufficient; a more precise analysis may be required, and a temporal sequence of configurations of the objects may need to be considered. VRCC-3D+ is the implementation of a Region Connection Calculus (RCC) that qualitatively determines the spatial relations between 3D objects, both in terms of connectivity and obscuration. Herein we discuss how this system can be used to validate hypotheses of spatial transformation over time, and generate possible intermediate configurations. Such analyses have been performed manually in research fields such as organism evolution and development. Automation of this process has the potential to improve both the efficiency and precision of such investigations.

1. INTRODUCTION

Increasingly, 3D models are being used in fields such as biology and engineering to visualize and analyse complex images. In particular, given a configuration of 3D objects at abstract time t_1, it may be of interest to know what configurations of those objects are possible at another time t_2. That is, we may seek to determine if it is possible for the configuration of objects at one time to have transformed into the configuration at another time, or, even more interestingly, exactly what intermediate spatial transformations could have occurred between time t_1 and t_2.

The determination of answers to such questions can be computationally prohibitive. Additionally, these problems require a mathematical model sufficient to characterize the pertinent spatial relations between the 3D objects.

VRCC-3D+ [11] is the implementation of a Region Connection Calculus (RCC) that quantitatively determines the spatial relations between 3D objects, both in terms of connectivity and obscuration. Herein we discuss how this system can be used to validate (or refute) hypotheses of spatial transformation over time, as well as generate possible intermediate configurations of the objects, effectively facilitating knowledge discovery.

For brevity, we have limited our discussion to consideration of a basic case: solid spheres that can only be related by disconnection or external connection. However, these results establish the groundwork for generalizing the solution to the full VRCC-3D+ model.

2. RELATED WORK

There is a considerable amount of research in the area of qualitative spatial reasoning (QSR) [2, 3], particularly using RCC models. Unfortunately, the vast majority of this research has been to propose new mathematical models of varying degrees of expressivity, without consideration of the computational complexities of implementation [9]. Among the most well-known of these models are RCC-5 [6], RCC-8 [6], RCC-23 [3], LOS-14 [4], ROC-20 [7], and RCC-3D [1].

Every relation in VRCC-3D+ is expressed as a composite relation that characterizes connectivity and obscuration. For a more in-depth discussion of the system, including how it compares to the aforementioned other models, see [11].

VRCC-3D+ uses a hierarchical structure of predicates (as shown in Figure 1) to distinguish the 3D connectivity part of the relation. A pair of objects, A and B, can be either disconnected (DC(A,B)) or connected (C(A, B)),. If connected they may be Externally Connected (EC(A,B)), Partially Overlapping (PO(A,B)), Equivalent (EQ(A,B)), or one of the objects may be a proper part (i.e., completely inside) of the other. Given two objects, A and B, we define the converse relation Proper Part Converse (PPc(A,B)) such that if B is a proper part of A (PP(B, A)), then PPc(A, B) is true. If PP(A, B) holds, the objects can either be tangential or non-tangential (TPP, NTPP). If PPc(A, B) holds, then we can use the TPPc or NTPPc relations.

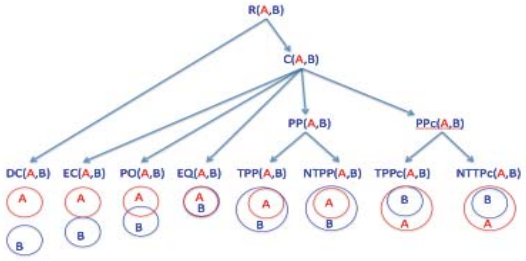


Figure 1. VRCC-3D+ hierarchical structure for connectivity relations

The obscuration part of each relation is determined by the overlap of the projections into the image plane as well as the object depth, as illustrated by Table 1. In this table, $\neg\Phi$ denotes a non-empty intersection of the 2D projection into the image plane, and Φ denotes an empty intersection. It should be noted that only certain combinations of connectivity and obscuration are possible; see [11] for a more detailed discussion.

	IntInt	IntExt	ExtInt	A InFront B
pObs	¬Φ	Φ	¬Φ	YES
pObs_c	¬Φ	{Φ,¬Φ}	Φ	NO
pObs_e	¬Φ	¬Φ	¬Φ	EQUAL
cObs	¬Φ	¬Φ	{Φ,¬Φ}	YES
cObs_c	¬Φ	Φ	¬Φ	NO
eObs_e	¬Φ	Φ	Φ	EQUAL
eObs_c	¬Φ	Φ	Φ	NO
eObs	¬Φ	Φ	Φ	YES
nObs_e	Φ	¬Φ	¬Φ	EQUAL
nObs_c	Φ	¬Φ	¬Φ	NO
nObs	Φ	¬Φ	¬Φ	YES

Table 1. Definition of 2D Obscurations Based on Interior and Exterior Intersections, and Depth

3. APPLICATION OF VRCC-3D+

We consider a *state* to be the set of V RCC-3D+ relations between each pair of objects in a scene at a particular abstract time. A *state transition* occurs when the objects morph from one state to another. A *step* occurs when two objects have gone from the current state to an *adjoining state*. An adjoining state can be reached from the initial state without passing through an intermediate state.

In order to prove that an object can go from one state to another in one step, we need a means of describing the system both visually and mathematically. This can be accomplished with a combinatorial approach. Figure 2 is a combinatorial map that illustrates how two solid spheres can be configured for the VRCC-3D+ DC and EC relations.

In Figure 2, C1 is the centroid of *Object 1*, the dark gray circle at the center of the map, R1 is the radius of *Object 1*, and R2 is the radius of *Object 2*, which is not explicitly shown, but is defined as another sphere with radius R2 located in one of the regions of the map. The different relation "regions" are determined by the the position of C2, the centroid of *Object 2* (not shown). A is the angle between two rays cast from the observer, tangential to *Object 1*.

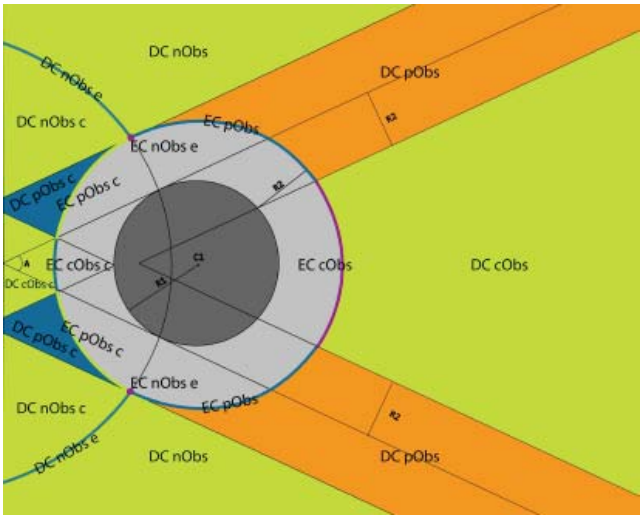


Figure 2. Combinatorial map of DC and EC relations

Some regions lie on an arc. If the centroid of *Object 2* does not lie on the EC curve, then the objects are perceived as DC. The line that divides nObs (no obscuration) and pObs (partial obscuration) was created by translating angle A behind the viewpoint. Similarly, the dividing line between pObs and cObs (complete obscuration) is created by translating the angle in front of the viewer.

The EC curve was created by generating a circle about the point C1 that has radius R1 + R2, which is the closest that *Object 2* can be to *Object 1* without overlapping. The nObs_e (no obscuration) curve was created by generating the circle centered at the intersection of the lines that separate pObs from nObs, and passing through the points where those dividing lines tangentially intersect the EC curve. These rules form the basis of the mapping, and allow combinatorial methods to be applied to the map to create a graph that can be used by an automated spatial reasoner.

Since this is a map, by the Four Color Theorem it is 4-colorable as shown in Figure 2 [8]. This allows each region used to be depicted using a perspective projection. Two examples can be seen in perspective projections of state transitions, shown in Figures 3 and 4.

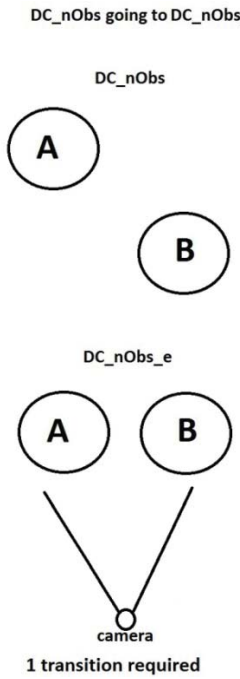


Figure 3. Perspective projection of transition from DC_nObs to DC_nObs_e

In Figure 3, the centroid of *Object A* (CA) does not lie on the EC curve of *Object B*. They are perceived by the camera to be DC. CA also lies outside of nObs. When the transition occurs, CA remains outside of the EC curve. The transition occurs when CA reaches the nObs_e arc. Referring to Figure 2, we can see that to go from the nObs region of DC to the nObs_e arc of DC requires only one spatial transition. Hence we can validate the hypothesis that the two configurations of *Object A* and *Object B* shown in Figure 3 are one transition apart.

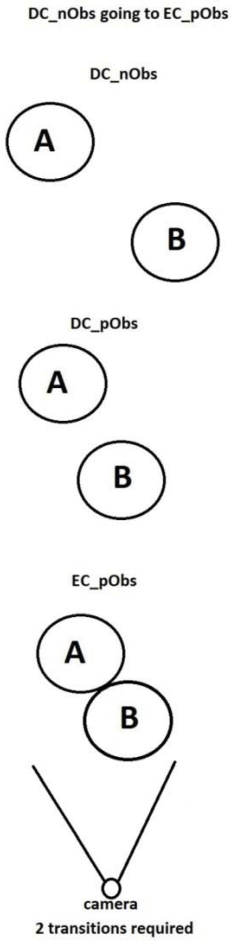


Figure 4. Perspective projection of transition from DC_nObs to EC_pObs

In Figure 4, we use a similar process to find the intermediate state(s) when we know the start and end states. Here the start state of *Object A* and *Object B* is DC_nObs (i.e., *Object A* is behind *Object B*). The end state is EC_pObs, where *Object A* is obscured by *Object B,* and they are edge connected. We assume that we are only interested in the shortest transition path that will take us from the start state to the end state. This can be determined by using the combinatorial map.

The centroid of *Object A* (CA) does not lie on the EC curve of *Object B*, and is outside of the nObs line. Getting to EC_pObs requires two transitions. CA must first cross the DC_nObs line into the DC_pObs region before it lies on the EC curve of *Object B* within that region. CA cannot reach to the EC curve from DC_nObs without crossing through some other intermediate transition. In this case, DC_pObs is the simplest choice as the (single) intermediate state.

An undirected graph can be constructed, using the adjacencies of each possible spatial region. Since EC_nobs_e is adjacent to many relations, it may be difficult to accurately represent it in a computer model due to round-off error. Figure 5 shows the undirected graph generated from the map. This graph also is represented as an adjacency matrix in Table 2 (albeit with 'T' entries for each vertex to itself, despite the fact that there are no loop edges in the graph).

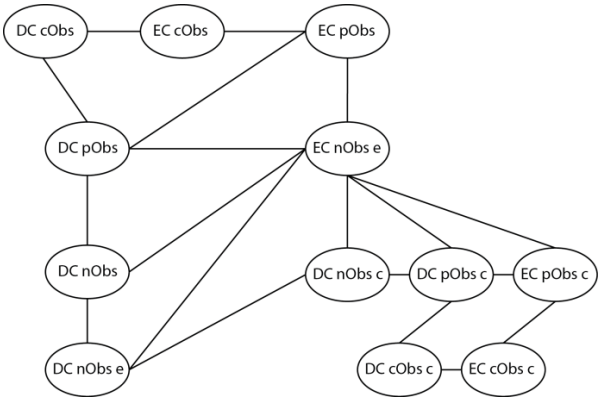


Figure 5. Undirected graph of DC and EC transitions

A list of all the simple paths from a given relation to a final relation can be created by recursively searching Table 2. This is accomplished using a recursive depth-first search through all possible one-step relation changes starting at

	DC_nObs	DC_nObs_c	DC_nObs_e	DC_pObs	DC_pObs_c	DC_cObs	DC_cObs_c	EC_nObs_e	EC_pObs	EC_pObs_c	EC_cObs	EC_cObs_c
DC_nObs	T	F	T	T	F	F	F	T	F	F	F	F
DC_nObs_c	F	T	T	F	T	F	F	T	F	F	F	F
DC_nObs_e	T	T	T	F	F	F	F	T	F	F	F	F
DC_pObs	T	F	F	T	F	T	F	F	T	F	F	F
DC_pObs_c	F	T	F	F	T	F	T	F	F	T	F	F
DC_cObs	F	F	F	T	F	T	F	F	F	F	T	F
DC_cObs_c	F	F	F	F	T	F	T	F	F	F	F	T
EC_nObs_e	T	T	T	F	F	F	F	T	T	T	F	F
EC_pObs	F	F	F	T	F	F	F	T	T	F	T	F
EC_pObs_c	F	F	F	F	T	F	F	T	F	T	F	T
EC_cObs	F	F	F	F	F	T	F	F	T	F	T	F
EC_cObs_c	F	F	F	F	F	F	T	F	F	T	F	T

Table 2. Table of Relations to Determine Possible One-Step Transitions

relation *R1* and ending at relation *R2*. The computational complexity for this approach is O (n^2), where n is the number of relations reachable from a starting state. The algorithm is shown in Figure 6.

```
LET Adj_Matrix = ADJACENCY_MATRIX
FUNCTION find_paths(R1,R2)
  recursive_search(R1,R2,Closed,Paths)
  SORT Paths FROM SHORTEST TO LONGEST
  OUTPUT Paths

FUNCTION recursive_search(R1,R2,Closed,Paths)
  IF R1 EQ R2:
    Closed = Closed + R2
    Paths = Paths + Closed
    RETURN
  ELSE:
    Closed = Closed + R1
    FOR R IN Adj_Matrix FROM R1:
      IF R Adj_Matrix[R][R1] <= 1 AND R NOT IN Closed
        recursive_search(R,R2,Closed,Paths)
```

Figure 6. Search algorithm for determining possible transitions

4. FUTURE WORK

Discussion in this paper is limited to solid spheres with static size and shape. It would need to be generalized to include the comprehensive VRCC-3D+ model, which has no restriction on the shape of the objects, and allows for objects to change size. Consequently, the scalability of the solution given here must be further investigated (i.e., to establish the validity of the argument for regions topologically equivalent to a sphere, and the correctness of results for regions not topologically equivalent to a sphere).

While the presented algorithm may not be practical for the full VRCC-3D+ model in terms of runtime efficiency, it will be a useful foundation for extending the solution.

5. SUMMARY

Currently 3D models are used in many problem domains. Typically the spatial relations between objects are approximated as necessary for human recognition of an image. This may be insufficient for some applications; a more precise analysis may be required, and a temporal sequence of configurations of the objects may need to be considered.

In this paper we have briefly discussed how one qualitative spatial reasoning system, VRCC-3D+, could be used to validate hypotheses of spatial transformation over time, and generate possible intermediate configurations. For years such analyses have been performed manually in research fields such as organism evolution and development. Automation of this process has the potential to improve both the efficiency and precision of such investigations.

6. REFERENCES

[1] Albath, Julia, Jennifer Leopold, Chaman L. Sabharwal and Kenneth Perry. "Efficient Reasoning with RCC-3D." *Knowledge Science, Engineering and Management: Lecture Notes in Computer Science* 6291, (September 2010): 470-481.

[2] Cohn, A.G. and S. M. Hazarika. "Qualitative Spatial Representation and Reasoning: An Overview." *Fundam. Inf.* 46, 1-2 (January 2001): 1-29.

[3] Cohn, A. G., Bennett, B., Gooday, J., and Gotts, N. M. "Qualitative spatial representation and reasoning with the region connection calculus." *Geoinformatica 1, 3* (Oct. 1997): 275-316.

[4] Galton, A. "Lines of sight". In *AISB Workshop on Spatial and Spatio-Temporal Reasoning*, (1994).

[5] Li, Sanjiang and Mingsheng Ying. *Region Connection Calculus: its models and composition table*. Beijing: Tsinghua University, 2010.

[6] Randell, D.A., Z. Cui, and A.G. Cohn. "A Spatial Logic Based on Regions and Connection." *KR* 92, (1992): 165–176.

[7] Randell, D., Witkowski, M., and Shanahan, M. "From images to bodies: modelling and exploiting spatial occlusion and motion parallax." In *Proceedings of the 17th international joint conference on Artificial intelligence - Volume 1 (San Francisco, CA, USA, 2001)*, (2001): 57-63.

[8] Roberts, F. S., and Tesman, B. A. Applied combinatorics (2. ed.). Prentice Hall, 2005.

[9] Sabharwal, C. and J. Leopold. "Reducing 9-Intersection to 4-Intersection for Identifying Relations in Region Connection Calculus." *The 24th International Conference on Computer Applications in Industry and Engineering (CAINE 2011)*, (Nov. 2011): 118–123.

[10] Simone Mantini, Maurizio Ripani. "Modern morphometry: new perspectives in physical anthropology." *New Biotechnology* 25, no. 5 (June 2009): 325-330.

[11] Sabharwal, C., Leopold, J., and Eloe, N. "A more expressive 3d region connection calculus". In *Proceedings of the 2011 International Workshop on Visual Languages and Computing (in conjunction with the 17th International Conference on Distributed Multimedia Systems (DMS'11))*, (2011), 307–311.

EEG-Based BCI Data Analysis on Visual-Perceptual Priming in the Context of a Museum of Fine Arts

Annalisa Banzi
Istituto di Comunicazione, Comportamento e Consumi "Giampaolo Fabris"
Università IULM
Milan, Italy
Annalisa.Banzi@gmail.com

Raffaella Folgieri
Dipartimento di Economia, Management e Metodi Quantitativi
Università degli Studi di Milano
Milan, Italy
Raffaella.Folgieri@unimi.it

***Abstract*— *Often students and lay public are not able to watch an art exhibit displayed in a museum of fine arts as scholars could wish. It has to be considered that watching is a learned skill that is neither innate nor spontaneous. Consequently, onlookers would benefit from a method that may enhance their visual skills. This consideration and the willing to give a contribution in artworks understanding guide our work. The goal of this paper is to assess whether a model based on visual-perceptual priming, a kind of implicit memory, may improve the methodology of looking. To support our model also with some kind of measures, we analysed the subjects' brain signals collected by an EEG-based Brain Computer Interface device during the verification phase of the performed experiments. In this work we also present some preliminary results obtained by this BCI-based promising approach.*

I. AIMS AND MOTIVATIONS

Museums are seeking to offer visitors a complete and participative experience, asking to contribute in the making of the museum experience. The main focus of this paper is on technology as a communication medium, but we wish to recall that technology also represents the tools we use to solve problems and develop the products that we need (e.g., many museums worldwide offer workshops and programs engaging children in learning technology, make tools, and express creativity).

Progress in technology and the multiple kinds of applications, devices and Real or Virtual offer museums the opportunity to involve visitors improving their experience. Thanks to technology, museums are also able to track visitors and understand what they are doing both in the museum and on-line. To receive a feedback and involve visitors in the process of experiencing Art, many museums have incorporated surveys or feedback mechanisms into their technology devices as well [1]. Another way adopted by museum educators in order to explore an exhibit with the lay public consists in the inquiry-based method. As an example, we wish to recall the GRAM (Grand Rapids Art Museum) experience. GRAM staff has provided a methodology based on an inquiry-based tour of the museum. The GRAM's[1] approach mainly stresses the importance of encouraging the development of viewers oral and written communication skills in relation to Visual Arts. To this aim, GRAM docents are trained to exploit specific questioning techniques in the course of the tour encouraging students' response.

Angela Lawler and Susan Wood [2] identify five steps that allow learning to look at art. At first, students observe artworks on their own and in complete silence. Subsequently, they describe the artwork objectively. After these two steps, within the third step, students analyse contents, such as colour, balance, space, line, value, technique, *etc*. The fourth step tries to interpret the works of art exploiting what students know and have seen. Finally, students make a critical judgment of the artwork.

As discussed, comparing and asking questions are currently the main methods used to analyse exhibits.

Shall we use other methodologies of looking in order to achieve the same goal? The considerations and the results presented in this paper starts by tackling this topic from a different point of view. The next section introduces the psychological phenomenon called *priming* that can be used to foster visual skills. Then, we describe the approach adopted to improve the methodology of looking with visual-perceptual repetition priming and the results obtained by experiments performed within a Museum of Fine Arts (Pinacoteca Ambrosiana, Milan – Italy). In the last section we present our conclusions and future works.

II. THE VISUAL-PERCEPTION PRIMING

At first, consider the following example: A person reads a list of words, including the word *cherry*. Subsequently the person is asked to complete a word starting with *ch*. The probability that he/she will answer *cherry* is increased because the word was previously primed. Therefore, if a stimulus is primed, later experiences of this stimulus will be processed more quickly and precisely by the brain.

[1] 2011. Learning to Look. Gathering Meaning Through Observation. Retrieved 11 April, 2011, from http://www.artmuseumgr.org/uploads/assets/LearningtoLookPacket.pdf.

Priming is a kind of implicit memory (a sort of memory unconsciously retrieved or observed). The peculiarity of this kind of memory consists in its permanence in a relatively stable level from age 3 to 80, while the performance of episodic memory based on explicit tasks initially improves with age and declines with advancing age.

For our purpose we are interested in the visual-perceptual priming. It is defined by enhanced processing of previously seen visual material, relative to the novel visual material. Wiggs and Martin [3] review the literature as to the main visual-perceptual priming experiments. They stress that perceptual priming is sensitive to changes in physical appearance only in some instances. In general, stimulus attribute alterations – such as colour, pattern, luminance, contrast, location, left-right reflection, and size – do not influence priming. At the same time, perceptual priming can be attenuated when stimuli are changed so as to affect the ability to identify stimulus form. Specifically, it is not affected by relatively small changes in orientation (e.g., rotations in depth up to 67°) but is eliminated by large changes in or ientation (e.g., rotations in depth > 80°). Furthermore, the phenomenon is diminished with changes in an object's exemplar (e.g., a different picture of the same-named object), and with changes in a word's typography from *study* to *test*. The above results suggest that physical attributes are essential to the representation of object form – such as line elements of the drawings, or written word form – and they influence visual-perceptual priming. The last consideration concerns «the degree of attention devoted to encoding» that, as stated in [3], «typically does not affect the magnitude of priming. Thus, when attention is divided during encoding, priming is no different than when attention is focus».

III. IMPROVING THE METHODOLOGY WITH VISUAL-PERCEPTUAL REPETITION PRIMING

The purpose of our research consist in developing a priming-based tools taking into account the most relevant experimental and physiological findings and to apply them to the museum environment.

The lay public and students are often unskilled visual onlookers because usually they lack a proper education. The ability to analyse the artwork formal qualities is intrinsic to complete understanding of the art-making process. Therefore people need to develop or improve visual skills. In order to achieve these aims we have chosen to analyse and apply priming to museum environment. This phenomenon possesses some interesting characteristics: perceptual priming effects are long-lasting in normal adults and amnesic patients, priming remains relatively stable from age 3 to 80 , the degree of attention devoted to encoding typically does not affect the magnitude of priming, and finally, but not less important, this phenomenon seems to be independent of cultural background.

Our aim consists in improving visitors' visual skills showing visual primed stimuli related to the artworks (colours, lines, shapes, and so forth). W e wish, so, t o analyse the effect of the primed stimuli in observing artworks.

The performed experiment consisted in three phases:

- Visitors of a Museum of Fine Arts have been submitted to primed stimuli (priming group). Others have been submitted to a neutral stimulus (neutral group).
- Participants visited the Museum of Fine Arts.
- At the end of the tour, people answered to several questions concerning the artworks.

In the last phase of the experiment, during the test performed by participants to verify the effects of priming stimuli, we also asked for participants to wear a BCI device. BCI devices are a simplification of EEG medical equipment. The difference consists in th eir flexibility, low cost and comfort for users. BCI devices are widely used in ICT for human-computer interaction [5, 6] and seems particularly suited to our aims, because in the literature they have been largely used also to measure human reaction to specific stimuli [7]. In this latter case the results in the literature demonstrate that BCI is comparable with the EEG medical equipment in measuring reactions to computer-based stimuli. We acquired subjects' EEG data while participants were asked to answer some questions about the experience. For its temporal resolution, EEG could give, in fact, important indications, confirming if and when, answering to the questions, participants present variation in EEG signals, due to stimuli recognition, frustration and/or change in attention levels.

A. *Experimental setup*

The final goal of our research is to model either technological devices and educational resources (wall and caption texts, booklet, *etc.*) based on priming process.

The experimental setup is based on the concept of visual-perceptual repetition priming. It has been modelled for on a museum tour (Pinacoteca Ambrosiana – Milan, Italy) where participants singularly watch prime stimuli on a screen under the supervision of researchers. On the whole, the experiment requires three statistically sampled groups of subjects: priming group, neutral group, and control group (visitors who did not receive stimuli). The neutral stimuli are unrelated to the artistic features of the works of art selected (e.g., pictures in black and white of sunglasses).

After a study of the documents related to the paintings and the direct observation of the artworks located in Pinacoteca Ambrosiana, we chose to work with colours (red, green, brown, white, and blue) selected for their:

- saturation, value, and hue;
- occurrence in the artworks;
- symbolic meaning;
- extent.

In this experiment, the portions of colours have been selected from 5 a rtworks positioned in Pinacoteca Ambrosiana.

B. Procedure

The priming group watched a session of 5 prime-stimuli (colours: red, green, blue, brown, and white) alternated with neutral-stimuli (objects in black and white not depicted in paintings: luggage, phone, baby's bottle, vacuum cleaner, and headband) for a short period of time (1 minute circa).

The participants chosen to receive neutral stimuli watched a trail composed by 10 stimuli completely unrelated to paintings (objects pictures in black and white such as sunglasses) for a short period of time (1 minute circa).

As told, a control group visited the museum without any previous visual stimulus. These visitors constitute the experiment baseline to which compare the results collected in the other (priming and neutral) groups. At the end of their tour, participants are asked to answer some questions in order to check if the prime-stimuli (colours) helped them remembering the artworks main features (*target*).

During the final test, we also registered subjects' EEG signal [8], using a Brain Computer Interface (BCI) devices [7] to collect EEG signals. We chose BCI to avoid influence in anxiety of participants, compared to a medical EEG equipment. BCI devices are a simplification of the medical EEG equipment and currently several kinds of low-cost, non-invasive BCI could be chosen for our research objective. We choose to collect EEG data using a Neurosky Mindwave™ BCI device, used in several research applications and yet used by authors in previous works [9, 10]. Before choosing this device, we compared it to other BCIs, such as, for example, Emotiv *Epoc*™. The comparison analysis showed that the *Mindwave*™ BCI results more comfortable for users, both for the easiness of positioning the device on the scalp, and because it uses a dry sensor instead of wet ones. The Neurosky MindwaveTM BCI device consists of a headset with an arm equipped with a single dry sensor acquiring brain signal from the forehead of the user at a sample rate of 512 Hz, transmitted via Bluetooth to a host computer. Moreover brain functions interesting our work are related to the premotor frontal cortex area, on which the Mindwave™ sensor is positioned. In fact, the signals from the frontal lobes are linked to higher states of consciousness.

BCIs collect several cerebral frequency rhythms. In our work, we concentrate on alpha, beta, theta and gamma band. We recall that activity in the alpha band (7 Hz – 14 Hz) is related to relaxed awareness, meditation, contemplation, etc. Beta band (14 Hz – 30 Hz) is associated with active thinking, active attention, solving concrete problems. Theta band (4 Hz – 7 Hz) is usually related to emotional stress, such as frustration & disappointment, while activity in gamma band (30 Hz – 80 Hz) is generally related to cognitive processing of multi-sensory signals.

C. Results

The goal of the present research is developing and training technological tools by means of priming. As previously mentioned, priming has interesting features, such as long-lasting effects, stability despite age, imperviousness to attention degree, and independence of cultural background, that can be exploited in a museum environment.

The results from the designed experiment have been analysed from two points of view: the outcomes given by the questions answered by participants at the end of the museum tour; the EEG signals collected by the BCI devices.

A first consideration is related to the observation of the participants' reaction during the question session. Participants demonstrated curiosity and interest in the experiment. During the test performed at the end of the experiment, some persons commented the questions, as they were remembering to have seen, in the artworks, the colour on which the question was based.

Encouraging results came also from the EEG signal analysis. Data collected by EEG signals have been analysed using MATLAB™ functions, to detect the presence of differences in brain activity during the final question session. To this aim, we calculated the average Power Spectral Distribution (PSD) [11] in alpha, beta, theta and gamma rhythms [8, 12] for all the groups of users.

The value of PSD gives information about how the power of a signal is distributed with frequency, and therefore the average values in each band indicate the overall behaviour of the brain activity eventually induced by the primed stimuli. For PSD we used the Welch's method [13] with Hamming window function (overlap 50%, segment length 64) [14].

We computed for each participant data the ratio between the average power in each band and the average power in the frequency interval between 0 Hz and 80 Hz.17, to compensate the different data ranges for each user, due to personal variability. We finally computed an average of these ratios in each band.

From the analysis of the obtained plot, it results that in subjects who received the primed stimuli we can see an increase in the attention level. This is, of course, particularly evident in Beta and Gamma bands, related to active thinking and attention. At the same time, in subjects who received the primed stimuli, Theta brainwaves decrease, indicating that participants did not feel frustration or disappointment, if compared to individuals who did not receive any stimuli and to subjects who received a neutral stimulus. In these two latter cases, in fact, Beta and Gamma bands decrease, compared to the first group, while we registered an increasing of Theta band.

For participants who received the primed stimuli, Gamma and Beta bands revealed an average band power significantly greater than individuals who did not receive a stimulus and to whom received a neutral one. The average band power for Alpha and Theta rhythms decrease for the first group, compared to the other two.

As an example, in figure 1 we show the EEG signal related to the attention level and to the Beta rhythm of a participant in the experiment who received a visual-prime stimulus.

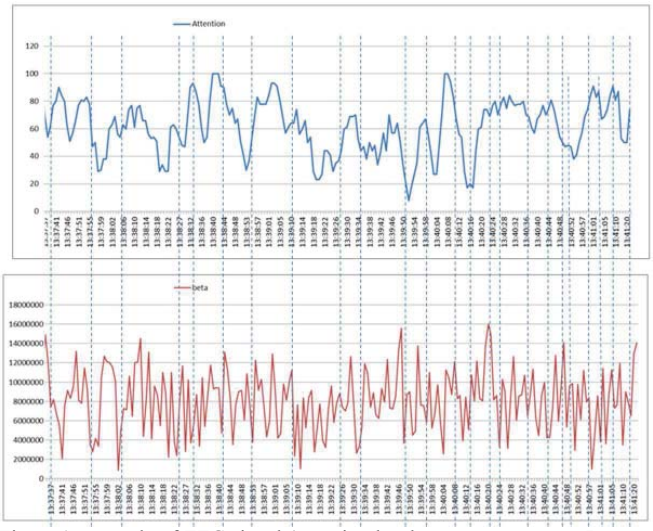


Figure 1: example of EEG signal (attention level – average power – Beta rhythm) of a participant who received a primed priming.

The described new approach appears promising for this kind of application, to measure the effects of visual stimuli. Of course, the application of EEG signal analysis to this kind of experiment represents also a challenge and we are aware that the presented experimental setups are not based on a specific task. In fact, the activity in the analysed bands is often considered and studied in well-defined task-based experiments. We are currently considering other experimental setups, based also on ERPs, in order to investigate the actual information given by the beta/alpha ratio as an index of attention in EEG based experiments on visual priming and to identify other measures, more suited for our research purpose.

IV. CONCLUSIONS

Developing and training visual skills using priming-based tools is the goal of the present research. As previously mentioned, priming has interesting features, such as long-lasting effects, stability despite age, imperviousness to attention degree, and independence of cultural background, that can be exploited in a museum environment. In this paper we presented encouraging results obtained submitting individuals to a museum tour where participants singularly watch prime-stimuli on a screen.

Participants in the experiment have been divided into three groups: a priming group (receiving the colour stimuli related to the paintings located in Pinacoteca Ambrosiana), a neutral group (receiving stimuli unrelated to paintings showed at the Museum) and a control group (in this case the participants did not receive any stimulus).

While participants were answering to the final questions, we registered their EEG signals using a non-invasive BCI device. The presented preliminary results shows that, compared to subjects who did not receive specific stimuli, in participants who received the primed stimuli, we registered an increasing of the attention level corresponding to questions related to the engagement of memory due to the visual stimuli. Also Beta and Gamma bands, related to active thinking and attention, presented a regular track on the same questions. Theta brainwaves did not show frustration symptoms and, correspondently, Alpha values, also related to meditation and contemplation, confirmed the relaxed attention state of subjects.

On one hand this new approach promises future improvement in exploring priming mechanisms, while, on the other hand, results represent just a preliminary step in improving the EEG use for our research aims.

Future works will have mainly the objective to individuate measures more significant for our aims. We also will have to perform more experiments to validate this innovative approach, representing, at the same time, a great opportunity and a challenge.

REFERENCES

[1] W. Witschey, H. J. Parry, E. Maurakis, D. Hagan,"Museums in Transition: Emerging Technologies as Tools for Free-Choice Learning", downloaded 19 March 2012, http://www.gyroscopeinc.com/articles/MuseumsInTransition.pdf

[2] A. Lawler, S. Wood, , "Learning to Look. Interdisciplinary Tour", K-12, Educator's Guide, 2011, retrieved 12 April, 2011, from http://museum.research.missouri.edu/pdfs/ LearningToLookGuide.pdf.

[3] C. L. Wiggs, A. Martin, "Properties and Mechanisms of Perceptual Priming", Current Opinion in Neurobiology, 1998, 8(2): 227-233.

[4] L. Nuel, "Art & visual literacy", in Journal of visual verbal languaging, fall, 1984, 77-79.

[5] R. Folgieri, M. Zichella, "Conscious and unconscious music from the brain: design and development of a tool translating brainwaves into music using a BCI device", in proceedings of AHFE, 2012, San Francisco, California, USA.

[6] D. Friedman et al., "Navigating Virtual Reality by Thought: What Is It like?", Presence, 2007, 16:1, 100-110.

[7] B. Z. Allison, E. W. Wolpaw, J. R. Wolpaw, "Brain-computer interface systems: progress and prospects", Expert Rev Med Devices, 2007, 4(4):463-74.

[8] E. Niedermeyer, F. Silva, "Electroencephalography: Basic Principles, Clinical Applications, and Re-lated Fields", Lippincott Williams & Wilkins, 2004.

[9] E. Calore, R.Folgieri, D.Gadia, D. Marini,"Analysis of brain activity and response during monoscopic and stereoscopic visualization", in proceedings of IS&T/SPIE's 24th Symposium on Electronic Imaging: Science and Technology, San Francisco, California, 2012.

[10] R. Folgieri,"VR for cultural heritage valorisation : a communication problem", Virtual Reality 21501, 2011.

[11] M. Priestley, "Spectral Analysis and Time Series", Academic Press, 1991.

[12] M. Bear, B. Connors and M. Paradiso, "Neuroscience: Exploring the Brain", 3rd edition, 2007, Lippincott Williams & Wilkins.

[13] P. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms", IEEE Transactions on Audio and Electroacoustic, 1967, 15(2): 70 –73.

[14] A. Oppenheim, R. Schafer, "Discrete-Time Signal Processing", Prentice-Hall, 1989.

Use of large multi-touch screens for informal learning

Rosa Lanzilotti, Maria Francesca Costabile
Dipartimento di Informatica
Università degli Studi di Bari "Aldo Moro", Bari, Italy
{lanzilotti, costabile}@di.uniba.it

Abstract — **The advent on the market of multi-touch devices has prompted studies that investigate the use of such devices in learning domain, demonstrating that games are an effective way of exploiting these new technologies that allow pupils to achieve informal learning and foster collaboration during their educational activities. We defined an educational format that combines traditional learning performed at school with educational multimedia games implemented on a multi-touch screen, set up vertically. A field study showed the effectiveness of this educational format, supporting the use of applications on the multi-touch display to help pupils consolidating the acquired knowledge. This paper reports a complementary study that analyses pupils' reactions to the use of a traditional desktop as compared with the use of the multi-touch screen. Results provided further evidence that pupils enjoy interacting with the multi-touch screen, because it allows them to collaborate and to use their hands, directly moving objects about on the screen.**

Multi-touch display; collaborative multimedia system; educational game in cultural heritage

I. INTRODUCTION

Researchers both in education and information and communication technology aim at understanding how technology coupled with skillful pedagogical solutions can help to innovate and improve learning at school. Since 2001, we have collaborated with educationalists, teachers, and students with the aim of developing systems for educational purposes in different domains, e.g. logic [13], history [9], and investigating the educational effectiveness of such systems and the overall user experience gained with these instructional instruments [9].

Several studies have demonstrated that technologies actually do improve student learning, but assessments conducted on technology programs have not always yielded significantly enhanced student performances [16], [20]. Our opinion is that technological tools cannot replace the teacher, but they can be considered a valid support for students in their learning paths [9], [13]. The availability of new technologies, ranging from latest generation mobile devices to gestural input, is bringing about profound changes into education. An effective way of exploiting technologies for learning purposes is to develop educational games [10], [12], [18]. Our experience in designing and evaluating educational games has demonstrated that games stimulate children's imagination, help them gauge their abilities and limits, and motivate them to learn through socialization, collaboration and constructive dialogues among themselves [4], [9]. Recently, we have investigated the use of large multi-touch screens in educational domain, because previous studies proved that this technology is enjoyable to use, promotes playfulness, supports awareness and fosters learning by allowing children to collaboratively enrich their knowledge [15], [17], [18]. Educational games are a type of *informal learning*, defined in [8] as a process whereby individuals acquire skills and knowledge from experiences, influences and resources in their whole environment, not restricted to school, from family and friends, from play, libraries, mass media, etc.

Educational games for large multi-touch screens that support children in consolidating knowledge they have acquired in more formal settings have been developed. Two examples of these games were presented at DET 2011, describing the design and development process of the games and reporting about the formative evaluation carried out in the university laboratory [3]. No summative evaluation at school was performed at that time.

Recently, we have integrated such games in an educational format, inspired by Bruner's discovery learning technique [7]. This format has been experimented in a field study involving 107 pupils of a primary school in Bari [5].

To further analyze how children were engaged by games on the large multi-touch screens, we let children interact with two different devices, i.e. a desktop vs. a large multi-touch screen. In this paper, the results of the data gathered and observations made during the study were reported.

The paper has the following organization. Section 2 briefly reviews games on multi-touch screens. The educational format and the results of its validation are summarized in Section 3 and Section 4, respectively. Section 5 presents the complementary study conducted for analyzing children's interactions with two devices (desktop vs. multi-touch screens). Section 6 concludes the paper.

II. EDUCATIONAL GAMES ON MULTITOUCH SCREEN

Since 2004, researchers have been increasingly

suggesting that tangible user interfaces might offer a fruitful support to children's learning, and various studies have been conducted since the advent of commercially available systems, such as Microsoft Surface and SMART table. To the best of our knowledge, in literature there are examples of the use of multi-touch interactive tabletop in the educational domain [1], [2], [15], [17], [19], but none about the use of a multi-touch vertical screen, except of "Us Hunters", an application that allows children to understand both the hunting strategy and the technique used by the painter to depict this activity [11].

In accordance with [11], we decided to use a vertical multi-touch screen. It is well known that multi-touch displays modify the traditional person-system interaction paradigm, not only because they allow different people to interact together at the same time but also because they can attract the attention of other people who are standing nearby. In a study we performed to analyze users' behavior and their experience with a large multi-touch display installed at an international conference [5], we observed a "honey pot effect", stating that people are more likely to interact with the device if it is already used by others. Our study results highlighted that people observing the interaction within an area of two meters around the screen moved towards the display and started to interact. For this reason, we decided to use a multi-touch screen vertically positioned.

The multi-touch technology we used consists of a MultiTouch Ltd 46-inch Full HD LCD display, called cell [14]. One of its peculiar characteristics is that a greater number of cells can be linked to create larger screens. Object manipulation is mediated by gestures performed with one or both hands (or some fingers). Students can: a) move an object on the screen with one hand just by touching the video; b) change the size of an object with the index fingers of both hands, by stretching or narrowing two corners of the object; c) rotate an object using the index fingers of both hands. In particular, we implemented two educational games aimed at stimulating pupils to exercise their knowledge about history and geography, called History-Puzzle and EuroFlags respectively [3].

III. GAMES ON MULTI-TOUCH SCREEN IN AN EDUCATIONAL FORMAT

The educational format we have defined integrates formal learning (traditional classroom lessons) with more informal and technology-based learning [6]. It, inspired by Bruner's Discovery Learning [7], organizes the learning activities in three phases and pupils get new information by: 1) attending the lesson(s) by their teacher in the classroom (*symbolic phase*), 2) acting in a real context (*active phase*), and 3) interacting with technological tools (*iconic phase*). While the activity carried out in the symbolic phase is always the classroom lesson(s), those performed in the other two phases (i.e. active and iconic) change according to the topic to be learned and the technology used. For example, if students have to learn about woodland habitat in their natural science curriculum, they firstly attend lesson/s at school, then they explore a woodland and lastly, back at school, they interact with an educational system.

We have experimented with this format for learning history and fostering a keen interest in cultural heritage in students aged 10 years old. Specifically, in the symbolic phase, children learn basic notions about ancient history by attending a lesson. After this, as an activity of the active phase, pupils go on a school trip to the archaeological park of the ancient city they are studying. Finally, back at school, they interact with History-Puzzle in the iconic phase of the educational format.

History-Puzzle is an educational game running on a multi-touch display whose aim is to support pupils learning the history of an ancient Roman city by proposing a set of puzzles, each of them depicts important places of the ancient site they have visited. Children have to reassemble puzzles by associating the two parts of a sentence that reports a historical notion. The nine square tiles covering a figure at the center of the display contain incomplete sentences. Players choose the rest of the sentence from the tiles displayed on the left and right sides of the puzzle and drag it into one of the nine tiles in the central zone (Figure 1).

Figure 1. The starting screen to solve the "Kiln" puzzle.

If the selected association is correct, the tile will reveal one ninth of the 3D reconstruction of the original place. Figure 2 shows a group of children playing with History-Puzzle that has discovered 6/9 of the "Fornace" (in English "Kiln") image. When the puzzle has been completed, different multimedia contents are proposed: a papyrus reporting all the puzzle sentences, a short video showing the 3D reconstruction of the place, sounds associated to the place (e.g. in the case of the kiln, people's voices, crackling fire); photos of the place, the half-section plan of a building.

A score of 5 points was awarded if the sentence was correctly combined. This score was reduced by 2 point every time participants moved a tile on a wrong one. The final score was stored and shown on a screen when the game was over. On this way, groups can know the score of the other companions.

Figure 2. A group of pupils interacting with History-Puzzle in a school laboratory. Six parts of the "Kiln" have been discovered.

IV. FORMAT VALIDATION

The educational format was validated during a field study involving 107 pupils (55 girls, average age 10) of the fifth class of the primary school "Clementina Perone" in Bari (Italy) conducted during November-December 2011. The study aimed at: 1) investigating the user experience of pupils interacting with multi-touch displays in a real context; 2) evaluating the effect of multi-touch technology as a means of consolidating knowledge; 3) assessing the effectiveness of the educational format. The study demonstrated that children enjoyed the game and were actively engaged. The multi-touch technology appeared very promising as a means of facilitating collaborative experiences. The study also highlighted the value of multi-touch technology as an important means to support knowledge consolidation. Children gave more correct answers to the true/false test performed after the multi-touch experience, confirming that technology is not an obstacle for children but, on the contrary, can be effective to support learning processes. Finally, results provided further evidence that the three phases of the educational format are important: information disregarded in one type of representation was then acquired in another. A more complete and comprehensive description of the results of this study is reported in [6].

V. DESKTOP VS MULTI-TOUCH SCREEN TO PLAY WITH HISTORY

A complementary study, involving the same children as the main study, was performed to analyze how children approach two different technologies (i.e. desktop vs. multi-touch display). The week after the main study, children went to the laboratory with their teachers and interacted with a game-based application for PC desktop. This application, called "Rino nella storia" (in English "Rino in history"), concerned history about ancient populations, i.e. Egyptians, Greeks, Romans, etc., whose study is part of the fifth class curriculum. Two days later, an entire class at a time went to the laboratory again, where they were free to play with History-Puzzle or with Rino nella storia. The week later, a focus group was held in each class, moderated by the teachers to gather information on the overall experience. Focus groups were video-taped and a research assistant noted down the most important observations.

Rino nella storia and History-Puzzle are different in terms of interface and interaction style. Rino nella storia runs on a traditional desktop while History-Puzzle is visualized on a multi-touch display. We are aware that such game differences might be seen as a study limitation, but we did not want to compare the two games, our goal was to investigate children's attitude with respect to the multi-touch application.

Observation revealed collaboration when children interacted both with History-Puzzle and Rino nella storia. Related to interactions with History-Puzzle, the children autonomously decided to divide up into groups. A total of 21 groups was observed. Most of them were composed of 3 children (5 groups), 4 children (6 groups) and 5 (6 groups). There were only 3 groups of 6 children and, finally, only 1 group of 8 children. Groups of 6 or 8 members had no success because some children did not have the chance to move the tiles, because the display was not big enough and those who remained on the sidelines did not become engaged in the game.

Children would have liked to play in groups also in the case of Rino nella storia but they were forced to play alone even if, in some cases, a child who was seated at a workstation was joined by her/his companion/s. Groups of 2 or 3 children were observed.

Twenty-one dispute situations arose, in which the use of the multi-touch display was a cause for argument. Although children decided to split into groups to ensure greater equity, they had to wait a long time, so many children kept going near the multi-touch display and insistently demanding: "Have you finished??". The answer they received was always the same: "We have to finish everything and then you can start!". If they did not succeed in interacting with History-Puzzle, they chose Rino nella storia, but this was viewed as second best.

During the focus group, all the children explicitly said that they enjoyed going to the laboratory and playing with the two games on the two different devices that allowed them to learn and explore concepts that their teachers had explained in classroom. But, they underlined that they preferred to interact with the multi-touch display. This was already evident, when children entered the laboratory: before the dynamics of the tasks they had to perform were explained, they moved directly towards the multi-touch display even if only to touch it and/or to move the tiles. Only 13 of 107 expressed a preference for Rino nella storia for the following reasons: they could choose among more different types of game; it could be played individually, thus reducing the risk of making mistakes for someone else's fault.

The multi-touch display tactile feature was one of the main reasons why pupils preferred to play with History-

Puzzle. Some of them were familiar with this technology, which is exploited by some cell phones, owned by many of them or known because their parents had them. Students expressed this important multi-touch characteristic with the words: "It was interesting and enjoyable playing with the multi-touch display because we had to search for the right answer using our fingers", or "It was like playing with a real puzzle!".

VI. CONCLUSION

Multi-touch technology is increasingly exploited by researchers to support students acquiring new knowledge. We implemented educational games on the multi-touch vertical screen supporting children in consolidating knowledge they have acquired. An educational format which combines traditional learning with more informal and technology-enhanced learning was proposed to try to overcome problems that might occur when technology alone is used at school. We have experimented with this format for learning history and a field study demonstrated that the educational format is effective and that applications on the multi-touch display can be a valid means for consolidating knowledge. In this paper, the results of a complementary study in which two different devices (i.e. desktop vs. multi-touch) were compared were reported. Results of this comparison highlighted that children much preferred interacting with the large multi-touch screen, that they found much more engaging, and they were enthusiastic about the chance to use their fingers and hands to move objects about on the screen without the need of other external and technological tools.

ACKNOWLEDGMENT

Partial support for this research is provided by Italian MIUR through grants "L4A". We thank Carmelo Ardito for his valuable contribution in all the design and evaluation phase of the multi-touch educational games, the Prof. Floriana Falcinelli of the University of Perugia for her important suggestions of the pedagogical issues of our educational format, and the teachers and students of the school “Clementina Perone” in Bari, Italy, for participating in the field study.

REFERENCES

[1] AlAgha, I., Hatch, A., Ma, L. and Burd, L. 2010. Towards a teacher-centric approach for multi-touch surfaces in classrooms. In Proc. of ITS 2010. ACM, New York, NY, USA, 187-196.

[2] Antle, A. N., Wise, A. F. and Nielsen, K. 2011. Towards Utopia: designing tangibles for learning. In Proc. of IDC '11. ACM, New York, NY, USA, 11-20.

[3] Ardito C., Buono P., Costabile M. F., and Lanzilotti R. 2011. Educational games on a large multi-touch screen. In Proc. of DMS 2011. Knowledge Systems Institute, Skokie, Illinois, USA, 242-245.

[4] Ardito, C., Buono, P., Costabile, M. F., Lanzilotti, R. and Piccinno, A. 2009. Enabling Interactive Exploration of Cultural Heritage: An Experience of Designing Systems for Mobile Devices. Knowledge, Technology & Policy, 22, 1, 79-86.

[5] Ardito, C., Costabile, M.F., Lanzilotti, R., De Angeli, A., and Desolda, G. 2012. A Field Study of a Multi-Touch Display at a Conference. In Proc. of AVI 2012. ACM Press, New York, NY, USA, 580-587.

[6] Ardito, C., Lanzilotti, R. and Costabile, M. F. 2012. Integrating traditional learning with educational games on large multi-touch displays. Submitted to an International Journal.

[7] Bruner, J. 1960. The Process of Education. Cambridge, MA, Harvard University Press.

[8] Conner, M. L. Informal Learning. Available at: http://marciaconner.com/resources/informal-learning, 2012.

[9] Costabile, M. F., Angeli, A. D., Lanzilotti, R., Ardito, C., Buono, P. and Pederson, T. 2008. Explore! possibilities and challenges of mobile learning. In Proc. of CHI'08. ACM, New York, NY, USA, 145-154.

[10] Kardan, K. 2006. Computer role-playing games as a vehicle for teaching history, culture, and language. In Proc. of Sandbox 2006. ACM, New York, NY, USA, 91-93.

[11] Kourakis, S. and Pares, N. 2010. Us Hunters. Interactive Communication for Young Cavemen. In Proc. of IDC ’10. ACM, New York, NY, USA, 89-97.

[12] Kristensson, P. O. and Zhai, S. 2007. Learning shape writing by game playing. In Proc. of CHI'07. ACM, New York, NY, USA, 1971-1976.

[13] Lanzilotti, R. and Roselli, T. 2007. An Experimental Evaluation of Logiocando, an Intelligent Tutoring Hypermedia System. Int. J. Artif. Intell. Ed., 17, 1, 41-56.

[14] Multitouch Ltd. http://multitouch.fi/. Jan, 12th, 2012.

[15] Olson, I. C. and Horn, M. S. 2011. Modeling on the table: agent-based modeling in elementary school with NetTango. In Proc. of IDC '11. ACM, New York, NY, USA, 189-192.

[16] Oviatt, S. L. 2012. The future of educational interfaces. Routledge Press, to appear.

[17] Rick, J., Marshall, P. and Yuill, N. 2011. Beyond one-size-fits-all: how interactive tabletops support collaborative learning. In Proc. of IDC '11. ACM, New York, NY, USA, 109-117.

[18] Rogers, Y., Price, S., Randell, C., Fraser, D. S., Weal, M. and Fitzpatrick, G. 2005. Ubi-learning integrates indoor and outdoor experiences. Commun. ACM, 48, 1 (2005), ACM, New York, NY, USA, 55-59.

[19] Sluis, R. J. W., Weevers, I., Schijndel, C. H. G. J. v., Kolos-Mazuryk, L., Fitrianie, S. and Martens, J. B. O. S. 2004. Read-It: five-to-seven-year-old children learn to read in a tabletop environment. In Proc. of IDC '04. ACM, New York, NY, USA, 73-80.

[20] Zucker, A. and Light, D. 2009. Laptop programs for students. Science.

Software Product Line Evaluation: Categorization and Evolution over the Years

Edson A. Oliveira Junior and Itana M. S. Gimenes
Informatics Department - State University of Maringá
Maringá-PR, Brazil
Email: {edson, itana}@din.uem.br

José C. Maldonado
Computing Systems Department - University of São Paulo
São Carlos-SP, Brazil
Email: jcmaldon@icmc.usp.br

***Abstract*—Research on software product line evaluation has resulted in a variety series of approaches and techniques. Such approaches and techniques have evolved all over the years. However, they are not clearly mapped, thus supporting the search for respective information. This paper aims to present the landscape of research on software product line evaluation approaches and techniques, highlighting their categorization and how they have evolved over the years. We conducted a systematic literature review to identify and categorize the most important software product line evaluation approaches and techniques and their information sources. Although many software product line evaluation approaches and techniques are proposed in the literature, there is a lack of structural-based SPL architecture evaluation approaches, such as, quality attribute metric-based evaluations and quantitative analysis.**

***Keywords*-Evaluation, Metrics, Quality Attribute, Software Product Line, Systematic Literature Review.**

I. Introduction

A Software Product Line (PL) represents a set of systems sharing common features that satisfy the needs of a particular market or mission segment [18], [22]. This set of systems is also called a product family. The family's members are specific products developed in a systematic way from the SPL core assets. The core asset has a set of common features as well as a set of variable parts, which represent later design decisions [22]. The composition and the configuration of such assets yield specific products.

SPL adoption has been increasing in recent years. Success stories from several companies have been reported in the literature, including: Philips, Bosch, Nokia, and Toshiba [18]. The benefits obtained with the SPL approach include [18]: better understanding of the domains, more artifact reuse, and less time to market. The adoption of a SPL approach requires long term management, since the benefits are not immediately obtained. A wide variety of products should be produced to consolidate the SPL adoption and, therefore, provide a return on investment (ROI) [2]. In order to produce SPL products and to take advantage of SPL benefits, there must be an intensive SPL architecture design and evaluation, as well as an efficient variability management.

Evaluation of SPLs is important from both academic and industrial view points due to its potential to increase the productivity and the quality of products, decrease the time to market, improve the SPL production capability [18], and to be used as a parameter for evaluating the SPL quality [11]. A SPL encompasses variabilities, which increase the effort to evaluate its quality [10], [11]. A SPL evaluation can reveal several issues including: (i) which quality attribute must be prioritized for the development and the evolution of SPL architectures by means of a trade-off analysis; (ii) which assets represent the main variabilities of a SPL and how they impact the overall SPL during maintenance activities; and (iii) whether the SPL business goals are satisfied for an organization.

Research on SPL evaluation comprises either: (i) an extension of an existing approach/technique, such as the ATAM (Architecture Tradeoff Analysis Method) [6] which is based on quality attributes and scenarios; (ii) the proposal of new approaches/techniques to conform to existing SPL development activities.

Although there are several approaches and techniques for SPL evaluation in the literature, to date there is not a clear categorization of their evolution. Therefore, feedback on how SPL evaluation approaches/techniques are categorized might stimulate and guide researchers on proposing or extending such approaches/techniques by taking into consideration their extents. Therefore, this paper draws a landscape of research on SPL evaluation, tackling the following issues:

- How SPL evaluation approaches and techniques can be categorized?
- How SPL evaluation approaches and techniques have evolved through the years?

In order to support answering such questions, a systematic literature review (SLR) was planned and conducted. Its results show that there is a lack of evaluation approaches and techniques based on the structure of a SPL, such as, SPL architecture metrics for quantitative analysis.

The remainder of this paper is organized as follows: Section II presents the data collection by means of the planning and conducting of the SLR; Section III analyzes and discusses the obtained results tackling the SPL evaluation

issues of this SLR; Section IV lists related work; and Section V presents conclusion and directions for future work.

II. Data Collection

This section aims at summarizing the collected data during our SLR. Our research goals (Section II-A) to identify our research of interest are presented, as well as the achieved results (Section II-B).

A. Planning and Conducting the Review

The systematic literature review is a process that rigorously drives one for identifying, evaluating and interpreting the research available for a certain topic of interest [3], [17]. A *primary study*[1] represents the smallest piece of work that provides evidence of a specific subject.

Two main goals are defined for our SPL evaluation SLR: (i) identify how existing studies are related to one another; and (ii) analyze the existing studies evolution over the years. Therefore, the primary question (PQ) and secondary questions (SQ) are defined as follows:

- **PQ#1: Which approaches and techniques have been applied to evaluate and assess software product lines?**
  - ***SQ#1:*** *Which approaches and techniques have been applied to evaluate and assess SPL and how they can be categorized?*
  - ***SQ#2:*** *How the SPL evaluation approaches and techniques have evolved over the years?*

In order to succeed on a SLR, an adequate search string must be defined. Such a string totally depends on the expertise of the researches involved on a SLR. It should be as comprehensive as possible to reach the maximum number of studies. Thus, the following string was defined for our SLR:

```
software AND (product-line OR product line OR
product-family OR product family OR family of
products) AND (assessment OR architecture assessment
OR architectural assessment OR evaluation OR
architecture evaluation OR architectural evaluation OR
architectural analysis)
```

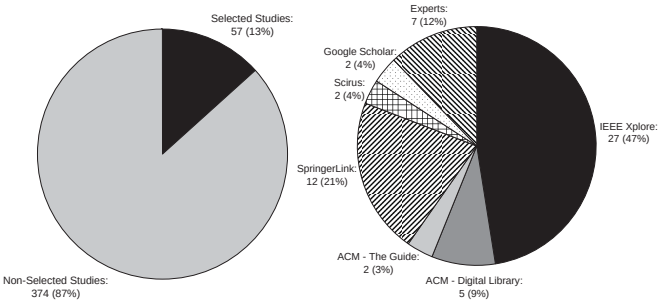
The search string is applied to predefined study sources, such as, IEEE Xplore, ACM, SpringerLink, and Elsevier.

The review procedures - preliminary selection, final selection, data extraction, and review documentation - follow the Biolchini et al.'s template [3]. For performing the preliminary selection, one goes through important parts of a study, such as, title and abstract to check whether such a study should be selected for full reading. In final selection step, one reads fully the studies and decides whether they are selected or should be discarded, according to predefined inclusion and exclusion criteria. In the data extraction step, data of interest is extracted and stored in proper forms. The SLR documentation is performed during the review process to allow replication of a SLR [3], [17].

[1]In this paper we use the shortest term *study* to refer to a primary study.

B. Results

The application of the search string (Section II-A) returned a total of 431 studies, from which 57 (13%) were selected (Figure 1a). Figure 1b presents the selected studies percentage by source as, for instance, IEEE Xplore which returned 27 studies, representing 47% of the total.



(a) (b)

Figure 1: Selected and Non-Selected Studies and Respective Sources.

Table I lists all selected studies sorted by the year of publication.

Table I: Selected Studies (Sorted by Year of Publication).

Ref	Title	Author(s)	Category	Year
1	Designing System on a Chip Products using Systems Engineering Tools	Hellestrand, G. R.	Other	1999
2	Stakeholder-Centric Assessment of Product Family Architecture	Dolan, T.; Weteringa, R.; Wortmann, J. C.	Quality attribute-based	2000
3	An Approach to Architectural Analysis of Product Lines	Gannod, G. C.; Lutz, R. R.	Quality attribute-based	2000
4	Software Product Lines: Practices and Patterns	Clements, P.; Northrop, L.	Quality attribute-based	2001
5	Product Line Software Engineering of Embedded Systems	Niemelä, E.; Ihme, T.	Context-based	2001
6	An Assessment Approach to Analyzing Benefits and Risks of Product Lines	Schmid, K.	Context-based	2001
7	MAP: Mining Architectures for Product Line Evaluations	Stoermer, C.; O'Brien, L.	Context-based	2001
8	Measuring Product Line Architectures	Dincel, E.; Medvidovic, N.; van der Hoek, A.	Structural-based	2002
9	Reviewing Product Line Architectures: Experience Report of ATAM in an Automotive Context	Ferber, S.; Heidl, P.; Lutz, P.	Quality attribute-based	2002
10	Experiences in Assessing Product Family Software Architecture for Evolution	Maccari, A.	Quality attribute-based	2002
11	Developing, validating and Evolving an Approach to Product Line Benefit and Risk Assessment	Schmid, K.; John, I.	Context-based	2002
12	Using the Architecture Tradeoff Analysis Method (ATAM) to Evaluate the Software Architecture for a Product Line of Avionics Systems: A Case Study	Barbacci, M.; Clements, P.; Lattanze, A.; Northrop, L.; Wood, W.	Quality attribute-based	2003
13	Creating Product Line Architectures	Bayer, J.; Flege, O.; Gacek, C.	Other	2003
14	Tricks and Traps of Initiating a Product Line Concept in Existing Products	Ebert, C.; Smouts, M.	Other	2003
15	Goal-Oriented Assessment of Product-Line Domains	Geppert, B.; Weiss, D. M.	Context-based	2003
16	Using Service Utilization Metrics to Assess the Structure of Product Line Architectures	Hoek, A.; Dincel, E.; Medvidovic, N.	Structural-based	2003
17	Software Product Family Evaluation	Linden, F.; Bosch, J.; Kamsties, E.; Känsälä, K.; Krzanik, L.; Obbink, J. H.	Other	2003
18	Analysis of a Software Product Line Architecture: an Experience Report	Lutz, R. R.; Gannod, G. C.	Quality attribute-based	2003
19	Experiences with Software Product Family Evolution	Riva, C.; Rosso, C. D.	Quality attribute-based	2003
20	Quality Prediction and Assessment for Product Lines	Zhang, H.; Jarzabek, S.; Yang, B.	Quality attribute-based	2003
21	An Evaluation of Aspect-Oriented Programming as a Product Line Implementation Technology	Anastasopoulos, M.; Muthig, D.	Other	2004
22	Scenarios, Quality Attributes, and Patterns: Capturing and using their Synergistic Relationships for Product Line Architectures	Babar, M. A.	Other	2004
23	Calculating ROI for Software Product Lines	Bockle, G.; Clements, P.; McGregor, J. D.; Muthig, D.; Schmid, K.	Context-based	2004
24	Analysis and Prediction of Performance for Evolving Architectures	Eskenazi, E. M.; Fioukov, A. V.; Hammer, D. K.; Obbink, H.; Pronk, B.	Structural-based	2004
25	Evaluating the Portability and Maintainability of Software Product Family Architecture: Terminal Software Case Study	Matinlassi, M.	Quality attribute-based	2004
26	Arcade Game Maker Product Line: Architecture Evaluation Report	McGregor, J. D.	Quality attribute-based	2004
27	Practical Evaluation of Software Product Family Architectures	Niemelä, E.; Matinlassi, M.; Taulavuori, A.	Context-based	2004
28	Metrics for the Structural Assessment of Product Line Architecture	Rahman, A.	Structural-based	2004
29	The Process of and the Lessons Learned from Performance Tuning of a Product Family Software Architecture for Mobile Phones	Rosso, C. D.	Quality attribute-based	2004
30	Introducing PLA at Bosch Gasoline Systems: Experiences and Practices	Steger, M.; Tischer, C.; Boss, B.; M uller, A.; Pertler, O.; Stolz, W.; Ferber, S.	Context-based	2004
31	Product-Line Architecture: New Issues for Evaluation	Etxeberria, L.; Sagardui, G.	Other	2005
32	Design Verification for Product Line Development	Kishi, T.; Noda, N.; Katayama, T.	Other	2005
33	Comparing Design Alternatives from Field-Tested Systems to Support Product Line Architecture Design	Knodel, J.; Forster, T.; Girard, J.-F.	Other	2005
34	Asset Recovery and their Incorporation into Product Lines	Knodel, J.; John, I.; Ganesan, D.; Pinzger, M.; Usero, F.; Arciniegas, J. L.; Riva, C.	Other	2005
35	From Scenarios to Aspects: Exploring Product Lines	Krüeger, I. H.; Mathew, R.; Meisinger, M.	Other	2005
36	Cost Estimation for Product Line Engineering using COTS Components	Lamine, S. B. A. B.; Jilani, L. L.; Gh_ezala, H. H. B.	Context-based	2005
37	Extending the ATAM Architecture Evaluation to Product Line Architectures	Olumofin, F. G.; Misic, V. B.	Quality attribute-based	2005
38	Experiences with Product Line Development of Embedded Systems at Testo AG	Kolb, R.; John, I.; Knodel, J.; Muthig, D.; Haury, U.; Meier, G.	Quality attribute and Structural-based	2006
39	A Software Fault Tree Metric	Needham, D.; Jones, S.	Other	2006
40	A Quality-Based Cost Estimation Model for the Product Line Life Cycle	Peter In, H.; Baik, J.; Kim, S.; Yang, Y.; Boehm, B.	Context-based	2006
41	Defining a Strategy to Introduce a Software Product Line using Existing Embedded Systems	Yoshimura, K.; Ganesan, D.; Muthig, D.	Context-based	2006
42	Fuzzy Inference System for Software Product Family Process Evaluation	Ahmed, F.; Capretz, L.; Samarabandu, J.	Context-based	2008
43	Evaluation of Quality Attribute Variability in Software Product Families	Etxeberria, L.; Sagardui, G.	Quality attribute-based	2008
44	Variability Driven Quality Evaluation in Software Product Lines	Etxeberria, L.; Sagardui, G.	Quality attribute-based	2008
45	Specification-Based Testing for Software Product Lines	Kahsai, T.; Roggenbach, M.; Schlingloff, B.	Context-based	2008
46	A Case Study on Architectural Maturity Evaluation: Experience in the Consumer Electronics Domain	Kim, K.	Structural-based	2008
47	A Case Study on SW Product Line Architecture Evaluation: Experience in the Consumer Electronics Domain	Kim, K.; Kim, H.; Kim, S.; Chang, G.	Structural-based	2008
48	Extending ATAM to Assess Product Line Architecture	Kim, T.; Ko, I. Y.; Kang, S. W.; Lee, D. H.	Quality attribute-based	2008
49	InCoME: Integrated Cost Model for Product Line Engineering	Nobrega, J.; Almeida, E.; Meira, S.	Context-based	2008
50	Software Performance Tuning of Software Product Family Architectures: Two Case Studies in the Real-Time Embedded Systems Domain	Rosso, C. D.	Quality attribute-based	2008
51	Using GQM for Testing Design Patterns in Real-Time and Embedded Systems on a Software Production Line	Bernardo, C.; Fernandes, D.; Dias, L.; Montini, D.; da Silva, D.; da Cunha, A.	Context-based	2009
52	Variability Assessment in Software Product Families	Deelstra, S.; Sinnema, M., Bosch, J.	Quality attribute-based	2009
53	Gathering Current Knowledge About Quality Evaluation in Software Product Lines	Montagud, S.; Abrahão, S.	Other	2009
54	Analyzing Structure-based Techniques for Test Coverage on a J2ME Software Product Line	Silva, L.; Soares, S.	Context-based	2009
55	An Organizational Maturity Model of Software Product Line Engineering	Ahmed, F.; Capretz, L.	Context-based	2010
56	A Business Maturity Model of Software Product Line Engineering	Ahmed, F.; Capretz, L.	Context-based	2010
57	Automated Test Data Generation on the Analyses of Feature Models: A Metamorphic Testing Approach	Segura, S.; Hierons, R. M.; Benavides, D.; Ruiz-Cortè ands, A.	Context-based	2010

Category column in Table I was proposed based on similar characteristics of the selected studies identified during the data extraction step (see Section III-A). Thus, the following categories of SPL evaluation were proposed:

- **Quality Attribute-based Evaluation** - research that focuses on the use of scenarios, such as the GQM (Goal-Question-Metric) approach [1], the ATAM and SAAM (Software Architecture Analysis Method) methods [6], and Architecture Description Languages (ADL) [19] for evaluating SPL architectures;
- **Structural-based Evaluation** - research that focuses on the use of metrics to evaluate SPL architecture components;
- **Context-based Evaluation** - research that focuses on the use of cost and economical models, as well as SPL testing; and
- **Other** - research with respect to the evaluation of a SPL in general, not taking a specific approach or technique into consideration.

Table II summarizes our SLR in terms of the number of selected studies per category. Note that in Table I, study Ref. #38 is categorized as Quality attribute- and Structural-based evaluation. Therefore, in Table II, at column 2006, half point is awarded to Quality attribute-based and half point to Structural-based. We decided to split points to keep the number of selected studies consistent with Table I.

Table II: Distribution of Selected Studies through the Years.

	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	Total	Average
Quality attribute-based	0	2	1	2	4	3	1	0.5	0	4	1	0	**18.5**	**1.54**
Structural-based	0	0	0	1	1	2	0	0.5	0	2	0	0	**6.5**	**0.54**
Context-based	0	0	3	1	1	3	1	2	0	3	2	3	**19**	**1.58**
Other	1	0	0	0	3	2	5	1	0	0	1	0	**13**	**1.08**
Total	**1**	**2**	**4**	**4**	**9**	**10**	**7**	**4**	**0**	**9**	**4**	**3**	**57**	**4.75**
Average	**0.25**	**0.50**	**1.00**	**1.00**	**2.25**	**2.50**	**1.75**	**1.00**	**0.00**	**2.25**	**1.00**	**0.75**	**4.75**	

Figures 2 and 3 depict the distribution shown in Table II. Figure 2 shows how the number of selected studies in each category varies along the years. Figure 3 shows the representativeness of the selected studies in each category over the years. We can learn from this figure that quality attribute-based evaluations are frequently performed as this category usually takes into consideration qualitative analysis which is based on widely known techniques as ATAM e SAAM. Moreover, structural-based evaluations are less applied as there is a small number of metrics for SPL. Note that in both figures the year 2007 has no selected study, thus it can be considered as an outlier.

III. Analysis and Discussion

In this section we present our analysis and discussion, based on the results of our SLR, which are focused on the main issues: (i) the categories of SPL evaluation; (ii) the evolution of SPL evaluation over the years; and (iii) the main sources of SPL evaluation research.

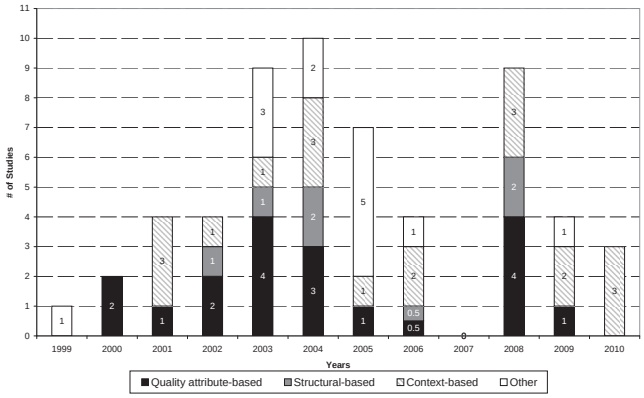


Figure 2: Number of Selected Studies in each Category along the Years.

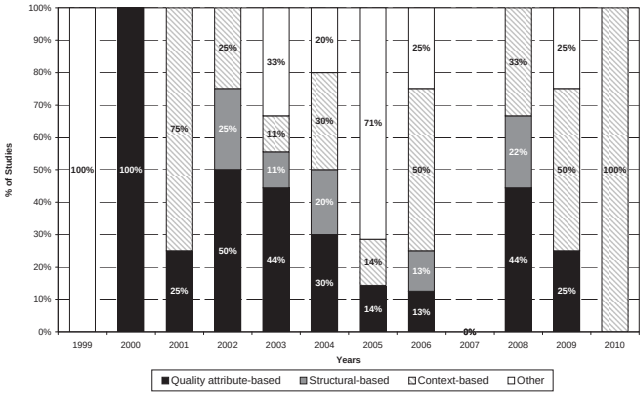


Figure 3: Representativeness of the Selected Studies in each Category over the Years.

The following sections try to answer the two questions presented in Section II-A.

A. Which approaches and techniques have been applied to evaluate and assess SPL and how they can be categorized?

After the reading of all selected studies (Table I) and the performing of data extraction, we classified the selected studies in the proposed categories of SPL evaluation, as follows:

- **Quality attribute-based** evaluation encompasses studies that take into consideration the SPL architecture and the application of:
  - scenarios (studies 3, 10, 18, 19, 25, 26, 29, 37, 50);
  - Architectural Description Language (ADL) (studies 3, 18, 48);
  - ATAM and SAAM methods (studies 2, 4, 9, 12, 26);
  - Goal-Question-Metric (GQM) (studies 20, 38); and
  - the quality attribute variability (studies 43, 44, 52).
- **Structural-based** evaluation encompasses studies with regard to the definition and application of metrics

to analyze SPL architecture components and quality attributes. Studies related to this category are: 8, 16, 24, 28, 38, 46, and 47.

- **Context-based** evaluation studies take into account:
  - SPL adoption (studies 6, 11, 30);
  - cost models (studies 23, 36, 40, 41, 49);
  - maturity models (studies 42, 55, 56);
  - candidate domain analysis (studies 5, 7, 15);
  - SPL testing (studies 45, 51, 54, 57); and
  - SPL overall analysis (studies 17, 27).
- **Other:** evaluation studies which encompass overall approaches and techniques: 1, 13, 14, 17, 21, 22, 31, 32, 33, 34, 35, 39, and 53.

We can observe in Table II that the Context-based evaluation category has the largest number of publications (19 out of 57 selected studies, i.e. 33.33% of the total) which represents 1.58 studies published in average per year. Quality attribute-based evaluation has the second largest number of studies (18.5 out of 57 selected studies, i.e 32.45% of the toal) which represents 1.54 studies published in average per year. Structural-based evaluation represents 11.40% of the total of selected studies and 0.54 study is published in average per year, whereas studies classified as Other represent 22.80% of the total and 1.08 studies published in average per year.

The year of 2004 had 10 studies published, whereas 2003 and 2008 had 9. On the contrast, 2007 had none studies published followed by 1999, 2000 and 2010, which had 1, 2 and 3 studies, respectively.

It can be observed, in this analysis, the lack of structural-based evaluation studies, which take into account mainly metrics for SPL measuring, compared to other categories. Metrics are essential to support SPL quantitative and improve qualitative analysis, as well as to estimate potential SPL products [2], [10], [11], [13], [23]. In addition, metrics can be applied to support both SPL quality attribute and context-based analysis.

According to Figures 2 and 3, metric-based studies for SPL only appeared in 2002 in Dincel et al.'s [7] study, representing 25% of the studies in such a year. The same occured in 2003 in Hoek et al.'s [13] study, when only one study was published. In 2004 (Eskenazi et al. [9] and Rahman [23]) and 2008 (Kim [15] and Kim et al. [16]) two studies were published in each year. Therefore, collected data provides evidence that might be difficult to analyze the structure of a SPL by means of its components and architectures. One of the reasons might be the fact that metric-related evaluation involves theoretical and empirical validations of metrics [4], [21], which are not trivial, but make evaluations more reliable.

Category combination is clearly feasible as shown in study #38 (Table I), in which metrics are used to measure SPL quality attributes for embedded systems.

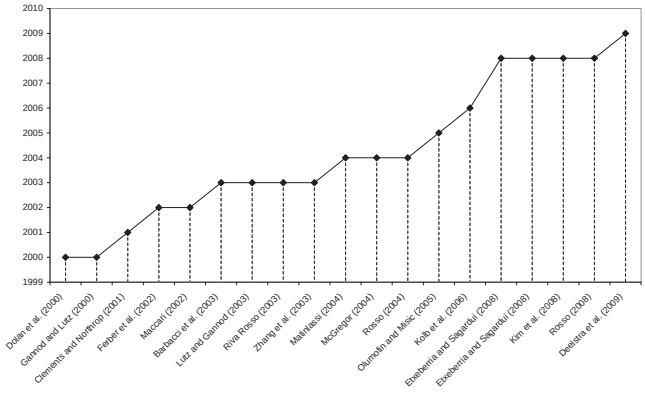


Figure 4: Quality Attribute-based Evaluation Evolution.

B. How the SPL evaluation approaches and techniques have evolved over the years?

Evolution of selected studies (Table I), in each proposed category, over the years is discussed in the following paragraphs.

Quality attribute-based: the first initiatives on quality attribute-based evaluation of PL (Figure 4, 2002-2003) took into consideration the ATAM/SAAM methods and scenarios. In 2003, Lutz and Gannod's and Zhang et al.'s researches applied ADLs and GQM, respectively, to represent variability and identify potential measures for PL. Etxeberria and Sagardui's and Deelstra et al.'s work investigated variability as quality attribute by applying combined techniques, such as scenarios and business drivers identification, to improve PL evaluation.

Context-based: as shown in Figure 5, from 2001 to 2003, context-based evaluation was characterized by research on the analysis of candidate domains, and adoption of PL. From 2004 to 2006, Bockle et al.'s, Lamine et al.'s, Peter In et al.'s, and Yoshimura et al.'s research aimed at define and apply cost models to improve PL evaluations. In 2008, Ahmed et al.'s and Kahsai et al.'s works were the first initiatives to introduce the concept of maturity models and testing into the PL domain.

Structural-based: Figure 6 depicts the evolution of structural-based evaluation of PL. In 2002, Dincel et al.'s research proposed a set of architectural metrics aimed at making decisions during design and evolution of PLs. Metrics to evaluate the PL architecture structure based on component service utilization were introduced in 2003. In 2004, Eskenazi et al.'s work proposed the APPEAR method to analyze and estimate PL architecture performance by applying metrics, such as response time, latency, and CPU average utilization. Kolb et al.' research defined metrics based on GQM for PL and products architectures conformance evaluation. Kim's and Kim et al.'s evaluation works are based on the static implementation of a PL architecture.

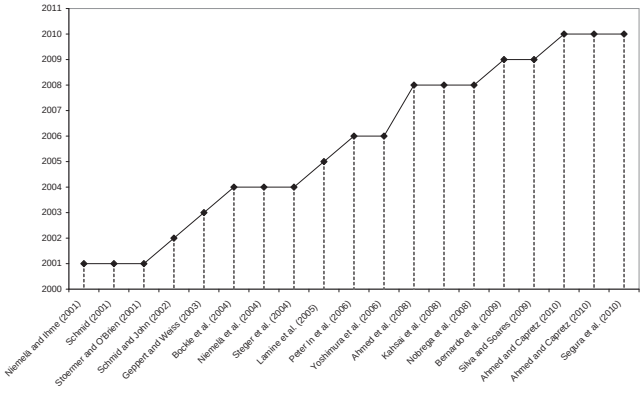


Figure 5: Context-based Evaluation Evolution.

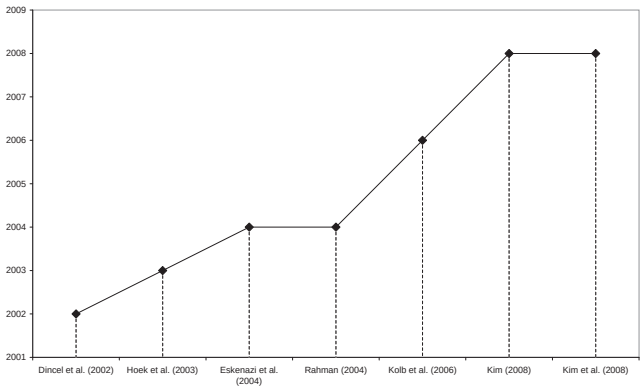


Figure 6: Structural-based Evaluation Evolution.

IV. Related Work

Montagud and Abrahão [20] present a SLR, limited to the IEEE Xplore, ACM and INSPEC digital libraries, about current knowledge of SPL evaluation. A total of 39 studies were reviewed. Such a SLR extracted important data, including: (i) type of approach (method or technique); (ii) lifecycle phase in which an evaluation takes place; (iii) artifacts evaluated; (iv) most widely used mechanisms for capturing quality attributes; (v) type of validation (survey, case study, controlled experiment, none); (vi) impact analysis; and (vii) actual application (academic or industrial). The Montagud and Abrahão SLR revealed that most evaluation takes place at the domain engineering activity, specifically at design time. Most evaluations takes into consideration the base architecture and the core assets. Qualitative evaluations are the most performed type of evaluation, whereas case studies are most carried out as validation procedures. Our SLR is closely related to Montagud and Abrahão's, as both try to answer related research questions by extracting complementary data.

Existing literature also presents related studies that aim at gathering knowledge related to SPL evaluation. Khurum et al. [14] present a systematic review of SPL economics solutions. It is essential as SPL evaluations can be used as a parameter to draw conclusions on SPL economics. Souza Filho et al. [12] performed a systematic review for analyzing domain design approaches, which might be useful to capture SPL architecture quality attributes to support SPL architecture evaluations. Chen et al. [5] performed a SLR about evaluation of variability management approaches in SPL. SPL variabilities are essential to perform SPL evaluations as they support the prediction and estimation of the SPL behavior based on its potential products. In addition, variability management has been widely investigated in the literature as a successful approach to postpone design decisions and produce a variety of products for a certain domain. Edwin [8] performed a SLR with regard to SPL testing. In such a SLR, research questions were addressed related to the approaches that can be used to develop test cases, and correlation between software reuse and reusability of test cases. Testing is essential to SPL as it can ensure the correctiveness of SPL artifacts to be taken into consideration for SPL evaluations.

V. Conclusion and Future Work

This paper presented a SLR which was planned and conducted to gathering essential knowledge about evaluation of software product lines. Such a SLR was performed in terms of which approaches and techniques have been applied over the years and how they can be categorized.

As a direct result, we proposed categories for SPL evaluation studies: Quality Attribute-based, Structural-based, Context-based, and Other. The 57 selected studies were classified taking into account the proposed categories, thus giving researchers directions on how they are grouped and can be rapidly related to similar studies. The Context-based category had 33.33% of selected studies, whereas Quality Attribute-based had 32.45% and Structural-based had 11.40%.

Evolution of the 57 selected studies of our SLR was presented, given researches a basis to understand which techniques have been applied over the years towards improvement to current research.

The SLR's overall results indicated that there are various approaches and techniques for SPL evaluation. Some of these approaches and techniques extend existing software architecture evaluation method as ATAM and SAAM by applying widely known techniques, such as, GQM, scenario-based analysis, and quality attributes capturing based on business drivers. Most of these techniques allows only qualitative analysis of SPLs. Although metrics, such as for structural soundness and service utilization, are proposed and applied to evaluate SPL architecture, they do not take into consideration the SPL variabilities. The lack of quantitative analysis approaches for SPL variabilities is clearly a research gap in the field. This kind of analysis gives SPL manager and architects a means to predict the SPL behavior

by collecting metrics from SPL models and analyzing the SPL architecture quality attributes by performing trade-off analysis.

Further SLRs must be planned and conducted to confirm our obtained results. Directions for future work might be: including other sources, such as, Science Direct, Scopus databases, and the INSPEC digital library; adding secondary questions to further map the existing literature of SPL evaluation taking into consideration the related work presented in this paper; and proposing a SPL evaluation approach based on the gathered knowledge.

REFERENCES

[1] V. R. Basili, G. Caldiera, and H. D. Rombach, "Goal Question Metric Paradigm," in *Encyclopedia of Software Engineering*. John Wiley & Sons, Inc., 1994, pp. 528–532.

[2] G. Böckle, P. Clements, J. D. McGregor, D. Muthig, and K. Schmid, "Calculating ROI for Software Product Lines," *IEEE Software*, vol. 21, no. 3, pp. 23–31, May 2004.

[3] J. Biolchini, P. G. Mian, A. C. C. Natali, and G. H. Travassos, "Systematic Review in Software Engineering," System Engineering and Computer Science Department. COPPE/UFRJ, Rio de Janeiro, Brasil, Tech. Rep. RT-ES 679/05, 2005.

[4] L. Briand, K. E. Emam, S. Morasca, K. El, and E. S. Morasca, "Theoretical and Empirical Validation of Software Product Measures," International Software Engineering Research Network, ISERN-95-03, 1995.

[5] L. Chen, M. A. Babar, and N. Ali, "Variability Management in Software Product Lines: a Systematic Review," in *Proceedings of the Software Product Line Conference*. Pittsburgh, PA, USA: Carnegie Mellon University, 2009, pp. 81–90.

[6] P. Clements, R. Kazman, and M. Klein, *Evaluating Software Architectures: Methods and Case Studies*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 2002.

[7] E. Dincel, N. Medvidovic, and A. v. d. Hoek, "Measuring Product Line Architectures," in *Proceedings of the International Workshop on Product Family Engineering*. London, UK: Springer-Verlag, October 2001, pp. 346–352.

[8] O. O. Edwin, "Testing in Software Product Lines," Master's thesis, School of Engineering - Blekinge Institute of Technology, Sweden, 2007.

[9] E. M. Eskenazi, A. V. Fioukov, D. K. Hammer, H. Obbink, and B. Pronk, "Analysis and Prediction of Performance for Evolving Architectures," in *Proceedings of the EUROMICRO Conference*. Washington, DC, USA: IEEE Computer Society, 2004, pp. 22–31.

[10] L. Etxeberria and G. Sagardui, "Product-Line Architecture: New Issues for Evaluation," in *Proceedings of the Software Product Line Conference*. Berlin, Germany: Springer-Verlag, 2005, pp. 174–185.

[11] ——, "Variability Driven Quality Evaluation in Software Product Lines," in *Proceedings of the Software Product Line Conference*. Washington, DC, USA: IEEE Computer Society, 2008, pp. 243–252.

[12] E. D. S. Filho, R. de Oliveira Cavalcanti, D. F. S. Neiva, T. Oliveira, L. B. Lisboa, E. de Almeida, and S. R. de Lemos Meira, "Evaluating Domain Design Approaches Using Systematic Review," in *Proceedings of the European Conference on Software Architecture*, Cyprus, 2008, pp. 50–65.

[13] A. v. d. Hoek, E. Dincel, and N. Medvidovic, "Using Service Utilization Metrics to Assess the Structure of Product Line Architectures," in *Proceedings of the International Symposium on Software Metrics*. Washington, DC, USA: IEEE Computer Society, 2003, pp. 298–308.

[14] M. Khurum, T. Gorschek, and K. Petterson, "Systematic Review of Solutions Proposed for Product Line Economics," in *Proceedings of the International Workshop on Management and Economics of Software Product Lines*, Limerick, Ireland, 2008, pp. 386–393.

[15] K. Kim, "A Case Study on Architectural Maturity Evaluation: Experience in the Consumer Electronics Domain," in *Proceedings of the OTM Confederated International Workshops and Posters on On the Move to Meaningful Internet Systems*. Berlin, Heidelberg: Springer-Verlag, 2008, pp. 364–373.

[16] K. Kim, H. Kim, S. Kim, and G. Chang, "A Case Study on SW Product Line Architecture Evaluation: Experience in the Consumer Electronics Domain," in *Proceedings of the International Conference on Software Engineering Advances*. Washington, DC, USA: IEEE Computer Society, 26-31 2008, pp. 192–197.

[17] B. Kitchenham, "Procedures for Performing Systematic Reviews," Department of Computer Science - Keele University and National ICT Australia Ltd., Tech. Rep. TR/SE-0401, 2004.

[18] F. J. v. d. Linden, K. Schmid, and E. Rommes, *Software Product Lines in Action: The Best Industrial Practice in Product Line Engineering*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2007.

[19] N. Medvidovic and R. N. Taylor, "A Classification and Comparison Framework for Software Architecture Description Languages," *IEEE Trans. Softw. Eng.*, vol. 26, pp. 70–93, January 2000.

[20] S. Montagud and S. Abrahão, "Gathering Current Knowledge About Quality Evaluation in Software Product Lines," in *Proceedings of the Software Product Line Conference*. Pittsburgh, PA, USA: Carnegie Mellon University, 2009, pp. 91–100.

[21] E. A. Oliveira Junior, I. M. S. Gimenes, and J. C. Maldonado, "Empirical Validation of Complexity and Extensibility Metrics for Software Product Line Architectures," in *2010 Fourth Brazilian Symposium on Software Components, Architectures, and Reuse*. Salvador-BA, Brasil: IEEE Computer Society, 2010, pp. 31–40.

[22] K. Pohl, G. Böckle, and F. J. v. d. Linden, *Software Product Line Engineering: Foundations, Principles, and Techniques*. Secaucus, NJ, USA: Springer-Verlag, 2005.

[23] A. Rahman, "Metrics for the Structural Assessment of Product Line Architecture," Master's thesis, School of Engineering - Blekinge Institute of Technology, Sweden, 2004.

Variability Management in Software Product Line Activity Diagrams

Diego R. Fiori*, Itana M. S. Gimenes*, José C. Maldonado† and Edson A. Oliveira Junior*
*Informatics Department - State University of Maringá - Brazil
Email: diegorafael.fiori@gmail.com, {itana, edson}@din.uem.br
†Computing Systems Department - University of São Paulo - Brazil
Email: jcmaldon@icmc.usp.br

***Abstract*—The software product line (PL) approach aims at promoting the generation of specific products based on the reusing of its core assets. Such core assets contain similar elements as well as features which vary according to specific products. The variability management has been an intensive and essential activity for the PL development. Several approaches are mainly focused on managing UML-based PL variabilities, more specifically in use cases and classes models. Therefore, this paper presents a proposal for managing variabilities in PL activity diagrams. Thus, the SMarty approach has been extended allowing the evolution of its UML profile, the SMartyProfile, and its variability management process, the SMartyProcess. The SEI's Arcade Game Maker (AGM) PL was used to illustrate the variability management in its main features activity diagrams.**

***Keywords*-Activity Diagram, Software Product Line, Variability Management, UML.**

I. Introduction

Software systems development is increasing every day as well its complexity and granularity. Reusing techniques can be seen as a realistic opportunity to decrease such a complexity by taking advantage of existing and well succeeded approaches, such as component-based development and software product lines. The software product line (PL) engineering approach has gained increasing attention over the last years due to competitiveness in the software development segment. The economic considerations of software companies, such as cost and time to market, motivate the transition from single-product development to the PL approach, in which products are developed in a large-scale reuse perspective [7].

A PL represents a set of systems sharing common features that satisfy the needs of a particular market or mission segment [7], [11]. This set of systems is also called a product family. The family's members are specific products developed in a systematic way from the PL core assets. The core asset has a set of common features as well as a set of variable parts, which represent later design decisions [11]. The composition and the configuration of such assets yield specific products.

The PL approach focuses mainly on a two-life-cycle model [7]: **domain engineering**, where the PL core asset is developed for reuse; and **application engineering**, where the core asset is reused to generate specific products. The success of the PL approach depends on several principles, in particular variability management [4], [6], [7], [11], [17]. However, most of the existing solutions are only applied to specific PL approaches. Thus, there is a lack of an overall reasoning about variability management applied to more general approaches [3] which take advantages of widely consolidated standard notations as, for instance, UML and its profiling extension mechanism [9] for specific domain applications. Although several approaches use this mechanism as a basis to represent variability [2], [4], [6], [18], most of them are not supported by a systematic process that provides guidelines to instruct the users on how to deal with variability issues in PL UML-based artifacts.

SMarty [8] is an approach for variability management in UML-based PL. This approach is supported both by a UML profile, the *SMartyProfile*, and a systematic variability management process, the *SMartyProcess*. *SMarty* makes it easier to deal with variability issues in PL. However, such an approach only supports use case, class and component models. Therefore, this paper presents an extension of the SMarty approach to encompass activity diagram models by means of the SMartyProfile and new guidelines for the SMartyProcess.

This paper is organized as follows: Section II presents basic concepts of UML activity diagrams; Section III discusses the main concepts of variability and the SMarty approach; Section IV proposes an extension of the SMarty approach by encompassing UML activity diagram models; Section V discusses related work; and Section VI presents the conclusion and directions for future work.

II. UML and Activity Diagrams

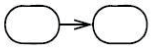
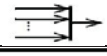
UML is a language for modeling object-oriented systems which is based in a wide vocabulary and concepts [9]. By means of its diagrams it is possible to represent software systems in such a way that stakeholders can communicate one another the systems main requirements and features.

Amongst the several UML diagrams, the activity diagram is one of the most important as it is responsible for communicating the actions via information/object flows. Moreover, such a diagram can represent artifacts produced during the execution of actions, as well as indicate partitions in which one might establish group of actions. The flow amongst

actions is usually based on decision nodes, as well as join and merge nodes.

Table I present the main elements of a UML activity diagram.

Table I
MAIN UML ACTIVITY DIAGRAM ELEMENTS [9].

UML Metamodel	Notation	Description
InitialNode		It indicates the beginning of a flow.
ActivityFinal		Final node of a diagram. It represents the end of a flow.
Action		It represents a step of an activity.
ControlFlow		It indicates a flow between two actions.
DecisionNode		It represents a choice among exit flows.
MergeNode		It merges various income flows to only one outcome flow.
ForkNode		It splits a flow in multiple concurrent flows.
JoinNode		It synchronizes multiple flows in only one exit flow.
ActivityPartition		It groups actions for a certain feature.

Activity diagrams have been used for several reasons as: (i) generating test cases [12]; (ii) Business Process Modeling (BPM) [10]; and (iii) modeling PL variabilities [13]–[15].

III. VARIABILITY MANAGEMENT IN SOFTWARE PRODUCT LINES

Variability is the general term used to refer to the variable aspects of the products of a PL. It is described through variation points and variants. A variation point is the specific place in a PL artifact to which a design decision is connected. Each variation point is associated with a set of variants that corresponds to design alternatives to resolve the variability [7], [11]. According to [7], [11], variability management is related to every activity of a PL approach and must comprise the following activities: **variability identification**, consists of identifying the differences between products and where they take place within PL artifacts; **variability delimitation**, defines the binding time and multiplicity of variabilities; **variability implementation**, is the selection of implementation mechanisms; **variant management**, controls the variants and variation points. These activities and the related concepts form the basis of the SMarty approach.

SMarty [8] is an approach for UML **S**tereotype-based **M**anagement of V**ar**iabili**ty** in PL. It is composed of a UML 2 profile, the *SMartyProfile*, and a process, the *SMartyProcess*. ***SMartyProfile*** contains a set of stereotypes and tagged values to represent variability in PL models. Basically, *SMartyProfile* uses a standard object-oriented notation and its profiling mechanism [9] both to provide an extension of UML and to allow graphical representation of variability concepts. Thus, there is no need to change the system design structure to comply with the PL approach. ***SMartyProcess*** is a systematic process that guides the user through the identification, delimitation, representation, and tracing of variabilities in PL models. It is supported by a set of application guidelines as well as by the *SMartyProfile* to represent variabilities.

The *SMartyProfile* represents the relationship of major PL concepts with respect to variability management. There are four main concepts: variability [1], variation point [11], variant [11], and variant constraints [1]. *SMartyProfile* is composed of the following stereotypes and respective tagged values:

≪**variability**≫ represents the concept of PL variability and is an extension of the metaclass `Comment`. This stereotype has the following tagged values: `name`, the given name by which a variability is referred to; `minSelection`, represents the minimum number of variants to be selected to resolve a variation point or a variability; `maxSelection`, represents the maximum number of variants to be selected in order to resolve a variation point or a variability; `bindingTime`, the moment at which a variability must be resolved, represented by the enumeration class `BindingTime`; `allowsAddingVar`, indicates whether it is possible or not to include new variants in the PL development; `variants`, represents the collection of variant instances associated with a variability; and `realizes`, a collection of lower-level model variabilities that realize this variability.

≪**variationPoint**≫ represents the concept of PL variation point and is an extension of the metaclasses `Actor`, `UseCase`, `Interface`, and `Class`. This stereotype has the following tagged values: `numberOfVariants`, indicates the number of associated variants that can be selected to resolve this variation point; `bindingTime`, the moment at which a variation point must be resolved, represented by the enumeration class `BindingTime`; `variants`, represents the collection of variant instances associated with this variation point; and `variabilities`, represents the collection of associated variabilities.

≪***variant***≫ represents the concept of PL variant and is an abstract extension of the metaclasses `Actor`, `UseCase`, `Interface`, and *Class*. This stereotype is specialized in four other non-abstract stereotypes which are: ≪mandatory≫, ≪optional≫, ≪alternative_OR≫, and ≪alternative_XOR≫. The stereotype ≪variant≫ has the following tagged values: `rootVP`, represents the variation point with which this variant is associated; and `variabilities`, the collection of variabilities with which this variant is associated.

≪**mandatory**≫ represents a compulsory variant that is

part of every PL product.

≪**optional**≫ represents a variant that may be selected to resolve a variation point or a variability;

≪**alternative_OR**≫ represents a variant that is part of a group of alternative inclusive variants. Different combinations of this kind of variants may resolve variation points or variabilities in different ways.

≪**alternative_XOR**≫ represents a variant that is part of a group of alternative exclusive variants. This means that only one variant of the group can be selected to resolve a variation point or variability;

≪**mutex**≫ represents the concept of PL variant constraint and is a mutually exclusive relationship between two variants. This means that when a variant is selected another variant must not also be selected;

≪**requires**≫ represents the concept of PL variant and is a relationship between two variants in which the selected variant requires the choice of another specific variant;

≪**variable**≫ is an extension of the metaclass `Component`. It indicates that a component has a set of classes with explicit variabilities. This stereotype has the tagged value `classSet` which is the collection of class instances that form a component.

Variability identification is a domain-dependent activity which requires abilities of the PL managers and analysts. Therefore, SMartyProcess provides guidelines to support the variability identification activity, which include:

G.1 elements of use case models related to the extend and extension points mechanism[1] [9] suggest variation points with associated variants which might be inclusive or exclusive alternative;

G.2 in class models, variation points and their variants are identified in the following relationships [9]: **generalization**, the most general classifiers are the variation points and the most specific ones are the variants; **interface realization**, the suppliers (specifications) are variation points and the implementations (clients) are the variants; **aggregation association**, the typed instances with hollow diamonds are the variation points and the associated typed instances are the variants; and **composite aggregation**, the typed instances with filled-in diamonds are the variation points and the associated typed instances are the variants;

G.3 elements of use case models related to the include (dependency) or association from actors relationships [9] suggest either mandatory or optional variants;

G.4 elements of class models related to the association relationship in which the `aggregationKind` attribute has value `none` [9], i.e., neither an aggregation nor a composition suggest either mandatory or optional variants;

[1]In SMarty, use case generalization is not used to represent variation points and variants as it does not represent the adding of specific actions from specialized use cases, as the extend relationship does [2].

G.5 variants that require the existence of other variants must tag their relationships as ≪requires≫;

G.6 mutually exclusive variants for a certain product must tag their relationships as ≪mutex≫; and

G.7 components, in component models, with variation point or variant classes are stereotyped as ≪variable≫.

IV. Managing Variability in Activity Diagrams

In order to gathering activity diagram variability management state of the art, a Systematic Literature Review (SLR) was performed. Such a SLR resulted in 223 primary studies, from indexed repositories (IEEE Xplore, ACM Digital Library, ScienceDirect and SpringerLink). From the total, 16 were selected for full reading. Thus, 5 studies were classified as relevant, which are discussed in Section V.

Based on the set of relevant studies, it was proposed an extension of the SMarty approach. Therefore, the following sections present such an extension, as well as an example of this extension applied to activity diagrams of the Arcade Game Maker (AGM) PL [16] main features.

A. Extending SMarty to Support Activity Diagrams

The proposed extension for SMarty is twofold: (i) an evolution of SMartyProfile by extending new metaclasses; and (ii) the addition of new guidelines for identifying and representing variability in activity diagrams.

The evolved version of the SMartyProfile (Figure 1) encompasses the extension of three metaclasses as follows:

- **`DecisionNode`** which might be considered a variation point and can be tagged with ≪variationPoint≫ as it allows the choice of different flows;
- **`Action`** which might be identified as a variant of any of the following types mandatory (≪mandatory≫), optional (≪optional≫), alternative inclusive (≪alternative_OR≫) or alternative exclusive (≪alternative_XOR≫); and
- **`ActivityPartition`** which might be defined with ≪variable≫ as it can encompass decision and action nodes with variability.

New guidelines were proposed for complementing the current SMartyProcess guidelines. Such new guidelines aim at identifying and representing variability in activity diagrams based on the three metaclasses extended in Figure 1:

G.8 DecisionNode elements in activity diagrams suggest variation points tagged with ≪variationPoint≫ as they explicitly represent multiple paths for distinct actions groups;

G.9 Action elements in activity diagrams might be defined as mandatory or optional variants tagged, respectively, with ≪mandatory≫ and ≪optional≫;

G.10 Action elements in activity diagrams which represent alternative exit flows from a DecisionNode

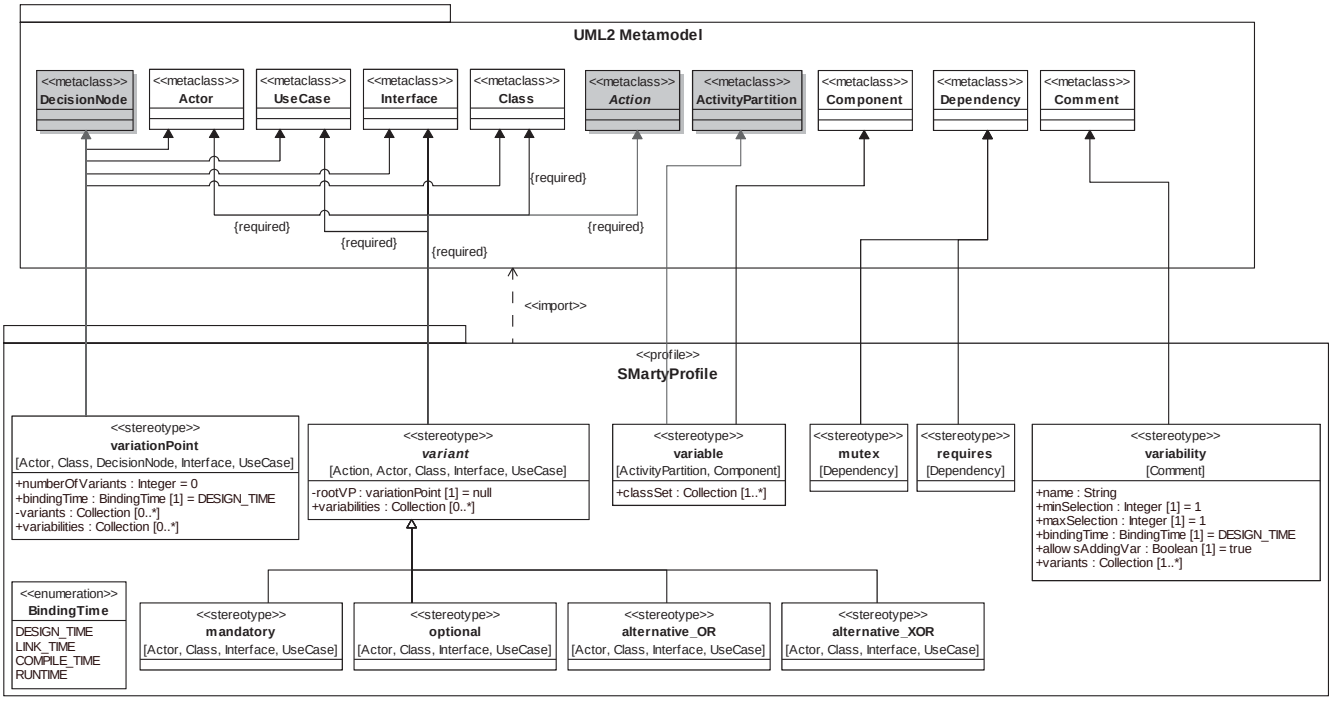


Figure 1. SMartyProfile Supporting Activity Diagram Elements.

suggest inclusive (≪alternative_OR≫) or exclusive (≪alternative_XOR≫) variants; and

G.11 ActivityPartition elements in activity diagrams which contain elements with associated variability, DecisionNode as a variation point or Action as a variant, must be tagged with ≪variable≫.

B. Application Example of the Extended SMarty

The Arcade Game Maker (AGM) [16] is a pedagogical PL for deriving games for mobile devices created by the SEI to support learning and experimenting based on PL concepts. Its main features are: play selected game, install game, save game, and load game. The essential AGM UML models [8] are the feature model, the use case model, the core asset class model, and the component architecture.

Figure 2 presents the `saveGame()` method activity diagram. Such a diagram was obtained by reverse engineering the AGM source code. This feature can be triggered by the `GamePlayer` actor represented by the ActivityPartition. A variation point is identified in a DecisionNode right after the action `String byteString = null;` according to guideline G.8. Thus, it is possible to decide which actions must be followed. The three possible actions are tagged with ≪alternative_OR≫ (according to guideline G.10) as a possible derived product might save more than one game. The DecisionNode path guard expressions decide which flow(s) must be followed according to guideline G.10. A variability named "record store" is associated with the variation point DecisionNode. Such a variability requires the selection of at least one (minSelection) and at most three (maxSelection) variants and it must be done at design time (bindingTime). In the case of the evolution of the AGM, this variability allows adding new variants (allowAddingVar). The ActivityPartition is tagged with ≪variable≫ according to guideline G.11 as it contains the variability "record store".

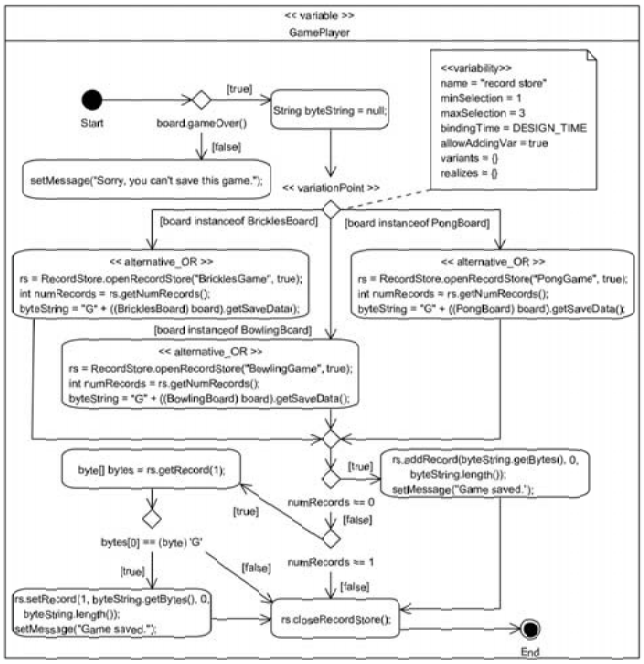


Figure 2. AGM saveGame() Method Activity Diagram.

Figure 3 presents the `loadGame()` method activity diagram. This feature can be triggered by the `GamePlayer` and `GameInstaller` actors represented by the ActivityPartition. A variation point is identified right after the action `byteString = new String(rs.getRecord(i));` according to guideline G.8. Thus, it is possible to decide which actions must be executed. The three possible actions are tagged with ≪alternative_OR≫ (according to guideline G.10) as a possible derived product might load more than one game. The DecisionNode path guard expressions decide which flow(s) must be followed according to guideline G.10. A variability named "record load" is associated with the DecisionNode. Such a variability requires the selection of at least one (minSelection) and at most three (maxSelection) variants and it must be done at design time (bindingTime). In the case of the evolution of the AGM, this variability allows adding new variants (allowAddingVar). The ActivityPartition is tagged with ≪variable≫ according to guideline G.11 as it contains the variability "record load".

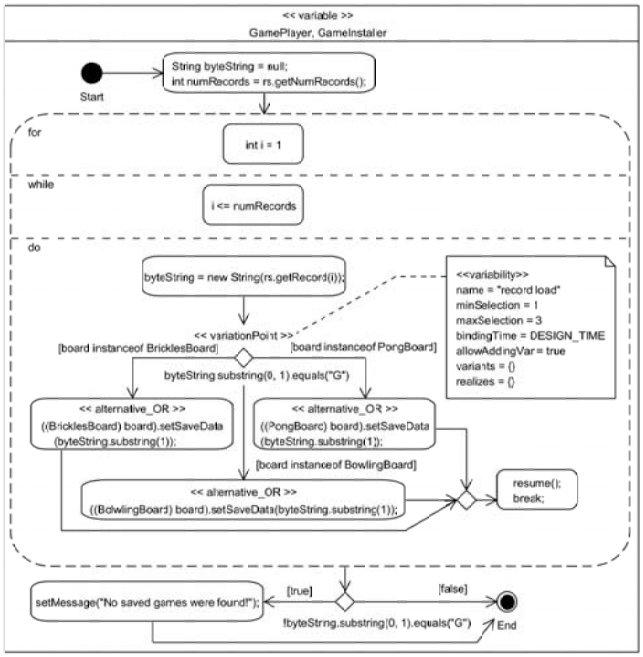


Figure 3. AGM loadGame() Method Activity Diagram.

V. Related Work

The existing literature provides a set of relevant related works which allow us to perform an overall analysis.

Robak et al. [14] propose the identification of variability in activity diagrams based on feature models with variability represented. Features are bound to representative methods or components of a PL. Each method/component is associated with an activity diagram in which variable actions are tagged with ≪variable≫. Moreover, each action is related to a corresponding feature. In the extended version of SMartyProfile, each action can be a variant related to one or more variation points. The variant's granularity is lower with relation to the strategy adopted by Robak et al. Therefore, it is possible to represent variations with lower granularity in an activity diagram, avoiding postponing them to components and their classes.

Reuys et al. [12] present an approach for deriving test scenarios from activity diagrams. Most of testing strategies is based on concepts as multiple paths and using/definition. Such paths can be seen as activity diagram inherent variabilities, from which case tests are defined. Thus, Reuys et al. propose the use of the stereotype ≪VP≫ to represent a variation point in a DecisionNode. Actions that represent each variation point decision node path are tagged with ≪variant≫. However, as discussed in previous works [3], [8], it is necessary to explicitly represent the kind of a variant (mandatory, optional, inclusive ou exclusive) in order to support the identification of the kind of variant-variation point relationship and guarantee the correct derivation of PL products. Therefore, SMartyProfile by means of its abstract stereotype ≪variant≫ and its concrete stereotypes ≪mandatory≫, ≪optional≫, ≪alternative_OR≫, and ≪alternative_XOR≫ contribute to make such a representation sharper.

Ripon et al. [13], as Reuys et al. [12], propose the use of the stereotype ≪variant≫ in activity diagram actions. However, such an approach aims at generating test cases, thus, providing no details with regard to variability-related concepts.

Shinieders [15] presents a variability mechanism-driven process for representing PL architecture. In such a process, macro-actions can be defined as variation points tagged with ≪VarPoint≫. Internally, each macro-action is composed of various actions and the flow amongst action pins can be variable ≪Variable≫. Such a representation approach is interesting. However, activity diagram with such a level of details were not identified in our SLR, thus, not yet taken into consideration by SMarty.

Heuer et. al. [5] present syntax and semantic formal definitions for variabilities in activity diagrams based on Petri nets. These definitions contributed to extend SMarty by analyzing how DecisionNodes are seen whereas they encompass possible executable paths. Such paths allow the variability representation in a PL as multiple paths can coexist in a same product. Moreover, Heuer et al.'s work indicates the need of adding formal expressions to the SMartyProfile variability models as, for instance, using the Object Constraint Language (OCL).

VI. Conclusion and Future Work

This paper proposed the extension of the SMarty approach for encompassing UML activity diagram models by

providing both new extended metaclasses support to the SMartyProfile and new guidelines to the SMartyProcess. An example of the extended SMarty approach was applied to two main features of the AGM PL: save game and load game.

The existing literature points out interesting related works, however none of them direct guided the SMarty extension strategy. Each of such works contributed in a specific basis, however several adjustments were done in order to make the extension as general as possible.

The extended SMartyProfile allows the application of its stereotypes to three metaclasses: DecisionNode, Action, and ActivityPartition. Variation points can be identified based on the proposed guideline G.8, which suggests the application of the stereotype ≪variationPoint≫ to DecisionNodes. Actions can be defined as variants according to guidelines G.9 and G.10, based on the concrete stereotypes ≪mandatory≫, ≪optional≫, ≪alternative_OR≫, and ≪alternative_XOR≫. Guideline G.11 suggests the application of the stereotype ≪variable≫ to ActivityPartitions that contain any type of internal variation.

This paper is limited with relation to formal definitions of variabilities in activity diagrams. Instructions of OCL or another formal language must support the SMartyProfile in order to guarantee the correct generation of PL products. Moreover, PL activity diagrams with more details must be used for validating the effectiveness of the extended SMarty.

Directions for future works are as follows: (i) apply OCL to SMartyProfile elements bringing out formalism to SMarty; (ii) analyze whether is possible to represent variability in pins and low granularity activity diagram elements; (iii) perform an experimental study for verifying the effectiveness of the extended SMarty; and (iv) investigate variability management in sequence diagrams as they are extremely rich in details and may contain essential variabilities for a PL.

REFERENCES

[1] J. Bosch, "Preface," in *Proc. Workshop on Software Variability Management: Software Product Families and Populations*. Groningen, The Netherlands: University of Groningen, 2004, pp. 1–2.

[2] A. Bragança and R. J. Machado, "Extending UML 2.0 Metamodel for Complementary Usages of the ≪extend≫ Relationship within Use Case Variability Specification," in *Proc. Software Product Lines Conf.* Washington, DC, USA: IEEE Computer Society, 2006, pp. 123–130.

[3] L. Chen, M. A. Babar, and N. Ali, "Variability Management in Software Product Lines: a Systematic Review," in *Proc. Software Product Lines Conf.* Pittsburgh, PA, USA: Carnegie Mellon University, 2009, pp. 81–90.

[4] H. Gomaa, *Designing Software Product Lines with UML: from Use Cases to Pattern-based Software Architectures*. Boston, MA, USA: Addison-Wesley, 2005.

[5] A. Heuer, C. J. Budnik, S. Konrad, K. Lauenroth, and K. Pohl, "Formal Definition of Syntax and Semantics for Documenting Variability in Activity Diagrams," in *Proc. Int. Conf. Software Product Lines*. Berlin, Heidelberg: Springer-Verlag, 2010, pp. 62–76.

[6] B. Korherr and B. List, "A UML 2 Profile for Variability Models and their Dependency to Business Processes," in *Proc. Int. Conf. Database and Expert Systems Applications*. Washington, DC, USA: IEEE Computer Society, 2007, pp. 829–834.

[7] F. J. v. d. Linden, K. Schmid, and E. Rommes, *Software Product Lines in Action: The Best Industrial Practice in Product Line Engineering*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2007.

[8] E. A. Oliveira Junior, I. M. S. Gimenes, and J. C. Maldonado, "Systematic Management of Variability in UML-based Software Poduct Lines," *Journal of Universal Computer Science*, vol. 16, no. 17, pp. 2374–2393, 2010.

[9] OMG, "UML Specification - Superstructure v2.2," 2009. [Online]. Available: http://www.omg.org/cgi-bin/doc?formal/09-02-02

[10] D. Peixoto, V. Batista, A. Atayde, E. Borges, R. Resende, and C. Pádua, "A Comparison of BPMN and UML 2.0 Activity Diagrams," in *Proc. Simposio Brasileiro de Qualidade de Software*, Florianopolis-SC, Brazil, 2008, pp. 1–14.

[11] K. Pohl, G. Böckle, and F. J. v. d. Linden, *Software Product Line Engineering: Foundations, Principles, and Techniques*. Secaucus, NJ, USA: Springer-Verlag, 2005.

[12] A. Reuys, S. Reis, E. Kamsties, and K. Pohl, "Derivation of Domain Test Scenarios from Activity Diagrams," in *Proc. Int. Workshop on Product Line Engineering*, 2003, pp. 1–7.

[13] S. H. Ripon, K. H. Talukder, and M. K. I. Molla, "Modeling Variability for System Families," *Malaysian Journal of Computer Science*, vol. 16, no. 1, pp. 37–46, 2003.

[14] S. Robak, B. Franczyk, and K. Poltitowicz, "Extending the UML for Modeling Variability for System Families," *Int. Journal of Applied Mathematics and Computer Science*, vol. 12, no. 2, pp. 285–298, 2002.

[15] A. Schnieders, "Variability Mechanism Centric Process Family Architectures," in *Proc. IEEE Int. Symposium and Workshop on Engineering of Computer Based Systems*. Washington, DC, USA: IEEE Computer Society, 2006, pp. 289–298.

[16] SEI, "Arcade Game Maker Pedagogical Product Line," 2010. [Online]. Available: http://www.sei.cmu.edu/productlines/ppl

[17] M. Svahnberg, J. v. Gurp, and J. Bosch, "A Taxonomy of Variability Realization Techniques: Research Articles," *Software Practice and Experience*, vol. 35, no. 8, pp. 705–754, 2005.

[18] T. Ziadi, L. Hélouët, and J. M. Jézéquel, "Towards a UML Profile for Software Product Lines," in *Proc. Product Family Engineering*, 2003, pp. 129–139.

Improving the identification of traceability links between source code and requirements

Daniel Hanspeter, Andrea Janes, Alberto Sillitti, Giancarlo Succi
Free University of Bozen/Bolzano
Center for Applied Software Engineering
Bolzano, Italy
daniel.hanspeter@stud-inf.unibz.it, {andrea.janes, alberto.sillitti, giancarlo.succi}@unibz.it

***Abstract*—**Software developers are interested in requirement traceability to e.g., verify if all requirements are covered by a system design specification. Based on the assumption that related artifacts contain related terms, researchers have developed, used, and extended algorithms that identify related terms and subsequently infer which artifacts are relat ed (i.e., there is a traceability link between them). Source code is not as verbose as a natural language description, which reduces the applicability of algorithms that precisely rely on such a commonality. This paper extends the Vector Space Model using tf*idf term weights to improve the identification of traceability links between source code and requirements. To this extent, we modify the way how requirements are identified and to include user feedback. We show that the inclusion of user feedback significantly improved the number of correctly identified requirements.

Keywords-requirement traceability; information-retrieval; recommender systems

I. INTRODUCTION

Requirement traceability is the capacity of "describing and following the life of a requirement, in both a forwards and backwards direction, i.e., from its origins, through its development and specification, to its subsequent deployment and use, and through periods of ongoing refinement and iteration in any of these phases" [1].

Software developers are interested in requirement traceability to e.g., verify if all requirements are covered by a system design specification, if and how software satisfies some given requirements [2], to understand which parts of the source code have to be refactored because a requirement changed [3][4], if all requirements are tested, to document [5][6][7][8] and predict the effort [9][10] and so on.

Traceability links – relations between a certain artifact in one document, like a requirement, with another artifact in another document, like a class – can be obtained using manual or automatic means. A manual assignment of code to a requirement is based on previously acquired knowledge or on the abilities to deduce the requirement behind a piece of code.

Automatic approaches – developed with the goal to reduce costs and to be able to tackle the complexity of identifying traceability links in large projects [11] – attempt to extract the links from the software development process itself, i.e., tracking its inputs, outputs, and activities and deducing from that why a given code was created.

Based on the assumption that related artifacts contain related terms, researchers have developed, used, and extended algorithms that identify related terms and subsequently infer which artifacts are related (e.g., [12][13][14][15][16][17][18][19][20][21][22][23][24][25]).

Information-retrieval (IR) techniques model the task of traceability link identification as an IR problem. The text contained in an artifact is used to formulate a query that retrieves artifacts that match the query. The retrieved artifacts are candidate traceability links. The presence of terms in code that are also contained in a requirement is interpreted as the trace to the requirement.

The outcome of such algorithms depends on the use of similar terms for similar things (i.e., a project glossary) and on complete and concise descriptions [12][26]. A complete description is exhaustive. If we describe every requirement exhaustively, we increase the probability that two related artifacts contain similar terms. A concise description contains only aspects that belong to it. This minimizes the probability that the same terms are present in two unrelated artifacts.

Algorithms that identify traceability links between natural language artifacts are less accurate when applied between formal language (like source code) and natural language artifacts [12][27]. Source code is not as verbose as a natural language description. For example, redundancies are avoided to improve maintainability, while in natural language redundancy can be used to ensure that something is understood correctly. Similarly, ambiguities in natural language is a form to communicate e.g., irony. Formal or controlled natural languages have explicit constraints on grammar, lexicon, and style to reduce ambiguity and complexity [28]. These differences reduce the applicability of algorithms that precisely rely on such a commonality [12].

This paper proposes to extend an IR algorithm to retrieve requirements based on source code and to include user feedback to improve the identification of traceability links between source code and requirements. The modifications are inspired from recommender systems in which similarity is used to improve recommendations and user feedback is used to personalize the recommendation [29].

The paper is structured as follows: section II describes related works in the field of IR based automatic traceability, section III our approach, section IV the evaluation. Section V concludes with the discussion of the results.

II. Related Works

We performed a systematic literature review [30] with the following research question: "which automatic ways, that make use of IR techniques, are described in literature to relate source code to requirements?".

We searched the terms "automated", "requirement", "traceability", "information-retrieval" in the title, keywords, and abstract in the online literature databases ACM Digital Library, IEEE Xplore, ScienceDirect, and Springer.

We included studies that a) report formal and informal approaches to requirement traceability in source code, b) contain a case study, a survey, an experiment, or a literature review, and c) were published in the time frame of January 2000 to February 2012. We excluded all studies that a) did not explicitly define what traceability means in the specific study, b) did not analyze traceability between source code and requirements, and c) did not describe the applied algorithm.

The result of the study selection process was a set of 30 primary studies, categorized by the technique used to identify traceability links: manually [31][32][33][34], automatic adopting IR algorithms [12][13][14][15][16], automatic extending IR algorithms [17][18][19][20][21][22][23][24][25][35], and not IR based [36][37][38][39][40][41].

For the purpose of this paper, we focused on automatic approaches, based on IR algorithms. In the following we briefly discuss the most popular approaches identified in our review.

The Vector Space Model (VSM) [42] is an approach that represents documents and search queries as vectors. The similarity between queries and documents is calculated by producing a similarity coefficient for each document's vector compared with the query's vector. Documents having a higher coefficient are considered the most pertinent.

The VSM uses basic term frequency, causing terms and derivatives (i.e., work and work*ed*) to be treated as separate entities. This produces poor similarity values even for those documents that contain derivatives and synonyms of terms contained in the search query. This is one of the major challenges of IR methods [43].

The VSM indexes content terms in documents (indexing) and assign of an importance level to each term (weighting) [44]. Term frequency – Inverse document frequency (TF*IDF) [45] and Latent semantic indexing (LSI) [46] represent the most used and common term weighting methods.

IDF (Inverse document frequency) is one of the most common term weighting methods, based on the assumption that the importance of a term associated to a document is inversely proportional to the frequency of occurrences of this term in all documents. IDF evolved into TF*IDF on the heuristic intuition that a term occurring in many documents is a less good indicator and should be weighted less than a term occurring in few documents. In other words, TF*IDF calculates the relevance of a particular word in a specific document by taking into consideration the frequency of this term in the global document collection, meaning that words that can be located in a few number of documents will be weighted more because of their rarity.

IR approaches suffer from the fact that queries often contain words that are not exactly those ones that are used for indexing and describing the resources they are searching for [43]. In fact, synonymy and polysemy of words in languages make it possible for users to create, produce and describe equal works using different terms, or different works containing equal terms. This implies that a query may contain terms that produce results that skip relevant documents, while keeping others that are not actually part of the search, just because of the term presence and not by the term's contextual meaning.

LSI (Latent semantic indexing) analyzes connections and occurrences of words in documents. The mathematical technique behind LSI is SVD (Singular Value Decomposition) which is used on a weighted term-document matrix to identify concepts contained in the analyzed documents.

The idea behind LSI is the usage of additional knowledge provided by the consideration of the semantic structure of documents and their content, based on the principle that words appearing in the same context incline to have similar meanings. In this way, also conceptually similar terms to those contained in search queries will be considered in the retrieval of results.

The challenge of applying LSI is the determination of the optimal number of dimensions for SVD. The selection of fewer dimensions allows a more expanded comparison of concepts in the documents, while a higher number enables a more detailed comparison.

TF*IDF and LSI assume that a term's frequency and representation inside a document reflects the possibility of that term to be an "agent" of this document. In fact, the more specific a term is to a topic the more straightforward a topic can distinguish relevant documents [42].

The studies we identified show promising results compared to a manual traceability link identification [7], but unfortunately suffer from low precision and recall and must be integrated with additional methods [8] to improve the results.

Two main strands of thought arise from the collected studies: a) one approach is to extend and adapt the base algorithms (as in [11], [12], and [13]) and b) by combining IR methods with other fields, like text mining, digital libraries [14], or clustering [9].

To our knowledge, the field of recommender systems remains still unconsidered in the field of traceability link identification. In this paper we combine IR concepts with a method used in collaborative filtering: user feedback [29]. In this work we understand feedback as a strategy for learning new search terms and improving the algorithm rather than evaluating the performance of the algorithm itself, as proposed by [47][48].

III. Approach

We divided the task of combining an IR algorithm with user feedback to improve the identification of traceability links between source code and requirements in four steps: 1) identify the most suited IR algorithm that identifies traceability links between source code and requirements in our context, 2) implement a basic version of the identified algorithm, 3) extend the algorithm to improve the precision for source code similari-

ty, and 4) consider user feedback.

A. Selection of the IR algorithm

From the systematic literature review we found out that the VSM in combination with TF*IDF or LSI is the most popular algorithm. In a first experiment, we determined which of the two term weighting methods are more suitable for us to identify requirement to source code traceability links.

We used 20 source code fragments, applied both algorithms – VSM/TF*IDF and VSM/LSI – to identify the originating requirement and compared the result with the (manually identified) correct requirement. To evaluate if the average number of correctly recognized requirements differ between VSM/TF*IDF and VSM/LSI, we performed a Mann–Whitney–Wilcoxon paired test [49]. We performed a one-tailed test with H_0: VSM/TF*IDF and VSM/LSI identify the correct requirement equally well and H_1: VSM/TF*IDF identifies more identify traceability links than LSI.

With a p-value of 0.03764 and Wilcoxon's test statistic W = 62 we rejected H_0 and confirmed that – in our context – VSM/TF*IDF performs significantly better (i.e., identifies more requirements correctly) than VSM/LSI. This confirms that VSM/TF*IDF is reported to have a higher prediction accuracy than VSM/LSI in contexts that are not characterized by the use of synonyms [50], such as in our context: short requirements containing few characterizing keywords. This makes it necessary to value the presence of rare terms in documents as VSM/TF*IDF does, and not their relation with other terms done by VSM/LSI.

B. Basic implementation of VSM/TF*IDF

We represent traceability links in a requirement traceability matrix (RTM) [51]. Such matrix relates two sets of artifacts and contains the type of relationship in the intersecting cell.

We implemented the VSM/TF*IDF algorithm within an Eclipse plugin. The plugin downloads the requirements from an issue management system and the source code from an configuration management system. The terms are obtained tokenizing the text representing the requirements and the source code after removing stop-words, boundaries, abbreviations, and phrases [52] and after reducing words to their stem form.

The plugin populates the RTM on startup, marking which requirement document r and which source code document d contains a specific term t. Tab. 1 shows an example of a RTM in which the term t_1 is contained in the source code file d_2 and the requirement r_2. This is marked by a dot in the intersecting cells of the row for the term t_1 and the column for the source code file d_2 and in the column for the requirement r_2.

When the programmer modifies source code, the terms contained in the method surrounding the current cursor position are extracted and become the query vector used to identify similar requirements.

As said in the introduction, source code is not as verbose as a natural language description. Consequently, methods – that are usually short units of source code – contain only few terms that can be used to search for related requirements, which renders this approach unfeasible in practice.

C. Extending the query vector to consider related source code for the similarity calculation

To reduce the problem of having too few terms to identify the related requirements, we extend the query vector to include more terms than only those of the modified method. This allows to predict requirements not only by their key identifying terms in the current modified method, but also by using those terms to identify similar requirements linked to methods having similar content.

Assuming that the modified source code fragment f is part of a document d_m and contains the terms t_q, for every term t in the RTM that is not yet part of t_q, we verify if f has a similarity of $sim(d_x, f)$ to another source code document d_x, $x \neq m$, higher than a defined threshold sim_{min} (we chose a threshold of 0.30 in our experiments. The optimal value of this threshold has to be evaluated in future investigations). If this is the case, we add the missing term t to f with a term weight of $sim(d_x, f) \times tf*idf(t, d_x)$ if and only if t is part of a requirement. $sim(d_x, f)$ is the VSM/TF*IDF similarity between d_x and f, $tf*idf(t, d_x)$ is the tf*idf weight of the term t in the document d_x.

This approach adds terms to the query vector that are contained in similar source code artifacts but with a weight that models also the similarity of the source code fragment f to d_x.

D. Adding user feedback

We use user feedback to improve the similarity calculation between source code and the requirements. The user can use the Eclipse plugin to a) see the predicted requirement and b) to change it.

When the user chooses a requirement r for source code fragment f, the plugin adds the terms contained in f to the terms contained in requirement r. When the similarity between the source code fragment and requirement r is calculated now, it is higher, since more similar words exist in both texts. This adding of terms and the requirement enhancing process described here allows the system to adapt and evolve predictions for the currently observed project's code and to specialize itself for the specific project's content.

While a user suggested requirement r for source code fragment f improves the similarity of both, it also improves similarity of r with other fragments f_x containing terms contained in f. Repetitive and reserved terms in code (e.g. "Session", "Request", "Response", etc.) are treated by the system as additional stop words allowing rare terms to be considered as more descriptive.

IV. Evaluation

Table 1. Initial requirement traceability matrix

Terms	Documents							
	Source code				Requirements			
	d_1	d_2	...	d_n	r_1	r_2	...	r_n
t_1		●				●		
t_2	●				●			
t_3		●				●		
...								
t_n								

The source code and the requirements used in our experiments are taken from projects developed using an agile approach, in which requirements are formulated in a short, concise way [53] and source code is (mainly) undocumented.

Source code without comments is criticized by practitioners and academics alike [54][55]. Nevertheless, promoters of agile software development [56] often develop code without or with only few comments if they think that it can be avoided [57] or compensated with "test-first" [58].

To develop "test-first" means to write the test case before the requirement is implemented. The test case (obviously) fails, but it is that failing test case that permits the programmer to write the code necessary to fulfill the test case.

The test case represents an acceptance test, written in a programming instead of natural language. It documents the intent of the tested code part. Since only code is written that is necessary to fulfill a failing test case, this practice also results in high testing coverage [59][60].

This chapter evaluates the proposed algorithm using the template proposed by [61]. The template provides a guideline to define the experiment's context, the observed hypothesis, the preparation, and execution of it.

The objective of this experiment was the evaluation of our automated traceability algorithm to observe how well source code artifacts are traced back to their originating requirement.

We tested our prototype with five different sets of source code and requirements. The first set (20 methods and 13 requirements) is from an Android application developed during a bachelor thesis, the second set (30 methods and 18 requirements) is from a Java project developed during a research project, the third set (30 methods and 13 requirements) is from Freemind[1], the fourth set (30 methods and 12 requirements) is from JGraphT[2], and the fifth set (30 methods and 13 requirements) is from JavaNCSS[3]. The last three projects are all Open Source Java software.

We performed the experiment using an automated test since it was possible to model the user behavior [62] and this allowed us to avoid errors caused by wrong artifact analysis or incorrect user feedback. The user's cursor change in code can be emulated by providing the algorithm any time with a different artifact representing source code snippets.

To emulate correct user feedback, artifact files where named respectively with the correct requirement number, which was then used by the testing algorithm to simulate user feedback.

We were interested to understand how the inclusion of similar source code documents to modify the search query and correct user feedback influences the tracing of source code artifacts back to their originating requirement. We formulated the null and alternative hypotheses as follows:

- $H_{0,1}$: The adaptation of the similarity calculation does not influence automated traceability of requirements in source code artifacts.
- $H_{0,2}$: The use of user feedback does not influence automated traceability of requirements in source code artifacts.
- $H_{a,1}$: The adaptation of the similarity calculation influences automated traceability of requirements in source code artifacts.
- $H_{a,2}$: The use of user feedback influences automated traceability of requirements in source code artifacts.

The preparation of the experiment involved uploading the requirements to the issue management system and adding the source code to the configuration management system. The source code files had to be named to contain the correct requirement identification.

The evaluation consisted in the comparison of results achieved using a) the use of pure VSM/TF*IDF and b) the modified VSM/TF*IDF including the correct user feedback for each artifact.

The null hypotheses has been tested using a paired, nonparametric test, the Mann–Whitney–Wilcoxon paired test [49] as suggested by [63]. With a significance level of 95%, the null Hypothesis $H_{0,1}$ was accepted, the null Hypothesis $H_{0,2}$ rejected. $H_{0,2}$ was rejected in the first project with a Wilcoxon's test statistic of W = 42, in the second project with W=63,5, in the third with W=76, in the forth with W=34 and in the fifth with W=62.5.

The results gathered for the projects show that user feedback has a significantly higher average correct requirement prediction.

V. Discussion

The experiment described in this paper showed that in our projects the adapted algorithm could not significantly improve the identification of traceability links. The consideration of user feedback on the other hand significantly increased the correctly identified the trace to the requirements, in the first, third fourth and fifth project the improvement is higher (see Figure 1 for project 1 as example) than in the second project (see Figure 2).

The reason for that difference is the one mentioned in the introduction: requirement traceability using IR based methods

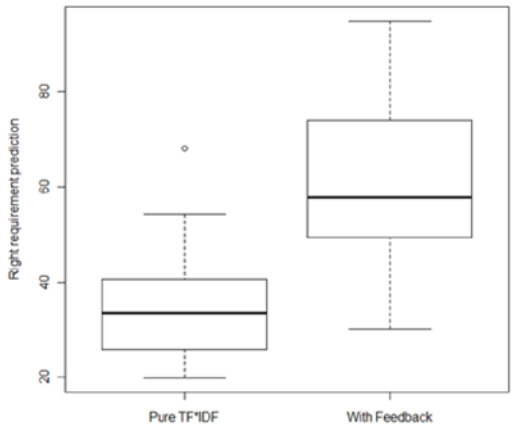


Figure 1. Correctly identified requirements in project 1.

[1] http://freemind.sourceforge.net
[2] http://www.jgrapht.org
[3] http://www.kclee.de/clemens/java/javancss

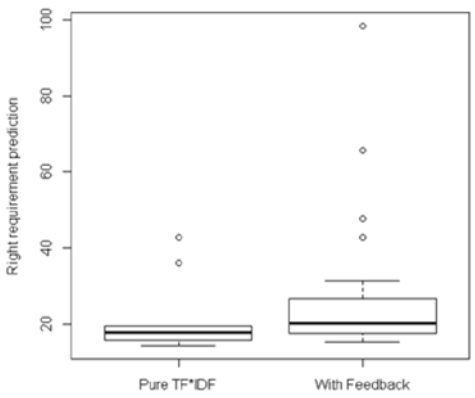


Figure 2. Correctly identified requirements in project 2

relies on the premise that programmers use meaningful (i.e., related, abbreviated, synonyms) names for types, classes, variables, functions and methods [12].

In project one three four and five, the artifacts made use of words for variables and methods that can be also found in the requirements. This allows TF*IDF to identify requirements in code already in its basic version. In project two we notice that requirement related words appear very rarely and are often distributed over similar and in some cases cross referenced [64] (i.e. one class makes use of methods of another class that are named using terms identifying requirements) coded artifacts. This is why most of the artifacts are recognized initially by the TF*IDF algorithm as being part of the same requirement. Feedback specializes the code for each requirement and in this way helps to identify also very similar methods and code parts as distinct requirements.

In the introduction we discussed the importance of automated traceability link identification. In this work we show that feedback can significantly improve the prediction, but we investigated if we can decrease the amount of feedback we ask for. We repeated the experiment only asking every 2nd time for feedback. Then every 3rd, 4th, and 5th. The results are shown in table 2. The precision [65] decreases rapidly from giving feedback every time to every second time. From giving feedback every second, third, fourth, or fifth time, the differences – while still decreasing – are not so big anymore.

As seen in the review of the state of the art, crosscutting requirements and non-meaningful named code need special adaptions of algorithms to allow their traceability. Crosscutting requirements define concepts that reflected in code terms mean particular functions and methods that can spread over all components. Common examples include transaction and error handling, security and logging. For example, a requirement saying "Every user interaction has to be logged" implies that every method capturing a user input to the system has to include one or more lines recording an entry in a log file.

Table 2. Precision and recall values with decreasing feedback

	1	2	3	4	5
Precision	75%	56%	56%	53%	51%
Recall	66%	74%	79%	84%	90%

This illustrates the difficulty of recognizing those kind of requirements majorly defined by only one line replicated all over the system. Future works will also address this issue.

REFERENCES

[1] O. Gotel and A. Finkelstein, "An analysis of the requirements traceability problem," RE, pp. 96-101, Colorado Springs, Colorado, 1994.

[2] H.G. Gross, M. Melideo, and A. Sillitti, "Self Certification and Trust in Component Procurement," Journal of Sc ience of C omputer Programming, Elsevier, Vol. 56, pp. 141-156, April 2005.

[3] I. Coman and A. Sillitti, "An Empirical Exploratory Study on Inferring Developers' Activities from Low-Level Data," SEKE, Boston, MA, USA, 9 - 11 July 2007.

[4] R. Moser, P. Abrahamsson, W. Pedrycz, A. Sillitti, and G. Succi, "A case study on the impact of refactoring on quality and productivity in an agile team," 2nd IFIP Central and East European Conference on Software Engineering Techniques, Poznań, Poland, 2007.

[5] A. Jermakovics, M. Scotto, A. Sillitti, and G. Succi, "Lagrein: Visualizing User Requirements and Development Effort," ICPC, Banff, Alberta, Canada, 26 - 29 June 2007.

[6] A. Sillitti, G. Succi, J. Vlasenko "Toward a Better Understanding of Tool Usage (NIER Track)," ICSE, Honolulu, HI, USA, May 2011

[7] I. Fronza, A. Sillitti, G. Succi, J. Vlasenko. "Understanding how Novices are Integrated in a Tea m Analysing their Tool Us age," Proceedings of the 2011 ICSP at ICSE 2011, Honolulu, HI, USA, May 2011

[8] L. Corral, A. Sillitti, G. Succi, J. Strumpflohner, J. Vlasenko. "DroidSense: A Mobile Tool to A nalyze Software Development Processes by Measuring Team Proximity," TOOLS Europe, Objects, Models, Components, Patterns, pp.17-33, 2012

[9] P. Abrahamsson, R. Moser, W. Pedrycz, A. Sillitti, and G. Succi, "Effort Prediction in Iterative Software Development Processes – Incremental Versus Global Prediction Models," 1st Int. Symposium on Empirical Software Engineering and Measurement, Madrid, Spain, 20 - 21 September 2007.

[10] P. Abrahamsson, I. Fronza, R. Moser, J. Vlasenko, W. Pedrycz. "Predicting Development Effort from User Stories," ESEM, 2011

[11] B. Ramesh, C. Stubbs, T. Powers, and M. Edwards, "Lessons Learned from Implementing Requirements Traceability," The Journal of Defense Software Engineering, pp. 11-16, April 1995.

[12] G. Antoniol, G. Canfora, G. Casazza, A. De Lucia, and E. Merlo, "Recovering Traceability Links between Code and Documentation," Transactions on Software Engineering, 28(10):970-983, 2002.

[13] J.H. Hayes, A. Dekhtyar, and S.K. Sundaram, "Improving after-the-fact tracing and mapping: supporting software quality predictions," IEEE Software, 22(6):30-37, 2005.

[14] A. De Lucia, F. Fasano, R. Oliveto, and G. Tortora, "Recovering traceability links in software artifact management systems using information retrieval methods," ACM Transactions on Soft ware Engineering and Methodology, 16(4):13-63, 2007.

[15] H.G. Gross and M. L ormans, "Towards Software Component Procurement Automation with Latent Semantic Analysis," Proceedings of the Third Int. Workshop on Coordination and Adaption Techniques for Software Entities, pp. 51-68, 2007.

[16] A. De Lucia, R. Oliveto, and G. Tortora, "Assessing IR-based traceability recovery tools through controlled experiments," Empirical Software Engineering, 14(1):57-92, 2009.

[17] J. Cleland-Huang, R. Settimi, D. Chuan, and Z. Xuchang, "Utilizing supporting evidence to improve dynamic requirements traceability" RE, pp. 135-144, La Sorbonne, France, 2005.

[18] D. Chuan and J. Cleland-Huang, "Clustering support for automated tracing," ASE, pp. 244-253, Atlanta, Georgia, USA, 2007.

[19] J. Cleland-Huang, W. Marrero, and B. Berenbach, "Goal-Centric Traceability: Using Virtual Plumblines to Maintain Critical Systemic Qualities," IEEE Trans. Softw. Eng., 34(5):685-699, 2008.

[20] Hsin-Yi Jiang, T.N. Nguyen, Ing-Xiang Chen, H. Jaygarl, and C.K. Chang, "Incremental Latent Semantic Indexing for Automatic

Traceability Link Evolution Management," ASE, pp. 59-68, L'Aquila, Italy, 2008.

[21] X. Wang, G. Lai, and C. Liu, "Recovering Relationships between Documentation and Source Code based on the Characteristics of Software Engineering," Proceedings of t he 2nd Int. Workshop on Harnessing Theories for Tool Support in Software, pp. 121-137, 2009.

[22] S. Winkler, "Trace retrieval for evolving artifacts," Proceedings of the 2009 ICSE Workshop on T raceability in Emerging Forms of Soft ware Engineering, pp. 49-56, Vancouver, Canada, 2009.

[23] X. Chen, "Extraction and visualization of traceability relationships between documents and source code," ASE, pp. 505-509, Antwerp, Belgium, 2010.

[24] C. Wiederseiner, V. Garousi, and M. Smith, „Tool Support for Automated Traceability of Test/Code Artifacts in Embedded Software Systems," IEEE 10th Int. Conference on Trust, Security and Privacy in Computing and Communications, pp. 1109-1117, Changsha, P.R. China, 2011.

[25] B. Dit, M. Revelle, and D. Poshyvanyk, "Integrating information retrieval, execution and link analysis algorithms to improve feature location in software," Empirical Software Engineering, Online First™, January 2012.

[26] J. Cleland-Huang, R. Settimi, and E. Romanova, "Best Practices for Automated Traceability," Transactions on Software Engineering, 60(6):27-35, 2007.

[27] A. Marcus, J. I Maletic, and A. Sergeyev, "Recovery of traceability links between software documentation and source code," Int. Journal of Software Eng. and Knowledge Eng., 15(5):811-836, 2005.

[28] R. I. Kittredge, "Sublanguages and controlled languages," Oxford University Press, 2003.

[29] F. Ricci, L. Rokach, B. Shapira, and P.B. Kantor. Reccomender Systems Handbook. 1st Edition, Springer-Verlag, 2011.

[30] B. Kitchenham, "Procedures for p erforming systematic reviews," Technical report, Keele University and NICTA, 2004.

[31] R. Pooley and C. Warren, "Reuse through Requirements Traceabilitym," The Third Int. Conference on Software Engineering Advances, pp. 65-70, Sliema, Malta, 2008.

[32] E. Denney and B. Fisher, "A Verification-Driven Approach to Traceability and Documentation for Auto-Generated Mathematical Software," 24th IEEE/ACM Int. Conference on Software Engineering, pp. 560-564, Auckland, New Zealand, 2009.

[33] P. Lago, E. Niemela, and H. Van Vliet, "Tool support for traceable product evolution," Eighth European Conference on Software Maintenance and Reengineering, pp. 261-269, 2004.

[34] S. Klock, M. Gethers, B. Dit, and D. Poshyvanyk, "Traceclipse: an eclipse plug-in for traceability link recovery and m anagement," Proceedings of the 6th Int. Workshop on Traceability in Emerging Forms of Software Engineering, pp. 24-30, Waikiki, Honolulu, HI, USA, 2011.

[35] Jun Lin, Chan Chou Lin, J.C. Huang, and R. Settimi, "Poirot: A Distributed Tool Supporting Enterprise-Wide Automated Traceability," 14th IEEE Int. Conference Requirements Engineering, pp. 363-364, Minneapolis, Minnesota, USA, 2006.

[36] A. Olszak and B.N. Jørgensen, "Remodularizing Java programs for improved locality of feature implementations in source code," Feature-Oriented Software Development, 77(3):131-151. 2012.

[37] B.S. Ali and Z.M. Kasirun, "3CI: A Tool for C rosscutting Concern Identification," Int. Conference on Computational Intelligence for Modelling Control & Automation, pp. 351-355, Vienna, Austria, 2008.

[38] G. Fraser and P . Ammann, "Reachability and Propagation for LTL Requirements Testing," The Eighth Int. Conference on Quality Software, pp. 189-198, Oxford, UK, 2008.

[39] C. Ziftci and I. Krueger, "Tracing requirements to tests with high precision and recall," ASE, pp. 472-475, Lawrence, Kansas, USA, 2011.

[40] P. Mäder and O. Gotel, "Towards automated traceability maintenance," Journal of Systems and Software, 25 October 2011.

[41] I. Omoronyia, G. Sindre, and T. Stålhane, "Exploring a Bayesian and linear approach to requirements traceability," Information and Software Technology, 53 (8):851-871, 2011.

[42] G. Salton and C.S. Yang, "On the specification of term values in automatic indexing," Journal of Documentation, 29(4):351–372, 1973.

[43] M.W. Berry, S.T. Dumais, and G.W. O'Brien, "Using linear algebra for intelligent information retrieval," SIAM Review, 37(4):573–595, 1995.

[44] D.D. Lewis, "Text representation for intelligent text retrieval: A classification-oriented view," Text-based intelligent systems: current research and practice in information extraction and retrieval, Lawrence Erlbaum Associates, Inc., New Jersey, USA, 1992, pp. 179-197.

[45] K. Sparck Jones, "A statistical interpretation of term specificity and its application in retrieval," Journal of Documentation 28, pp. 11–21, 1972.

[46] S. Deerwester, S.T. Dumais, T.K. Landauer, G.W. Furnas, and R.A. Harshman, "Indexing by latent semantic analysis," Journal of American Society of Information Science, 41(6):391–407, 1990.

[47] J.H. Hayes, A. Dekhtyar, and S.K. Sundaram, "Advancing Candidate Link Generation for Requirements Tracing: The Study of Methods," IEEE Trans. Softw. Eng., 32 (1): 4–19, 2006.

[48] De Lucia, Andrea, Rocco Oliveto, and Paola Sgueglia. 2006. "Incremental Approach and User Feedbacks: a S ilver Bullet for Traceability Recovery." Proceedings of t he 22nd IEEE ICSM, Washington, USA, pp. 299–309.

[49] W.J. Conover, Practical Nonparametric Statistics, 3rd Edition, Wiley, 1998.

[50] S. Dumais, "Latent Semantic Analysis," Review of Information Science and Technology, 38(4), 2004.

[51] A.M. Davis, "The analysis and specification of systems and software requirements," IEEE Systems and Software Requirements Engineering, pp. 119-144. 1990

[52] R. Grishman, "Natural language processing," Journal of the American Society for Information Science, 35(5):291–296, 1984.

[53] A. Sillitti and G. Succi, "Requirements Engineering for Agile Methods," in Engineering and Managing Software Requirements (Eds.: Aurum A., Wohlin C.), Springer, 2005.

[54] J. Raskin, "Comments are More Important than Code," Queue –Patching and deployment, 3(2):61-63, 2005.

[55] D. E. Knuth, "Literate Programming," The Computer Journal 27 (2):97-111, 1984.

[56] K. Beck, Extreme Programming Explained: Embrace Change (2nd Edition), Addison Wessley, 2004.

[57] M. Fowler, "CodeAsDocumentation," martinfowler.com, http://martinfowler.com/bliki/CodeAsDocumentation.html, 2005.

[58] James Shore and Shane Warden, The Art of Agile Development (First ed.). O'Reilly. 2007.

[59] L. Madeyski, "The impact of Test-First programming on branch coverage and mutation score indicator of unit tests: An experiment," Information and Software Technology 52(2):169–184, 2010.

[60] M. Siniaalto and P. Abrahamsson, "A Comparative Case Study on the Impact of Test-Driven Development on Program Design and Test Coverage," Empirical Software Engineering and Measurement, pp. 275-284, 2007.

[61] C. Wohlin, P. Runeson, M. Host, M. C. Ohlsson, B.Regnell, and A. Wesslen, Experimentation in Software Engineering – An Introduction, Kluwer, 2000.

[62] E. Dustin, J. Rashka, and J. Paul, Automated Software Testing: Introduction, Management and Performance, Addison Wesley, 1999.

[63] E. Skovlund and G. Fenstad, "Should we always choose a nonparametric test when comparing two apparently nonnormal distributions?," In Journal of Clinical Epidemiology 54, pp. 86–92, 2000.

[64] L. Rosenhainer, "Identifying Crosscutting Concerns in Requirements Specifications," Proceedings of OOPSLA Early Aspects: Aspect-Oriented Requirements Engineering and Architecture Design Workshop, Vancouver, Canada, October 2004.

[65] G. Salton. Automatic Information Organization and Retrieval. McGraw Hill Text, 1968.

An Integrated Architecture for Multiagent Virtual Worlds for Performing Adaptive Testing Games

Sandeep Virwaney Fuhua Lin Grant McClure Robert Heller
Athabasca University, Canada
oscarl@athabascau.ca

Abstract — **One of the key success factors that contribute towards the creation and sustenance of online (2D and 3D) virtual worlds for learning might be to provide game-style educational activities. However, there is no development platform available which can meet the inherent system requirements including usability of platform and scalability to modern massively multiplayer online games, yet focused on engaging learning for the individual user. Work has been done with software agents in the context of multi-agent systems (MAS), and it makes sense to try to leverage that work when it comes to modeling functional modules, controlling realistic non-player characters (NPCs), and Personal Assistants for Learning (PALs) in a virtual learning world. There are challenges to integrating a multi-agent system into a virtual world including concerns with synchronization, communication, monitoring, efficiency, and control. This paper describes the design and implementation of an integrated architecture for performing and facilitating quiz games for adaptive testing with a multi-agent system JADE/Jason and a 3D virtual world engine Open Wonderland.**

Keywords: virtual learning environments; multiagent systems

I. INTRODUCTION

Educational theories classify learning activities as constructivist, constructionist, or situated [14]. The three types of learning are well served by virtual worlds due to its ability to mediate world exploration and construction, to map a user to any character of choice, and to provide shared virtual three dimensional space. Through its interactive environment, repetition, and one-to-one experimentation, virtual worlds can help improve knowledge retention and student motivation. By simulating chemistry or physics experiments, knowledge can be transferred without increased safety risks associated with real laboratories. Through the Internet, virtual worlds can facilitate distance learning in places where schools or specialized teachers are scarce, and so on.

There have been relatively mature technologies for physical modeling for virtual objects, modeling virtual humans or non-player characters (NPCs) in virtual worlds in their appearance, gestures, kinematics, and physical properties. Avatars are controlled by the user during interaction. Like objects, virtual humans or NPCs are controlled by a simple script.

To build a smart virtual world, personalized and adaptive virtual learning environment, it is desirable to incorporate intelligence to virtual worlds. Work has been done with software agents in the context of multi-agent systems, and it makes sense to try to leverage that work when it comes to modeling functional modules for pedagogical purposes, controlling realistic non-player characters (NPCs), and maintaining Personal Assistants for Learning (PALs) in a virtual learning environment.

The research is interdisciplinary in nature. The domains involved include human-computer interaction, psychology, and education. The problem is that there are no agent development tools that are intended to be easy to use by non-programmers. To develop such tools we need to have a novel integrated system architecture that can incorporate Artificial Intelligence (AI) technologies such as machine learning, data mining, natural language processing, affective computing, and models and algorithms for building the agents.

A. MAS in Virtual Worlds

Multi-agent systems (MAS) have been adopted by researchers in virtual worlds. MAS is particularly well-suited to application domains where virtual entities (agents) are self-directed and can actively pursue their goals within an environment that they can interact with, including interactions with other entities that are also in pursuit of their own goals. It is easy to see how they are ideally suited for modeling people — they are active and social in a way similar to people.

There are two perspectives to consider when applying MAS into virtual worlds. One is from the artificial life (ALife) community while the other is from classical AI community. In the ALife community, the term complex dynamical systems is usually preferred, because it also includes physical inorganic systems, where the individual agents or components, such as molecules or sand grains, only have limited agent characteristics. ALife has strong roots in biology rather than psychology, and focuses on the *emergence* of behavior in large populations of agents [1].

In the AI community, an agent is assumed to have certain elementary sensory-motor abilities, so that it can perceive aspects of the environment and, depending on this information and its own state, perform certain behaviors. Rather than looking for emergence, as is common in the field of ALife, MAS is usually employed to solve some particular problems or achieve particular tasks, such as path-finding, coordinated control [2]. The individual agents are endowed with centralized control similar to that employed in the classical approach.

B. QuizMASter

In classrooms, teachers usually use quiz games to create some interesting activities. The purposes of quiz games for the classes include:

(1) reviewing and reinforcing previously taught material,

(2) warming up or ending lesson with a socially engaging activity,

(3) encapsulating the basic unit of conversation.

We have been designing and developing engaging pedagogical agents to enhance the engagement of students in a game-based learning environment QuizMASter [4-6], using a multi-agent system (MAS) approach. It helps students undergo adaptive testing and collaborative learning through friendly competition. Conceptually, QuizMASter is designed to be similar to a TV game show, where a small group of contestants compete by answering questions presented by the game show host. A prototype of QuizMASter using a JADE MAS was successfully implemented on Open Wonderland [7]. The implementation enabled an opportunity to explore the affective dimension of gaming situation by controlling the affective channels of information between avatars. Specifically, the QuizMASter agents can be enhanced to operate within the dimensions of Transformed Social Interaction (TSI) [8] to maintain high levels of engagement. Bailenson et al. (2008) proposed the theory of Transformed Social Interaction to describe the transformation of interaction in virtual environments. According to the theory, real-time transformations can be classified into three categories or dimensions: self-representation, sensory-abilities and situational context [8]. These transformations empower avatars to complement human perceptual abilities [12]. For the autonomous agent and the avatar it represents, the ability to process eye gaze or motion data is advantageous in any implementation of virtual world applications. Situational context deals with transformations that alter the spatial or temporal structure of a conversation. For example, the communication and positions between agents and students can be optimally configured in terms of the geographical setup of a classroom. A class of 20 students can sit directly in front of a virtual instructor, and perceive the rest of the students as sitting farther away. Furthermore, by altering the flow of rendered time in the communication session, users can implement strategic uses of rewind and fast forward during a real-time interaction to increase comprehension and efficiency.

This paper focuses on the roles of agents in multi-agent virtual learning environments and some architectural designs we have proposed and developed for multi-agent virtual learning environments. We investigate the effectiveness of the use of agents to facilitate virtual world based quiz games.

II. Related Work

Two different applications towards multi-agent virtual learning environments are surfacing: virtual classrooms and fantasy worlds. Virtual classrooms have the look and feel of a regular classroom and are often made to look like a replica of the sponsoring school. Second Life (SL) is often the choice for building the 3D virtual classrooms. The results are a mixed bag with some reporting that SL has too many difficulties and is not very game-like as such classrooms lack game scenarios. Fantasy worlds teleports the learner to a different place or time. For example, Virtual Singapura (VS) [15], an example of a fantasy world takes the learner to nineteen century Singapore in the throes of disease epidemic. The VS project is one of the first to integrate an intelligent agent architecture with a virtual world to explore ways in which adaptive synthetic characters might enhance the learner's experience in a virtual world and perhaps to enhance learning as well [15].

There has been some work done over the last decade when it comes to creating agent control mechanisms for 3D virtual environments, for example, Gamebots [9] and Pogamut [10]. Dingum et al. (2008) provide an excellent overview of the state of the art when it comes to incorporating MAS concepts with gaming systems or other 3D virtual environments [11]. According to the literature, it looks like most work to date has relied on the Unreal Tournament engine and Gamebots. There are a number of positives to this technology including the robustness of the environment. The downsides include the fact that it is not open source, and there are therefore limitations in what can be done in terms of ensuring agents have all the information they need and that synchronization works perfectly. There are also limitations in communications that means that agents aren't able to take full advantage of the social actions in a MAS environment.

III. The Roles of Agents

Agents are encapsulated computer systems that are designed to behave flexibly and somewhat autonomously to achieve some goal(s). Agents are situated in some environment and have some autonomy and capabilities to observe that environment. They can communicate those observations to other agents. This makes them particularly suited for distributed virtual learning environments. Such environments are open in nature as either students or teachers can enter or leave the virtual world at will.

They may be either visible or invisible. We classify the agents in a multi-agent virtual learning environment into the following types:

A. Game agents

First, the game agents are designed to realize the game functions. The game agents include game management agents, game control agents, user management agents, and domain knowledge management agents. Gaming frameworks are quite mature when it comes to optimizing graphics and physics—current games are visually stunning. There is more to a gaming experience than what a player sees. How a game is structured and the kind of narrative it plays also lends itself to the credibility of the virtual world a player interacts with. This includes the behavior of NPCs in the environment. If an NPC does not act in a believable way it detracts from the realism of the player's game experience.

B. Pedagogical agents.

Pedagogical agents are agents that could provide pedagogy to the games, including Instructor agents that work with the Host Agent to ensure the subject, difficulty, and content of the quiz are appropriate for each game; Behavior agents that during each game collects information, such as response time and score needed for PAL to update the student model.

C. PALs

Personal agents or Personal Assistants for Learning (PALs) are designed for analyzing the user's profile, learning attributes, and competencies keeping track of individual interests, preferences, motivations, and goals of the human participants (i.e., learners) and building student models.

D. NPCs

To make the game-based learning environment more realistic and immersive, a MAS is ideal for modeling virtual students and a virtual audience which are implemented as NPCs or bots. Many researchers in the virtual worlds community have noted that city replicas, malls, bars, or educational places often look interesting but abandoned at the same time. This socially isolating aspect could be mitigated by deploying MAS controlled NPCs who could inhabit virtual worlds to create impressions of the societies we live in. Also, they could be used to substitute the work of real people by fulfilling different tasks such as demonstrating objects, explaining manuals or guiding the way to new locations.

AI techniques have been used to drive NPCs, but this is usually done in a prescriptive manner that doesn't give a lot of freedom to NPC actions. These NPCs require simple scripting as the tasks they are following are simple procedures and need almost no interaction with the environment. Nevertheless, NPCs enhance the immersive feeling in a virtual world and provide answers to users that are seldom possible by other means. Moreover, we can increase the intelligence of the NPCs through implementing MAS *communication* and *coordination mechanisms* and further sensory feedback to actually enable interaction between human avatars and bots that are representing people.

Finally, the communication and coordination among these three types of agents would be extremely important to make the learning environment smart. To enable game-based learning, it is crucial for the pedagogical agents to make the quizzes interactive, interesting, appropriate in the degree of difficulty, and adaptive in topics given a group of students with different student models. Based on the user models, the game agents will also be able to make decisions intelligently, e.g. to optimize the grouping and seating/standing positions of participants according to their preferences.

IV. IMPLEMENTATION

The primary goal of this research is to construct an integrated intelligent virtual world for performing and facilitating quiz game-style adaptive testing. For a 3D VW, Open Wonderland [7] is used. For a MAS, a combination of JADE and Jason is used, leveraging the strengths of both where appropriate. As such, the framework needs to have interfaces to both. We have proposed an architecture, the WL-MAS interface, that is the foundation on which the educational game, QuizMASter, has been built. The WL-MAS interface are used by agents controlling NPCs as well as agents controlling other aspects of the game, such as score keeping and user modeling.

The following work has been completed to ascertain the technical feasibility and limitations creating agent-controlled NPCs in a 3D virtual world like Open Wonderland.

A. JADE-OWL Integration

As an initial step, we integrated Open Wonderland and JADE by starting a JADE server separate from Open Wonderland [12]. JADE (http://jade.tilab.com/) is a middleware that facilitates the development of multiagent system. JADE agents can be started inside the Open Wonderland world by simply using *jade.core.Runtime* and which enables the agent to send pertinent information back to OWL in the form of messages to a TCP port. The interface is a modified OWL module to which code has been added to start a JADE agent. The JADE agent is started via a runtime call. Therefore, the agent will be started on a local computer. However, the rest of the agents that make up the MAS can be on other computers. The agents started by the OWL module will communicate with the rest of the JADE agents using the FIPA specified protocol. One of the agents in the MAS communicates to a TCP port that has been activated on another or the same OWL module.

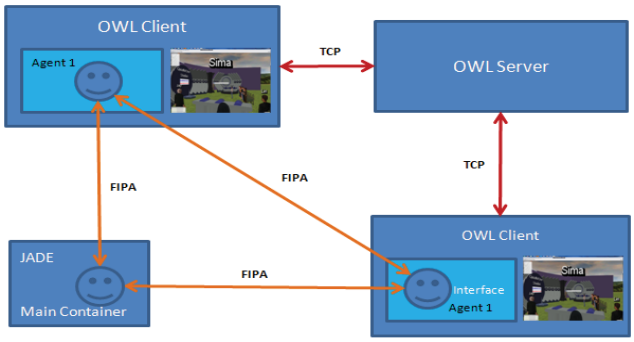


Figure 1: JADE-OWL integration.

JADE followed the FIPA assumption that only the external behavior of system components should be specified, while leaving the implementation details and internal architectures to agent developers. It implements a very general agent model that can be easily specialized to realize both reactive and Belief, Desires, and Intentions (BDI) architectures. Moreover, the behavior abstraction of our agent model allows simple integration of external software into one of the agent tasks.

B. Jason/JADE-OWL Integration

The results of our previous work showed the possibilities and challenges that arise when the systems aren't coupled enough [12]. These challenges informed our decision to simplify and integrate everything right into Open Wonderland. Therefore, our second model added a Jason module to the Open Wonderland server that takes an NPC and controls its movement in the environment (see Figure 2). Jason (http://jason.sourceforge.net/Jason/Jason.html) is an interpreter for an extended version of AgentSpeak and a platform for the development of multiagent systems.

The NPC runs on the Open Wonderland server and pulls an AgentSpeak() asl file from the file system (in the Open Wonderland server cache), sets up its belief system and goals, and then runs. As yet it doesn't get any perceptual information from the environment, but is able to guide the direction of the NPC. It will be easy to extend this to have the NPC make any available action that an avatar can do within Open Wonderland. The percept feedback to the Jason agent still needs to be completed. The MAS and the Open Wonderland will be more tightly coupled than what had originally been planned so that development, deployment, and management of QuizMASter

are easier when handled within a s ingle environment. The Open Wonderland module system makes it s imple to extend Open Wonderland to include the framework and any other functionality that is required [7]. Figure 3 shows the screen shot of Jason and Open Wonderland integration.

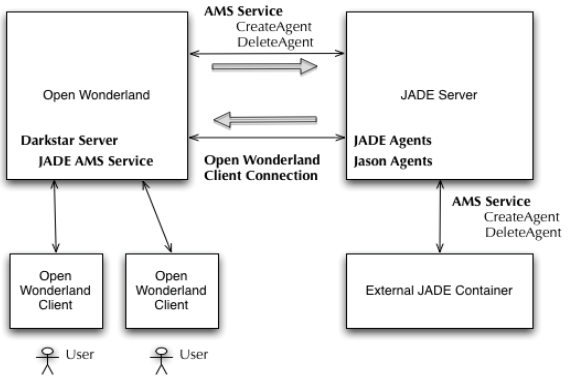


Figure 2: Jason/JADE-OWL integration.

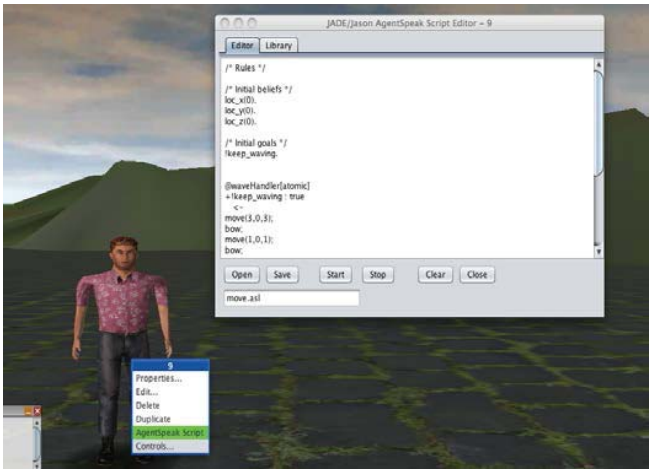


Figure 3: Jason and Open Wonderland integration.

C. Jason-CArtAgo-OWL Architecture

Our third model is CArtAgO (Common ARTifact infrastructure for AGents Open environments) [4]. CArtAgO is a general purpose framework that makes it possible to program and execute virtual environments for multi-agent systems. CArtAgO is b ased on the Agents & Artifacts (A&A) meta-model for modeling and designing MASs. Since these artifacts can be modeled in Java, some of the core components of the Open Wonderland. In the current implementation as a prototype, some agents are temporally implemented as artifacts. A MaryTTS server is added to the system to realize speech synthesis function with Open MARY TTS which is an open-source, multilingual Text-to-Speech Synthesis platform written in Java (http://mary.dfki.de/).

One of the main advantages of leveraging the capabilities of a Jason-CArtAgO MAS is that it enables customization of the rendering of visible artifacts for different clients. This advantage was crucial for implementing the three principles of TSI theory [8]. In particular, the principle of self-representation in NPC creation is described below.

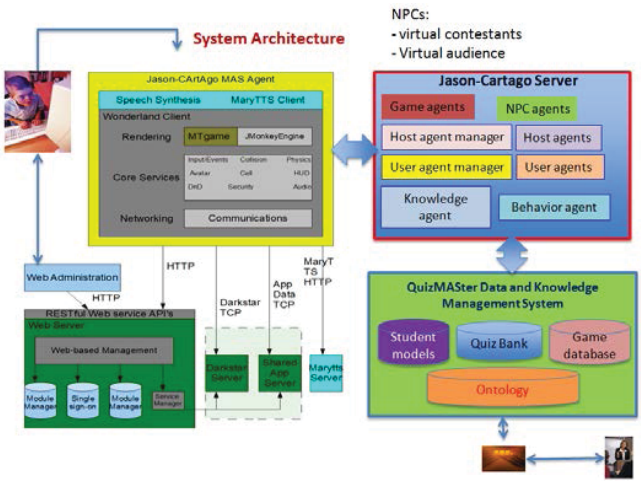


Figure 4: Jason-CArtAgo-OWL architecture.

1) NPC Creation and Facial Morphing

Evolver (http://evolver.com) was an online tool t hat enables one to model, texture, and rig one's own 3D character in minutes. It h as the capability to design the face of an NPC using a 2D photograph and further customize the body type, eyes, skin-complexion and wearable costumes of the avatar being developed. This capability of *Evolver* was used to create an NPC of one of the contestants, the host agent for player Oscar and player Steve as show in Figure 5 below. As shown by Bailenson et al. (2008), when facial images incorporate personal features, user satisfaction ratings increase.

- *Self-representation* – morphing of real faces into the avatars

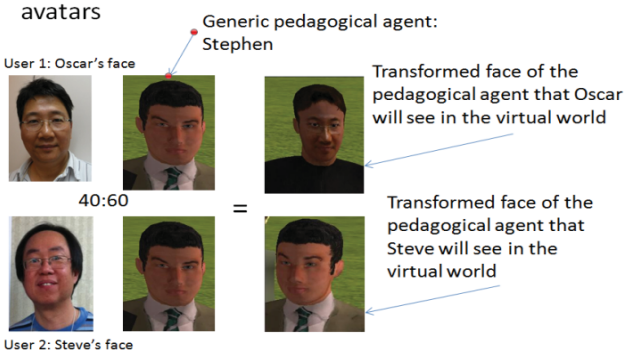


Figure 5: Creating the face of the host agent of players.

Using the evolver-multimesh module, all NPCs were imported into the Wonderland virtual environment. Furthermore, using the scripting-component module and the NPC module, these NPCs were given capability to produce animations such as clapping, cheering and public speaking. In CArtAgo model, NPCs have been wrapped as artifacts in order f or the Jason agents to manipulate, manage, and communicate with them. They can converse with contestants and agents that could react to external events such as contestant answers to qu estions, calculate and compute scores amongst other things. Figure 6

shows the audience performing the cheering and clapping animation.

Figure 6: The virtual audience performing the cheering and clapping animation.

Figure 7 shows the **Host agent** performing public speaking animation. By synchronizing the cheering and clapping animation with appropriate sound effects, and also host public speaking animation of the host with synthesized speech delivered by MaryTTS, an attempt was made to make these gestures seem as real as possible to the contestant.

Figure 7: the host agent performing public speaking animation.

2) Agent interactions

First the **Game agent** starts the Open Wonderland client. In order to achieve this goal, its first plan is to create an artifact called *WLStarterArtifact*. Then, a focus is done on the *WLStarterArtifact* so that any signal generated by *WLStarterArtifact* is observable by all agents in the current default workspace.

The next goal of the **Game agent** is to invoke the main entry point of the wonderland client which is *JmeClientMain*. Once the invocation of the wonderland client is successful, the *WLStarterArtifact* sends a *readytostart* signal which is perceivable by all agents within the default workspace. As soon as the **Host agent** percepts *readytostart* signal, it starts executing the plan with *QMIntroArtifact*. It goes on to execute the plan of *doIntroductionStage*.

Once the Introduction stage is done, the **Host agent** sends a signal "*introduce"* which is perceivable by **the Game agent**. After perceiving the "*introduce"* signal, the **Game agent** executes the plan with *ClockArtifact* – which starts and stops the timer for every question, and communicates with the **Host agent** to carry out the entire question and answer session, including displaying the questions, making sure that the **Host agent** delivers the question verbally along with necessary comments to the contestant. The Host agent performs this by communicating with the **Instructor agent and** *ClockArtifact*. Figure 9 shows the result of this stage in the form of a screen captured image.

Figure 9: A player is answering the question with a mouse.

At the beginning of the QuizMASter show during the initialization stage, contestants were given control of their avatars by enabling them to change the look direction using the left and the right arrow keys. When the show starts, during each question and answer sessions, the movements of the user are monitored by **the Behavior agent** using *CellTransform* which holds the translation and rotation parameters of each cell in Open Wonderland environment. All contestant movements are recorded in a log file in the name of the contestant. If it is detected that the contestant is looking away from the **Host agent**, the contestant cell is transformed using *CellTransform* so that he/she is facing the direction of the Host agent and the **Host agent** requests the user to pay attention to the quiz. This is how the 2nd dimension of TSI is implemented in QuizMASter. Figure 10 shows an example of the log files, which is generated for contestant named Steve.

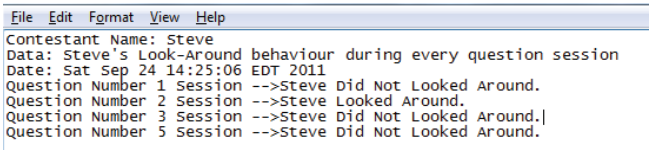


Figure 10: Contestant Steve's log file of his look_around behavior during Every Question Answer Session.

We successfully tested the prototype with an online course ENLI 255 of Athabasca University and two real players located in different places in Canada. The OWL server is installed in Northern Alberta of Institute of Technology of Canada.

V. CONCLUSIONS

This paper introduces an architectural design and prototype implementation of adaptive testing games within a virtual world.

In order to retain the functionality of having a MAS within the QuizMASter framework that was previously developed [12], we enhanced the architecture by incorporating a more versatile BDI agent model with AgentSpeak using Jason into it. Furthermore, there was an opportunity to implement the Agent and Artifact model proposed by CartAgO tool, that fully supports agents implemented using Jason and provides a bridge for agents developed using Jason. By leveraging the power of CartAgO and being able to drive and control the entire Open Wonderland client using Jason and CartAgO, QuizMASter could be developed primarily using artifacts whereby it can be driven entirely by intelligent agents. The seemingly endless possibilities of using QuizMASter as a test-bed for conducting current and future experiments using 3D virtual world along with sophisticated intelligent agents will provide a valuable research tool to investigate game based learning and adaptive assessment.

The main advantage of Jason-CArtAgO MAS is that it bridges the gap between designing MAS and programming MAS.

Thus far, we have identified the roles of agents and the three different architectures we developed to investigate MAS in a game environment. To make environments more effective for educational applications, we need to provide a personalized and customized environment where students feel encouraged. We are working on user modeling for such environments but more work about the coordination mechanisms for dynamic game formation and adaptive quizzes sequencing, and incorporation of the intelligence of the agents with the architecture proposed needs to be done.

To meet the needs of the next generation learners, educators must use innovative methodologies and approaches for solving problems in collaborative, disparate environments in a manner that improves learning outcomes. We will continue to explore the roles of agents and agent technologies in facilitating and supporting assessment-based learning wherein learners engage in self-directed learning and are able to test out of education or training by successfully completing validated assessments outside of structured and formal learning instances.

ACKNOWLEDGEMENT

We thank NSERC discovery grant of Canada and Canada Foundation for Innovation (CFI), Athabasca University, and communities for Open Wonderland, Jason, and JADE for their support and help to this research.

REFERENCES

[1] Pfeifer, R. and J. Bongard, (2007). How the body shapes the way we think: a new view of intelligence, MIT Press.

[2] Weiss, G. (2000). Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence, MIT Press.

[3] Avlett, R., M. Vala, P. Sequeira, A. Paiva, (2007). FearNot! – An Emergent Narrative Approach to Virtual Dramas for Anti-bullying Education, Virtual Storytelling. Using Virtual Reality Technologies for Storytelling, LNCS 4871/2007, 202-205

[4] Ricci, A., M. Viroli, and A. Omicini. (2007). CArtAgO: A framework for prototyping artifact-based environments in MAS. In D. Weyns, H. V. D. Parunak, and F. Michel, editors, Environments for MultiAgent Systems III, volume 4389 of LNAI, pages 67–86. Springer Berlin / Heidelberg.

[5] Dutchuk, M., K. A. Muhammadi, & F. Lin (2009), QuizMASter - A Multi-Agent Game-Style Learning Activity, EduTainment 2009, Aug 2009, Banff, Canada, Learning by Doing, (eds.), M. Chang, R. Kuo, Kinshuk, G-D Chen, M. Hirose, LNCS 5670, 263-272.

[6] Weng, M., F. Lin, T. K. Shih, M. Chang, I. Fakinlede, (2011). A Conceptual Design of Multi-Agent based Personalized Quiz Game, The 11th IEEE International Conference on Advanced Learning Technologies (ICALT), Athens, Georgia, USA

[7] Kaplan, J. and Yankelovich, N. (2011). Open Wonderland: An Extensible Virtual World Architecture, *IEEE Internet Computing*, **15**(5), 38 - 45.

[8] Bailenson, J. N., Yee, N., Blascovich, J. & Guadagno, R.E. (2008). Transformed Social Interaction in Mediated Interpersonal Communication. In Elly A. Konijn, Sonja Utz, Martin Tanis & Susan B. Barnes (Eds.), *Mediated Interpersonal Communication.* 77–99. New York and London: Routledge.

[9] Kaminka, G. A., Veloso, M. M., Schaffer, S., Sollitto, C., Adobbati, R., Marshall, A. N., et al. (2002). GameBots: a flexible test bed for multiagent team research. Communications of the ACM, 45(1).

[10] Gemrot, J., Kadlec, R., Bída, M., Burkert, O., Píbil, R., Havlíček, J., et al. (2009). Pogamut 3 Can Assist Developers in Building AI (Not Only) for Their Videogame Agents, 1–15.

[11] Dignum, F., Westra, J., van Doesburg, W. A., & Harbers, M. (2009). Games and Agents: Designing Intelligent Gameplay. International Journal of Computer Games Technology, 1-19.

[12] Blair, J. and F. Lin, (2011). An Approach for Integrating 3D Virtual Worlds with Multiagent Systems, ISeRim workshop - IANA 2011, Singapore.

[13] Thalmann, D., S. Musse, and M. Kallmann, (2000), From Individual Human Agents to Crowds, Informatique, 2000(No. 1), 6-11.

[14] Moshell, M. and C. Hughes (2002), Virtual Environments as a Tool for Academic Learning, in K. Stanney (Ed.), The Handbook of Virtual Environments Technology, Erlbaum, Mahwah, NJ, 893-910.

[15] Michael J. Jacobson, Beaumie Kim, Chunyan Miao, Zhiqi Shen, Mark Chavez, *Design Perspectives for Learning in Virtual Worlds*, in Jacobson, M. J., & Reimann, P. (Eds.). (2009). Designs for learning environments of the future: International perspectives from the learning sciences. Springer, 111-141.

Optimimal resources utilization for indexing within a distributed multimedia retrieval system: an implementation for a video surveillance use case

Dana Codreanu, Ana-Maria Manzat, Florence Sedes
Université de Toulouse – IRIT – UMR 5505
31062 Toulouse, France
{codreanu, manzat, sedes}@irit.fr

Abstract— *Effective and flexible solutions for enabling reduced resource consumption and handling content and associated processing (indexing algorithms) diversity (in terms of location, technical characteristics, execution constraints, requested features) are essential in multimedia management systems. In this paper we present a distributed multimedia retrieval system that enables efficient and effective content processing and access. Though the system's architecture is a generic one, in order to explain system's functionalities we will present an instance of the architecture for a video surveillance use case. The proposed system addresses all important topic research concerning this type of systems (architecture, indexing, query) and proposes a novel two steps indexing approach (at acquisition time: implicit indexing and at query time: explicit indexing). We also present in this article how the user access management is done.*

Keywords-component; multimedia information system; implicit explicit indexing; reduced resource consumption; video surveillance

I. INTRODUCTION

A huge quantity and diversity of multimedia contents is generated in multiple domains (video surveillance, medical records, libraries), with different acquisition contexts. With the constant growth of the multimedia contents volume, the indexing process becomes very important in their efficient management and retrieval. This process is the most resource consuming, in terms of memory, CPU and transactions over the network. This overconsumption is due to the fact that the indexing is accomplished by executing a fixed set of indexing algorithms over the entire multimedia collection on indexing dedicated servers.

From an operational point of view the problem of *resource consumption* is how to efficiently manage the huge volume of contents and the contents and contents processing techniques diversity in the context of a large scale information system in order to provide users with a high quality multimedia experience. The resource consumption reduction for indexing (the most consuming process) can be achieved in two ways: by limiting the multimedia content transfer over the network, and by using only the most appropriate algorithms, only over the relevant content.

Our solution addresses these two points by reducing as much as possible the resource consumption through (1) a distributed indexing technique that avoids multimedia transfer; (2) a dynamic indexing process that is accomplished in two steps: (a) at the acquisition moment and (b) at the query execution moment. Our proposal is based on generic description models associated to multimedia contents, indexing algorithms and servers. These models are explained in detail in [1, 2, 3].

These solutions are implemented within a *generic framework*, developed in the context of the LINDO[1] (Large scale distributed INDexation of multimedia Objects) ITEA2 project, which is intended to guide the formalization and the development of a distributed multimedia information system, while favoring a reduced resource consumption, in terms of data transfers over the network, storage and CPU utilization.

The video surveillance systems become more complex (geographically spread) and the volume, diversity and heterogeneity of the contents and acquisition contexts generate new requirements for multimedia information systems. The purely manually approach (agents watching the videos and detecting events) becomes insufficient. Also, the video surveillance domain raises challenges like: real time constraints (in some situations the optimal compromise between precision and processing time needs to be found) and emergency situations management. Therefore, the automatically access to the semantic content of surveillance audiovisual content is a challenging research area. All these characteristics and constraints led us to choose to instantiate the architecture in the video surveillance domain. We present a system that manages a network of surveillance cameras placed in buses, around the bus stations and the bus ticket machines in Paris.

In the remainder, the paper presents a synthesis of the existing works related to the addressed problems. The Section II**Erreur ! Source du renvoi introuvable.** provides an overview of the solutions adopted by some industrial projects that address the multimedia management problems. Then, the LINDO generic architecture is briefly presented in Section III. In the following the Implicit and Explicit Indexing Processes

[1] http://www.lindo-itea.eu

are explained in Section III.A and the emergency and security management in Section III.B. A concrete implementation of a video surveillance system using the LINDO architecture is detailed in Section IV. Conclusions and further work directions are exposed in the end of the paper.

II. RELATED WORKS

In the following we present some distributed systems that manage multimedia contents by emphasizing the architectural choice and the adopted solution for multimedia indexing.

Many distributed systems that were proposed for the management of distributed textual documents. [4, 5, 6, 7] adopt a peer-to-peer architecture where the contents are stored and indexed on distributed peers. The resulted index is distributed over the network using a hashing table. These systems benefit from the fact that the textual documents indexing is based on the identification of terms and the same model is employed for their descriptions. This is not the case of multimedia documents. The indexing algorithms used in this case are specific to each media type the generated outputs and metric spaces are heterogeneous. Thus, using peer-to-peer architectures for the management of multimedia contents is more difficult. The following projects addressed this problem in different manners.

The SAPIR (Search on Audio-visual content using Peer-to-peer Information Retrieval) project [8, 9] proposes a hybrid peer-to-peer architecture for the management of multimedia contents. It uses three specialized indexing servers (for images, texts and audio-visual contents) where each peer sends its ingested content in order to be indexed. The resulted metadata is sent back to the peer that ingested the multimedia content for storage.

The DISCO (Distributed Indexing and Search by Content) project[2] has chosen a structured peer-to-peer architecture for the management of multimedia contents [10]. The indexing is accomplished at each peer, at the moment when the content is stored. Each peer could run a different set of indexing algorithms, with an important constraint: the usage of the same descriptors and associated metrics. Each peer sends a summary of its index that is concatenated to a global index which is sent to all the other peers. The query is executed on this global index.

The CANDELA project (Content Analysis and Network DELivery Architectures)[3] is focused on the video content analysis and retrieval, in a distributed architecture [11]. Within this architecture the content is stored and indexed on distributed servers. The indexing is accomplished on all remote servers using a pre-defined and fixed set of indexing algorithms. The proposed solution was implemented for several use cases: personal mobile multimedia management [12], video surveillance [13, 14].

The WebLab project[4] proposes an integrative infrastructure that enables the indexing algorithms management as web services in order to be used in the development of multimedia processing applications [15]. These indexing services are handled manually through a graphical interface. For each specific application a fixed set of indexing tools is run.

The VITALAS (Video & image Indexing and retrieval in the Large Scale) project[5] capitalizes the WebLab infrastructure in a distributed multimedia environment [16]. The architecture enables the integration of partner's indexing modules as web services. The multimedia content is indexed off-line, at acquisition time, on different indexing servers.

The KLIMT (Knowledge InterMediation Technologies) project [17] proposes a Service Oriented Architecture middleware for easy integration of heterogeneous content processing applications over a distributed network. The indexing algorithms are considered as web services. The user query is limited to pre-defined patterns that match a set of rules for the algorithms' execution sequence. After such a sequence selection, the content is analyzed and the metadata is stored in a centralized database.

TABLE I. COMPARATIVE STUDY OF MULTIMEDIA MANAGEMENT PROJECTS

System	*Architecture*			*Indexing*		
	P2P	***SOA***	***Central Control***	***Uniform***	***Fixed***	***Content Transfer***
SAPIR	x	-	-	x	x	x
CANDELA	-	-	x	x	x	-
VITALAS	-	x	-	x	-	x
KLIMT	-	x	x	x	-	x
WebLab	-	x	x	x	x	x
DISCO	x	-	-	x	x	-

The comparative study of these systems (see Table I) shows that usually no matter the architectural choice, the content indexing is usually done on dedicated servers (and the content associated resulting metadata are transferred over the network) using a pre-programmed set of algorithms. These algorithms are executed on all ingested multimedia content (uniform indexing).

The main goal of LINDO project is to define a distributed system for multimedia content management, while focusing on the efficient use of the resources in the indexing and query processes. Thus the adopted distributed architecture enables: (1) to centralize the management of indexing algorithms and their deployment on demand on different remote servers; (2) to execute different indexing algorithms, in parallel, on different remote servers; (3) to dispatch user queries only on some specific remote servers that might contain some needed information; (4) to integrate new algorithms at any time into the architecture; (5) to accomplish the indexing in two steps: an implicit indexing that is executed over the multimedia contents from each remote server at the content acquisition and an explicit indexing that could be executed, on demand, on a specific remote server. In the following section, we present the adopted architecture, the implicit and explicit indexing processes and how the emergency and security management is done.

[2] http://www.lamsade.dauphine.fr/disco/index

[3] http://www.hitech-projects.com/euprojects/candela

[4] http://weblab-project.org/

[5] http://vitalas.ercim.org

III. THE LINDO ARCHITECTURE

The architecture adopted in LINDO is a distributed with central control one. Our architectural solution is divided into two main components: (1) *remote servers* which acquire, index and store multimedia contents, and (2) *a central server* which has a global view of the system and orchestrates the indexing and query processes. This architectural solution makes LINDO system scalable because it has the following advantages: (i) each remote server is independent and may have different characteristics (e.g., we can have a remote server in Paris where BBC video journals are stored and another in Toulouse where video surveillance contents are acquired); (ii) a remote server is composed of modules with different functionalities that communicate between them, with no given indication concerning the implementation; (iii) at each moment we can integrate a new remote server; (iv) the central server can send relevant indexing routines or queries to relevant remote servers, while the system is running.

The originality of this solution is that: (a) the content is not moved to indexing servers, but indexing algorithms are deployed on the servers where the content is acquired; (b) implicit/ explicit indexing (see Section III.A). Therefore, the proposed architecture can guide the design of distributed multimedia information systems, while favoring a reduced resource consumption, in terms of data transfers over the network, storage and CPU consumption. Actually, the idea was not to define yet another information retrieval model or indexing engine, but rather to encapsulate any existing frameworks for indexing and retrieving multimedia contents, like those used and/or defined by the projects presented in Section II.

A detailed presentation of the LINDO system architecture and each of its modules can be found in [3]. In this paper we will present a proposition of implementation for a video surveillance use case. This domain raises constraints like: real time queries, security, privacy and emergency issues etc. In the following we explain the implicit and explicit processes and how we handle emergency and security problems.

A. Implicit and Explicit Indexing Processes

Within a system that adopts the architecture that we presented above, the content is acquired and stored on the remote servers, and the collection of indexing algorithms is stored and managed on the central server. This collection is variable, at any moment we can integrate new algorithms with different functionalities, execution constraints and performances. Executing all these algorithms on all multimedia content will: (1) overload the system and (2) produce metadata that might never be used. Therefore, in order to save server resources, multimedia content indexing is realized at ingest time *implicit indexing* and on demand *explicit indexing*.

When a new multimedia content is ingested by a remote server, the Storage Manager (SM) (see Figure 2) module first stores it. Afterwards, the Feature Extractor Manager (FEMrs) module starts the *implicit indexing process*. This process consists in the execution of a set of indexing algorithms, i.e., implicit indexing algorithms (iIA), on the acquired multimedia content in order to extract some metadata that will be further queried. For that purpose, the FEMrs retrieves the multimedia content from the SM and executes these algorithms on the content. The obtained metadata is then stored by the Metadata Engine module (MDErs) in its metadata collection (in a format defined within the project that is not presented in this paper). Periodically, a Web Service sends to the central server a metadata summary that contains the essential information detected on each remote server.

The query workflow begins with the user query specification by using the Terminal Interface module (TI) (Figure 3). Each user query is sent to the Request Processor (RP) module which analyses the query in order to translate it into the formal language used by the MDErs and MDEcs (the metadata engines from the remote servers and from the central server). After this transformation, the RP selects, based on the location and on the metadata summary, the remote servers that might have results to the query. The RP sends the query to the MDErs of the selected remote servers, where the query is executed.

At this step of the query process it is possible that not all remote servers were selected for executing the query. This does not necessarily mean that they do not have results for the user's query. It is possible that their multimedia contents were not indexed with the right indexing algorithms. Thus, the RP module selects the remote servers that should be re-indexed and the supplementary indexing algorithms, i.e., explicit indexing algorithms (eIA). These algorithms are deployed on the selected remote servers in order to produce additional metadata that could provide answers to the user's query. The user is informed that supplementary indexing is performed in order to retrieve other results to his query. The selection algorithm of iIA and eIA based on user's queries (e.g., in a car park video surveillance system it is not necessary to apply algorithms which detect persons to all video sequences, but at a certain time someone might need information about persons that crossed the parking area) and execution constraints (e.g., an algorithm conceived to run only on Linux machines will not be deployed on a remote server that uses a Windows operating system) is presented in detail in [2].

At this step, the RP decides, according to the user's query (e.g., the time interval specified in the query), on what multimedia content these explicit indexing algorithms have to be executed. The FEMcs is in charge of the deployment of these explicit algorithms on the concerned remote servers, if necessary. From this point the FEMrs takes the control of the indexing, which is accomplished as for the implicit indexing. When the explicit indexing is finished, the MDEcs informs the RP that other results are available. At this moment, the RP module sends the user query to the MDErs of remote servers where the *explicit indexing* was accomplished. The results obtained on the two ways (Figure 1) are merged and displayed.

B. Emergency and security management

The two red points in Figure 1. Before selecting the algorithms for the explicit indexing, the RP checks if the user's role allows him to run explicit algorithms as these algorithms analyze more in detail the multimedia content and usually they consume a lot of resources (see Section IV.B for concrete examples that illustrate this problem).

Before displaying the results the Results Aggregator module (RA in Figure 2.) takes into consideration in one hand the user's context and role (access rights) and the permitted use and distribution rights of the content on the other hand. After matching these two the RA decides if content is displayed as it is or if it has to be modified (e.g., resize, apply a blur region, display only some parts of the content). (see Section IV.B for examples). The control access mechanism is the one presented in [18] and uses XACML and RBAC technologies.

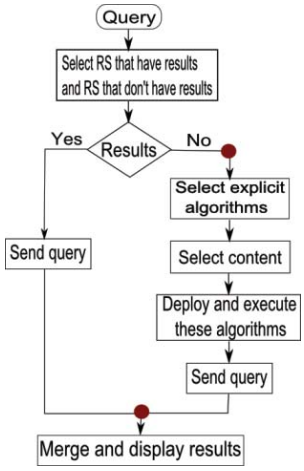


Figure 1. Explicit Indexing Process

IV. SYSTEM IMPLEMENTATION IN A VIDEOSURVEILLANCE USE CASE

This use case was proposed by one of the industrial partners. It concerns a public transportation company that placed surveillance cameras in their buses, around the bus stations and the bus ticket machines. Each video stream is recorded with a real time association to a common time stamp and GPS location. A typical query may be to find all videos containing a specific event, such as a person with a red pullover, who took the bus in Montparnasse station, Paris, on Monday 3rd May, between 6pm and 8pm.

A. Architecture instantiation

The analysis of the use case highlighted two types of contents: those generated by video surveillance cameras situated in fixed locations (the bus stations) and those situated in the buses. This observation led us to use several remote servers, each one managing contents acquired in similar situation (e.g., in bus stations; in buses). These remote servers store and index the video contents acquired in real time. The central server stores information about the users and their access rights (the MDEcs module), the remote servers (the indexing algorithms that are installed and a metadata summary that are produced on each remote server) and enables content querying. Thus, the system architecture is composed of a central server and three remote servers (see Figure 2). The type of content stored on each remote server, the acquisition context and the physical characteristics of the server influenced the choice of the implicit indexing algorithms.

In the following, we detail the particularities of each architecture module instantiation on each one of these servers, either remote or central.

The generic architecture of a remote server was instantiated for each one of the three remote servers. The instantiations maintained the architecture's modules, while adopting a different implementation of their functionalities, realized by different partners. For example, for the SM module two different proprietary versions were developed, one in C and the other one in Java language. For the FEM modules, two different implementations (in Java and C#) were adopted for the remote servers. The same *MDE* module was integrated in all three remote servers. This module was developed in Java and uses the XML native Oracle Berkley DB XML[6] database for storing metadata.

The metadata provided by the indexing algorithms used for this use case respect the XML data format presented in [1]. Each indexing algorithm available in the system has associated a description that respects the model described in [2]. All this information is stored by the *MDErs*.

For the communication between the different modules of a remote server we used the JMS[7] (Java Message Service) API. The communication between the remote servers and the central server is accomplished through Web Services.

The **first remote server (RS_1)** manages videos acquired in real time by two video surveillance cameras situated in the Trocadero Bus station, in Paris. The implicit indexing algorithms for video content are in charge of crowd detection and abandoned luggage detection [19].

The **second remote server (RS_2)** manages audio and video contents acquired in real time from video surveillance cameras situated at the entrance of the Montparnasse Station. It contains the following indexing algorithm: abandoned luggage detection, crowd detection and speech detection.

The **third remote server (RS_3)** manages video contents acquired in real time from video surveillance cameras installed on the busses that are used on the lines 1 to 18. The implicit indexing algorithms are: person counting and shout detection.

All the modules of the **Central Server** were implemented for this use case. As mentioned, the FEMcs manages the global indexing algorithm collection where, alongside with all the implicit indexing algorithms installed on the remote servers, a supplementary set of explicit indexing algorithms are stored (shape detection, color detection, person detection, etc. [20]). The MDEcs has a global view of the system, through the multiple data that it collects: descriptions of each remote server, metadata summaries from each remote server, descriptions of indexing algorithms, users access roles and rights and supplementary knowledge (e.g., bus time tables and GPS information).

This video surveillance system is used by security agents that work in the stations and also by policemen. Each of them has different rights, for example a security agent is not allowed to execute explicit algorithms or to see some audiovisual contents. A policeman is allowed to visualize all multimedia contents. An important constraint of the system is that the explicit indexing can be accomplished only from the control room.

[6] http://www.oracle.com/technology/products/berkeley-db/xml/index.html

[7] http://java.sun.com/products/jms/overview.html

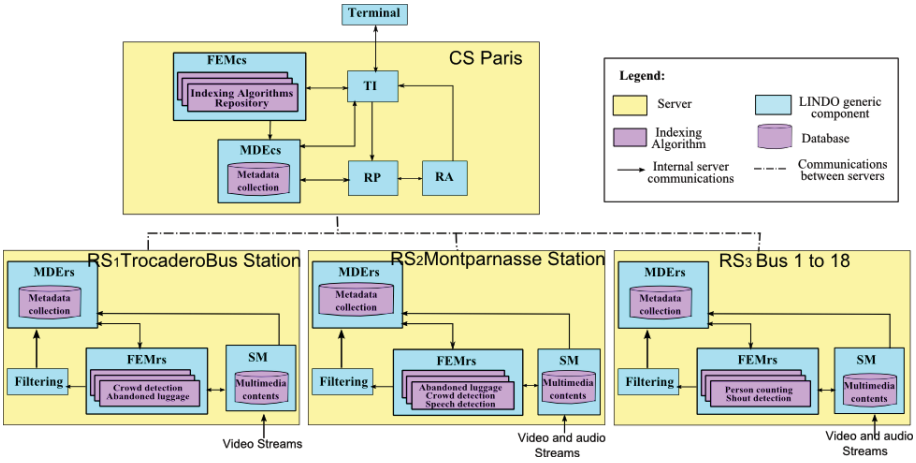


Figure 2. The LINDO testing system topology

B. System utilization examples

Let's suppose the following queries:

Query 1, entered by a security agent, on his PDA, in the station: show me all au diovisual content related to a bag forgotten in Montparnasse station, Paris, on Wednesday, 29 of February, between 10 a.m. and 2 p.m..

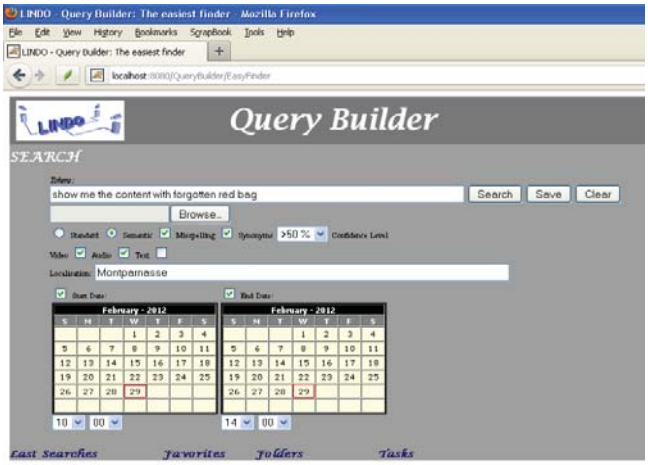


Figure 3. The query interface

Based on the localization specified in the query the remote server corresponding to Montparnasse station is selected. Then the query is an alyzed and sent to this remote server for execution (all the features required in the query (forgotten bag) are extracted by implicit algorithms on this remote server). Responses are found and they are sent to the RA for ranking. At this moment the RA checks if the security agent has the right to access all multimedia results. In our case the agent is not allowed to hear the audio contents. Thus, the audio parts of the contents are blocked.

Query 2, entered by a security agent, on his PDA, in a station: show me all v ideos with women in red, on Wednesday, 29 of February.

Because there is no localization specified in the query, all RS are selected. After analyzing the query and matching its required features with those extracted by the implicit algorithms none of the remote servers is selected as might having results. Thus the explicit indexing is necessary. At this moment, the RP checks the rights of the agent in the current situation. He does not have the right to run explicit algorithms and his context does not describe an emergency situation. So, the explicit indexing is not accomplished and a message is displayed to the user in order to let him know that no results were found for his current context.

Query 3, entered by a policeman, on his PDA, in a bus station: show me all audiovisual content with child dressed in red in Trocadero bus station, Paris, on Wednesday, 29 of February, between 10 a.m. and 2 p.m.

Figure 4. Example of detection

Based on the localization, the RS2 and RS3 are selected because RS2 is located in Trocadero station and RS3 stores the contents of bus line 2 which passes nearby Trocadero station. The query is analy zed in order to see if these servers might have results. In this case the explicit indexing is needed because on these servers the person and color detection algorithm is n ot executed. The RP checks the rights of the policeman in the current situation. Normally he does not have the right to run explicit algorithms, but his current context unlocks temporarily this restriction. (a policeman in a bus station might describe an emergency situation). Thus, the RP selects the explicit indexing algorithms based on the user's query. More precisely the algorithm that detects persons and

colors is selected. This algorithm is executed on the video content that corresponds to the specified time period. After the execution of the algorithm the query is executed on the remote server selected before and the results are transmitted to the RA and displayed to the user. An example of result is presented in Figure 4.

V. CONCLUSION AND FUTURE WORKS

In the context of the general process of distributed multimedia content management, this paper addressed the problem of resource consumption. The solutions proposed for this problem consist in: (1) adopting generic distributed architecture that avoids content transfer for indexing, (2) implementing a dynamic indexing process accomplished at acquisition time and at query execution time. These solutions were implemented and validated in the LINDO project. A concrete implementation for a video surveillance use case is presented in the paper. Some system utilization examples are described in order to illustrate the functioning of the system and how it manages the security and emergency situations.

As future work, we will improve the indexing algorithm selection process by taking into account multiple characteristics (e.g., their complexity, execution time). In the case of complex indexing needs, it is often necessary to execute into a certain order a chain of algorithms (e.g., French words detection algorithm must be executed only after a positive detection accomplished by the speech detection algorithm). Therefore, we intend also to include some pre-conditions and post-conditions specifications in the algorithms descriptions and to exploit them in order to automatically determine such chains specific to a complex indexing need. The set of indexing algorithms on each remote server might be variable (e.g., the performances of indexing algorithms depend on weather conditions). Therefore some tests were already been done in order to integrate remote site characteristics and context, as well as the user queries history in the dynamic selection of indexing algorithms.

ACKNOWLEDGMENT

This work has been supported by the EUREKA project LINDO (ITEA2-06011).

REFERENCES

[1] M. Brut, S. Laborie, A.-M. Manzat, and F. Sedes, "A Generic Metadata Framework for the Indexation and the Management of Distributed Multimedia Contents," In 3rd International Conference on New Technologies, Mobility and Security, 2009

[2] D. Codreanu, S. Laborie, and F. Sedes, "Modelling Multimedia Indexing Algorithms Description for Implicit and Explicit Indexation," In 5th International Conference on Multimedia and Ubiquitous Engineering, pp.232-238, 2011

[3] M. Brut, D. Codreanu, S. Dumitrescu, A.-M. Manzat, and F. Sedes. "A distributed architecture for flexible multimedia management and retrieval." In Proceedings of Database and Expert Systems Applications, pp. 249-263, 2011

[4] T. Suel, C. Mathur, J.-W. Wu, J. Zhang, A. Delis, Mehdi, X. L. Kharrazi, and K. Shanmugasundaram, "Odissea: A peer-to-peer architecture for scalable web search and information retrieval." In International Workshop on Web and Databases, pp. 67-72, 2003.

[5] S. Michel, P. Triantafillou, and G. Weikum, "MINERVA infinity: A scalable efficient peer-to-peer search engine." In Middleware, pp. 60–81, 2005.

[6] I. Antony, T. Rowstron and P. Druschel. "Pastry: Scalable, decentralized object location, and routing for large-scale peer-to-peer systems." In Middleware, pp. 329–350, 2001.

[7] L. Toan, F. Klemm, I. Podnar, M. Rajman, and K. Aberer, "Alvis peers: a scalable full-text peer-to-peer retrieval engine." In the int. workshop on Information retrieval in peer-to-peer networks, pp. 41–48, 2006

[8] M. Agosti, E.D. Buccio, G.M.D. Nunzio, N. Ferro, M. Melucci, R. Miotto, and N. Orio, "Distributed information retrieval and automatic identification of music works in SAPIR.", In the 15th Italian Symposium on Advanced Database Systems, pp. 479-482, 2007.

[9] B. Michal, F. Fabrizio, L. Claudio, N. David, P. Raffaele, R. Fausto, S. Jan, and Z. Pavel,"Building a web-scale image similarity search system" in Multimedia Tools and Applications, vol. 47, pp. 599-629, 2010

[10] F. Boisson, M. Crucianu, and D. Vodislav, "Publication Framework for Content-Based Search in Heterogeneous Distributed Multimedia Databases". Scientific Rapport CEDRIC No 1585, 2008

[11] M. Petkovic, and W. Jonker, "Content-Based Video Retrieval: A Database Perspective", Kluwer Academic Publishers, Boston, Monograph, 2003

[12] P. Pietarila, U. Westermann, S. Jarvinen, J. Korva, J. Lahti, and H. Lothman , "Candela-storage, analysis, and retrieval of video content in distributed systems: Personal mobile multimedia management", in IEEE International Conference on Multimedia and Expo, pp. 1557-1560, 2005

[13] P. Merkus, X. Desurmont, E. Jaspers, R. Wijnhoven, O. Caignart, J.-F. Delaigle, and W. Favoreel, "Candela - integrated storage, analysis and distribution of video content for intelligent information systems", in European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology, 2004

[14] E.G.T. Jaspers, R.G.J. Wijnhoven, A.H.R. Albers, X. Desurmont, M. Barais, J. Hamaide, and B. Lienard, "Candela-Storage, Analysis and Retrieval of Video Content in Distributed Systems: Real-Time Video Surveillance and Retrieval", in IEEE International Conference on Multimedia and Expo, pp.1553-1556, 2005

[15] P. Giroux, S. Brunessaux, S. Brunessaux, J. Doucy, G. Dupont, B. Grilheres, Y. Mombrun, A. Saval, "Weblab : An integration infrastructure to ease the development of multimedia processing applications", In the 21st Conference on Software and Systems Engineering and their Applications, 2008

[16] M.-L. Viaud, J. Thièvre, H. Goëau, A. Saulnier, and O. Buisson. "Interactive components for visual exploration of multimedia archives." In the 7th ACM International Conference on Image and Video Retrieval, ACM, pp. 609-616, 2008

[17] V. Conan, I. Ferran, P. Joly, and C. Vasserot, "KLIMT: Intermediations Technologies and Multimedia Indexing." In 3rd International Workshop on Content-Based Multimedia Indexing , pp. 11-18. 2003

[18] D. Al Kukhun, and F. Sedes, "Adaptive Solutions for Access Control within Pervasive Healthcare Systems". In ICOST, pp 42-53, 2008

[19] R. Miezianko, and D. Pokrajac, "Detecting and Recognizing Abandoned Objects in Crowded Environments" In 6th International Conference on Computer Vision Systems, pp. 241–250, 2008

[20] C. G. Snoek, and M. Worring, "Multimodal video indexing: A review of the state of the art." In Multimedia Tools and Applications, vol. 25, pp. 5- 35, 2005

Functionalities and Flow Analyses of Knowledge Oriented Web Portals

Daniele Cenni, Paolo Nesi, Michela Paolucci
DISIT-DSI, Distributed Systems and Internet Technology Lab
Dipartimento di Sistemi e Informatica, Università degli Studi di Firenze, Firenze, Italy
http://www.disit.dsi.unifi.it/, cenni@dsi.unifi.it, paolo.nesi@unifi.it, paolucci@dsi.unifi.it

Abstract

The concept of knowledge starts from the individual, from his critical analysis and on how one interacts with people and objects around. The goal of this paper is (i) to investigate the treatment of knowledge, (ii) highlight that Social Networks (SNs) and Best Practice Networks (BPNs) are the best Web technologies available on the market for its management, (iii) make a comparison among the most diffused Social Platforms with the Best Practice Networks, in terms of functionalities offered for the knowledge management, analyzing the knowledge flows in all their complexity and entirety, (iv) evaluate all the previous aspects, taking into account our BPN applied in a Social Portal. Our aim is to study a new model to make a comparison among the most diffused SNs and the BPNs, that we are developing. This model takes into account the functionalities available in the portal, and the treatment of the knowledge.

Topics: collaborative and social multimedia systems and solutions, management and fruition, intelligent multimedia computing, user profiling.

1. Introduction

In recent years, organizational aspects related to knowledge management and to intellectual capital management have been largely studied; many methods and models of interaction and communication among people have been developed, in order to study the organizational implications related to the creation, storage, movement and use of knowledge [1], [2], [3], [4], [5]. In many of them the knowledge is not seen as a static concept but as a continuous becoming: it is a set of ideas that are born, spread and grow up through the interactions among people. Moreover the activities of creation, acquisition, management, dissemination, use and application of knowledge involved in working flows, involving individuals, are heavily influenced by the dynamics of social and organizational type [6].

These models have been fundamentally applied on business contexts. However, the methodologies adopted are so general that can be easily extended and applied in a variety of contexts. One of the most relevant knowledge treatment models is the SECI, studied by Nonaka and Takeuchi. This model sees the knowledge as a collection of facts, information, skills that a person acquires in the course of his/her life. Basically, two types of knowledge are outlined: explicit and tacit. Explicit knowledge is a formal, systematic language, that can be expressed and shared in the form of data, in scientific formulas, specifications, manuals and so on; it can be processed, transmitted and stored relatively easily; tacit knowledge is in the human mind and it is highly personal and hard to formalize, externalize or mediate, represented by using the normal channels of communication. The SECI model talks about how the tacit knowledge is converted in explicit and vice versa. The SECI model defines four modes of knowledge conversion: Socialization (from tacit to tacit knowledge); Externalisation (from tacit to explicit knowledge); Combination (from explicit to explicit knowledge); and Internalization (from explicit to tacit knowledge) [2] (see Fig.1).

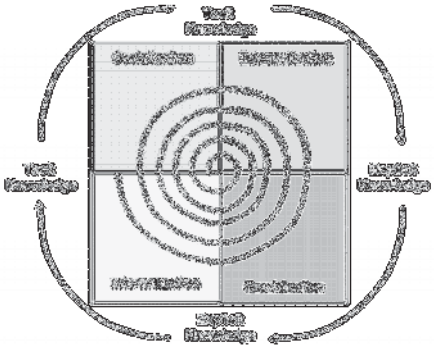


Figure 1: The SECI spiral model, inspired by [1].

In Fig. 1 it is also possible to see the Nonaka's spiral-like process: the knowledge is transformed through exchanges among individuals and groups, within the organization, across the four modes of knowledge conversion described above. This process does not stop once it has closed the circle, but continues into a new knowledge-creating spiral [3] [4]. This result increases the intellectual capital in its various components: organizational and relational. An appropriate organizational structure and the correct technology solutions, adopted in support of such a cycle, may enhance the effects of knowledge management in the growth of intellectual capital [6]. The most advanced organizations, structure their intellectual capital through the ICTs, often using Social Networks; from this point of view, many studies have been developed to match ICTs and SECI Model, such as [7] [8]. Some Social Networks are mainly focused on contents (e.g., YouTube, Flickr, LastFm) [9], whereas others are more focused on establishing relationships among users

(Facebook, MySpace, Orkut, Friendster, etc.) [10]. SN Analysis is typically focused on analyzing relationships among users and group of users, in order to identify which are the most central users and groups and, on the other hand, which are those that are frozen out, namely being those running the risk of losing interest in the network activities, due to a serious lack of involvement [11] [12]. On the other hand, Best Practice Networks are thematic SNs, where smaller groups of users share content, contacts and information with a common goal. In this paper we analyze the differences among the SNs and the BPNs, in terms of functionalities and knowledge flows; the goal is to study a methodology that can be used to classify and establish the most efficient instrument, in terms of stimuli given to the users, to enhance their creativity and to develop new knowledge. After this, we apply this methodology and verify its validity, through the analysis of data related to users' activities on our BPN, applied in a real context (i.e., the ECLAP portal). The paper is organized as follows: section 2 describes the SECI knowledge model applied on SNs and BPNs; in section 3 a comparison among SNs and BPNs in terms of functionalities is presented; in section 4, the functionalities are described as flows of knowledge; section 5 an analysis related to our Best Practice Network is presented.

2. SECI model, SNs and BPNs

Through the SECI conversion process, tacit and explicit knowledge expands in both quality and quantity, both in the Epistemological Dimension and also in the Ontological plan, that takes into account the human types of interactions (see Fig. 2).

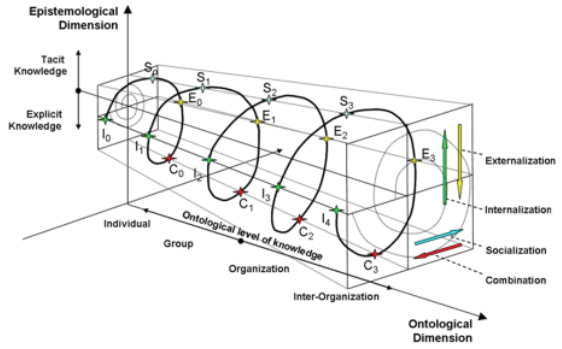


Figure 2: The SECI spiral model.

Given the wide spread of SNs and BPNs, the more relevant functionalities provided by these technologies will be analyzed and related to the SECI model. One of the aims is to determine which are the most capable platforms, to provide stimuli and incentives for the users, that are capable to assess the effectiveness of a platform, with respect to the others. Both the SNs and the BPNs provide for their users a set of functionalities, allowing them to make actions on the portal. Each user action has its flow of knowledge, basing on the SECI spiral model: it starts when people use the platform (Externalization), passes through the platform elaboration (Combination) and ends creating new knowledge or giving stimuli to the users (Internalization), to elaborate new ideas also involving the network of users (see Fig. 3). Talking or meeting other people both in the virtual spaces provided by the Platform (e.g. friends, forum, blog, thematic groups, suggestions, etc.) but also in the real world (Socialization), facilitates the users in exchanging their ideas, thus allowing the growth of the knowledge level, and stimulate the users to start again to put new knowledge in the platform (Externalization mode).

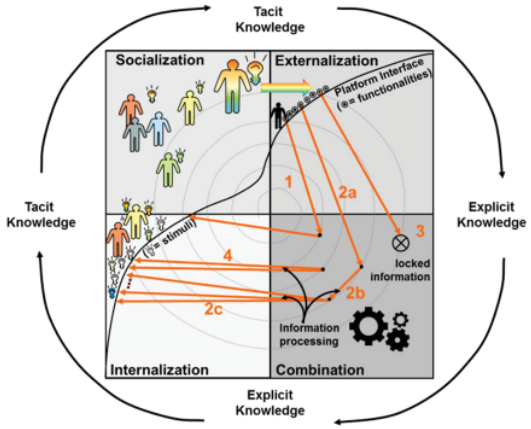


Figure 3: SECI – BPNs, SNs: actions flow.

The flows of knowledge, that are present in a Social Platform (SNs or BPNs), can be classified in four different types:

- *Direct flow* (User to Users): includes actions carried out by users and addressed to other users of the platform, with a little work (in terms of knowledge-reprocessing transformation) made by the platform itself (e.g. make a comment on a content, add a new friend, etc.);
- *Mediate flow* (User to Platform and Platform to Users): includes actions brokered by the platform, actions in which the users put their knowledge on the portal (uploading content, filling their user profile, ranking, etc.). The portal processes the information and produces stimuli for them or for other users; in this case the Portal itself generates new knowledge. This type of flow is composed of three main steps: *i*) actions done by the users on the Platform; *ii*) analysis and elaboration of data made by the platform; *iii*) production of stimuli for the end users (e.g. recommendations, suggestions, etc.).
- *Locked flow* (User to Platform): it occurs when users put their knowledge in the platform or make actions on it, and the platform does not register the information given, with a consequent loss of information;
- *Platform flow* (Platform to Users): it includes all the functionalities automatically provided by the platform that do not depend on the user's actions or on the knowledge that they have put in the platform (e.g. basic functionalities such as: give

the possibility to upload content, to make queries on content, to search new friends, etc.).

All the flows that begin on Externalization are made by the users: they produce new knowledge because they put their ideas, interests and creativity on the Portal. All the flows that begin in Combination, involve the Platform processing capabilities that can be minimal (for example when the user writes a web page, or makes a query on the portal), or can be very complex and capable to create new information, useful to stimulate the users (this is the recommendation sample). Moreover the more are the functionalities offered to the users (in Fig. 3, they are represented by the arrows ending on Internalization), and the more are the stimuli given to them. Following the SECI spiral model, users can do actions on the portal (Externalization) and, if they receive interesting stimuli from the Platform (Combination and Internalization), they can chat, share information or competencies, talk together (passing from the Socialization mode of knowledge conversion) and put new knowledge on the Portal (again starting their actions on Externalization, this time with a more high level of knowledge, see Fig. 3).

3. Comparison among the BPNs and the Social Networks on Functionalities

Both the SNs and the BPNs provide for the users a set of functionalities, allowing them to make actions on the portal. In order to analyze what type of platform can be considered the most effective and efficient to stimulate the creativity of the users, enhancing them to develop new flows of knowledge, we have realized a comparison in terms of functionalities. We have chosen YouTube (SN oriented to content), Facebook (SN oriented to users) and our Best Practice Network. We have selected the following functionalities as the most effective to give stimuli to the users:

- Content Download and View
- User Friendship
- Portal Registration
- Suggestions
- Searching
- Content Ranking

It is possible to see the result of this comparison in Fig. 4. The functionalities can be offered both to Public Users (PU) and to Registered Users (RU), and can have a different level of importance (high, low or medium), depending on the Platform mission or structure.

Functionalities for Public Users (PU) & Registered Users (RU)		YouTube	Facebook	Our Best Practice Network
Content download & view	Audio	-	-	high (PU&RU)
	Video	high (PU)	low (RU)	high (PU&RU)
	Images	-	High (RU)	high (PU&RU)
	Documents	-	Medium (RU)	high (PU&RU)
	Cross Media	-	-	high (PU&RU)
Portal Registration		low (PU)	high (PU)	high (PU)
User Friendship		low (RU)	high (RU)	high/medium (RU)
Suggestions to users	Same for all	high (PU&RU)	low (RU)	high (PU&RU)
	Different for each user	medium (RU)	medium/high (RU)	high (RU)
Searching	Content	high (PU&RU)	-	high (PU&RU)
	User	-	high (RU)	high (RU)
Content Ranking		high (PU&RU)	medium (RU)	medium (RU)

Figure 4: Functionalities on YouTube, Facebook and Our BPN.

It is possible to observe that:

- YouTube is focused on contents; user registration to the portal is not so relevant: the main services are provided for the Public Users and are obviously related to the contents;
- Facebook practically provides functionalities only for registered users, and it is mainly centred on making new friends and relationships among users, even if it keeps attention also on contents;
- BPN is equilibrated in terms of functionalities and on the type of users for which they are provided.

4. Analyzing functionalities and flow of knowledge

In this section it is detailed how the Portal's functionalities can be matched with different types of flows of knowledge. Each flow is composed of actions made by the users or by the platform itself; each action (see Fig. 5) starts and ends in one of the four SECI model modes of knowledge conversion, and corresponds to one of the flows (or even only to a part of them, in the case of type 2 flow), identified in the preceding paragraph. Each mode of conversion involves a knowledge transformation:

- Externalization: from tacit to explicit (T→E)
- Combination: from explicit to explicit (E→E)
- Internalization: from explicit to tacit (E→T)
- Socialization: from tacit to tacit (T→T)

Actions that transform the knowledge from tacit to tacit are almost never supported by virtual media. These actions usually take place in the real world, and are not supported in their entirety by SNs or BPNs platforms (conversations, meetings, reunions, etc.). In Fig. 5 is presented the model to compare the functionalities and the types of knowledge flows in a SN or in a BPN (detailed in section 2).

Functionalities and flows of knowledge					
	Start		flow type	End	
	Kind of user	Action		Result of the action	Kind of user
Functionality	Individual, Platform	T→E E→E (E→T)	1,2, (3) or 4	E→E E→T (T→E)	Platform, All users

Fig. 5: Functionalities and flows of knowledge model.

It is useful to make some examples and take into account the legend here after.

Action: Write a public Web page. Knowledge flow type: *Direct flow*.

- Start: the user starts from its own knowledge (tacit), uses editing tools and creates a web page on the platform (explicit). So the passage of knowledge is from tacit to explicit: T→E;
- Result: this is the result of the user action. The user has produced a Web page, visible to all the users of the platform (explicit); the users can read it and get information (tacit): E→T.

Action: Recommendations. Knowledge flow type: *Mediate flow*.

- Start: the platform produces and offers to the user a list of recommendations (explicit) and the user sees them. In this case the knowledge flow is from explicit to explicit: E→E;
- Result: the user reads the recommendations (explicit), acquires and reworks the information provided by the platform (tacit): E→T.

Platform functionalities and flow types					
	Start		flow type	End	
	Kind of user	Action		Result of the action	Kind of user
Content view	Platform	E→E	4	E→T	All users
Content download	Individual	T→E	1	E→T	All users
Portal Registration	Individual	T→E	2a	E→E	All users
User Friendship	Individual	T→E	2	E →T	Individual (friend)
Suggestions to users	Platform	E →E	4	E →T	All users
Searching	Individual	T→E	2	E →T	Individual
Content Ranking	Individual	T→E	1	E →T	All

Figure 6: BPN, sample on user functionalities.

In Fig. 6 are described the top functionalities, previously used to compare the SNs and the BPNs and their relations among the different knowledge flow types analysed in section 2. It is possible to note that:

- YouTube is focused on functionalities starting on the Combination mode of knowledge conversion;
- Facebook provides functionalities that involve flows mainly starting on Externalization (new knowledge from the users);
- The BPN is equilibrated: the functionalities involve all the knowledge mode of conversion.

5. BPN functionalities analysis

The analyzed BPN has been developed on the basis of an infrastructure called XLMS (Cross Media Learning Management System). Other BPNs has been deployed in various contexts by using XLMS infrastructure. The most relevant BPN from the point of view of the knowledge management analysis is ECLAP. ECLAP Portal is a BPN and content aggregator for the ECLAP European Project, [13]. The collected data are related to both content present in the platform and user behaviors. ECLAP Portal resources are divided in Cross Media and Drupal contents ([14]), accessible through Searching or Menus of the portal. Cross Media types include: archive, audio, braille music, collection, crossmedia, document, event, epub, excel, flash, html, image, pdf, playlist, slide, smil, tool and video. Drupal Content types include: blogs, forums, groups and pages. Cross Media contents may be annotated, recommended to other users, added to a playlist or to a personal collection, edited, downloaded, marked as favorite or featured, commented, tagged and voted. Drupal contents may be commented and voted. Also, each Cross Media or Drupal content may be linked and suggested to users in other Social Networks, with 1-click icons (i.e., del.icio.us, digg.com, Facebook, Twitter, MySpace, Messenger, Orkut, Plaxo, LinkedIn, Google). The following analysis has been conducted in the period September 1st – November 30th 2011. The ECLAP community consisted in 606 registered users, with 23544 digital items.

5.1 Cross Media Playback and Download

When playing multimedia contents (videos, audios, documents, images, and cross media), a right panel shows related metadata in various languages, content taxonomical classification, IPR information, and geolocation data. Cross media content access is one of the most exploited functionalities on the ECLAP portal. Cross media contents collected 7084 views and 1080 downloads. Top accessed contents include pages, pdfs, images and videos. These kinds of content can be accessed through searching, popular content lists or email notifications. A small fraction of the users accessing Cross Media content decided to download it. Images and videos are viewed by users, but generally not downloaded. Downloaded Cross Media contents include resources that are not usually intended for leisure or entertainment, so that the typical downloader is a technical/professional user. Playlists and Collections were the less exploited contents on the portal, with few creations and accesses from users.

5.2 Drupal Resources View

The main Drupal contents include pages and groups. Registered users can create, edit and translate these contents, with an online rich text editor (plain or html).

The ECLAP portal features two types of blogs: general and groups related. The general blog is the main project repository page, containing updated news about the project related tasks and activities. This single page collected nearly 100 clicks per day from the ECLAP menu. Groups related blog accesses are about 6.2% of all blog accesses. Pages are accessible through searching, groups, newsletters, or content lists; despite the relatively limited number of page items on the portal, this kind of content was fully exploited by users, with more than 100 accesses per day in the considered period. The most 3 accessed pages, collected about 26.4% of total page accesses (respectively 9.6%, 8.5% and 8.4%). After groups, pages were the top accessed Drupal resources (44.17% of Drupal content accesses, 33.09% of total views on the portal). The general ECLAP page collected 673 accesses (7.47 per day, about 1 access each 12 visits to the portal). During the sample period, the 28 groups created since the start of the project collected 9789 views, more than 100 per day. This was a very popular resource highly exploited by users, and the top most exploited resource, collecting about 46.27% of all Drupal views (34.66% of total views). Drupal resource had 74.91% of total views performed on the portal.

5.3 User Friendship

Each registered user can receive connections and friendship requests from other users in the ECLAP BPN; each request is notified in a right box, and can be accepted, ignored or denied. A list of potential colleagues is provided for the logged user. Users have a personal page that includes: specialization and job, general information, proximity details, list of colleagues, messages and subscriptions. Registered users have inbox and sent folders, and can send and receive private messages. Friendships and connections between users were established mostly by partners (92.26%); messages were exchanged mostly by partners too (66.67%). These numbers suggest that the registered social activities were about technical tasks between the project partners.

5.3 Portal Registration

Anonymous users can register on the ECLAP BPN with the Register button in the top right of the portal. After entering a captcha secure code and registering, the user receives an email with a confirmation link, for profile activation. In the considered period, there were 396 user registrations, with an average of one registration each 20 visits on the portal.

5.4 Suggestions

Content suggestions are available at access level or menu level. Similar contents were accessed 2864 times. This means that on average 9% of users who accessed contents, decided to play a content proposed to be similar to the accessed one. Similar content views per visit was 0.35. Featured content was the most exploited content list, with 343 clicks (about 30% of total content list accesses), followed by Last Posted (23%) and Popular contents (20%). Content lists collected 1140 clicks from users, about 0.14 clicks per visit and 12.67 clicks per day.

5.5 Searching

Search is divided in simple and advanced. Simple search allows full text search of ECLAP resources, eventually filtered by type. Query services were one of the top exploited functionalities in the portal. Most of the queries were of simple type (about 99%), performed from the main page, and a limited fraction were advanced queries. 65% of queries were performed from public users. Users performed about 0.50 simple queries per visit on the portal. As for advanced queries, faceted queries were not significant in number; 76.56% of faceted search were performed by public users. Queries through keyword or query clouds were about 4.9% of total simple queries. 46.34% of advanced queries were performed by public users. 1764 items was viewed after performing a query, thus 39.33% of all queries performed resulted in a click by the user on some result. 411 viewed contents after performing a query were of type Drupal (23.2%), and 1353 cross media (76.70%). This is in line with the data accounting for 74.91% of Drupal content views and 25.08% cross media views, so that user preferred more Drupal related contents to cross media, despite the limited amount of Drupal items, compared to Cross Media contents. Top search result clicks were on groups (12.14%).

5.6 Content Ranking

ECLAP contents can be ranked by registered users, using a 5-star based scale. Top ranked items are collected into a separate list, on the right menu of the portal. Contents report their average score. Votes assigned to contents can be changed at any time.

Functionality		Target	Source Path	Total users' Accesses	
Oriented to manage content	Crossmedia View/Play	RU&PU	Search Filter (*2*) Content lists (*2*) Keyword Cloud (*2*) Query Cloud (*2*) Groups (*2*)	*7084* (pdf: *2543*, image: *2477*, video: *1614*, …)	**Total users' actions on content** *12656*
	Searching (*Simple, Advanced, Faceted, Cloud search*)	PU&RU	Search Form (*0*) Menu (*1*)	*4484*	
	Crossmedia Download	RU&PU	Search Filter (*2*) Content lists (*2*) Keyword Cloud (*2*) Query Cloud (*2*) Groups (*2*)	*1080* (pdf : *884*, document: *70*, slide: *39*, …)	
	Crossmedia Ranking/ Voting	PU&RU	Featured (*2*) Search Filter (*3*) Content (*3*) Keyword Cloud (*3*) Query Cloud (*3*) Groups (*3*)	*8*	
Oriented to manage users and social interactions	Drupal contents View (*blogs, pages, groups, forums*)	RU&PU	Groups (*3*) Menu (*1*) Search (*2*)	*20111*	**Total users' social actions** *20522*
	Portal Registration	PU	Register Button (*1*)	*396*	
	User Friendship (*direct messages, invitations*)	RU	Menu (*3*)	*15* (messages: *9*, invitations: *6*)	

Figure 7: BPN, sample on user functionalities.

5.7 User Behavior

Public users typically started their session by clicking on the ECLAP menu in the top home page (24.73%), viewing a content (cross media 21.83%, drupal 15.83%, group 14.77%), or performing a query (12.79%). Most

popular second and third clicks nearly exhibited the same behavior. Users finished their sessions by viewing a page, clicking on a menu or accessing a group page. Registered users started their sessions logging in the portal (50.83%), then clicking on the ECLAP menu, and viewing an event page or a generic page (8.97%). Most popular second clicks include menu access (28.32%), viewing event pages (14.33%), and pages (11.6%). Third clicks generally were menu access (31%), viewing event pages (20.93%), and viewing groups pages (11.24%). Top sessions last clicks were on event pages (23.25%), cross media content access (54%) and menus (17.6%). In Fig. 7, we classify the main functionalities, just described as oriented to manage content or to enhance social interactions. Taking into account the table it is possible to note that:

- functionalities are mainly addressed both to Public Users and to Registered Users;
- user actions (evaluated through their access on the portal) are equally: *i*) finalized to manage content (total access: 12.656) and to establish social connections on the portal (access: 20.522) *ii*) distributed in terms of SECI modes of knowledge conversions (with the exception of the Socialization, that is mainly based on relations among people in the real world, and not through the use of web platforms);
- social connections are mainly oriented to contact the ECLAP Partners, because of the Portal mission. In general, the functionalities provided are oriented to enable the Partners to do their work on the BPN, and not created to enhance leisure or for entertainment (aspects that could be more interesting for new users).

6. Conclusions and Future work

In this paper we have described a new modality to analyze the capabilities of the SNs and the BPNs, starting from the SECI model, analyzing the functionalities offered by the Social Platforms in terms of flows of knowledge. We have collected and evaluated data related to our Best Practice Network applied in a real context (i.e., the ECLAP portal). Future work includes: *i)* increase the user behavior analysis, especially taking into account the sequence of the actions made by the users during a session. This, in order to establish if a functionality provided by the platform has really stimulated the users (to produce new knowledge, to make new friends, to chat with colleagues, etc.); *ii)* establish metrics to evaluate what are the points of strength or weakness of the platform, to give new stimuli to the users: how to increase the data processing capabilities (Mediate and Platform flows) and decrease the loss of information (Locked flow); *iii)* make a comparison among our portal, eventually applied in other contexts, and the most diffused Social Portals at data level and not only basing on functionalities (through the data crawling, etc.).

7. References

[1] Ikujiro Nonaka, Katsuhiro Umemoto and Dai Senoo. “From Information processing to knowledge Creation: A Paradigm Shift in Business Management”. *Technology In Society*, Vol. 18. No. 2, pp. 203-218, 1996.

[2] Ikujiro Nonaka, Ryoko Toyama and Noboru Konno. “SECI, Ba and Leadership: a Unified Model of Dynamic Knowledge Creation”. *Pergamon Long Range Planning 33* (2000) 5±34.

[3] Malin Brännback. “R&D collaboration: role of Ba in knowledge creating Networks”. *Knowledge Management Research & Practice* (2003) 1, 28–38 & 2003.

[4] John L. Rice & Bridget S. Rice. “The Applicability of the SECI Model to Multi-Organizational Endeavours: an integrative review”.

[5] V. Richard Benjamins, Dieter Fensel, Asunción Gómez Pérez. “Knowledge Management through Ontologies”. Proceedings of the Second International Conference on Practical Aspects of knowledge Management, 1998.

[6] M. Grimaldi, P. Rippa. M. Ruffolo. “A Methodology to Evaluate the Organizational Impact of IT on Knowledge Management: An Italian Case Study”. *Journal of Information Technology Case and Application Research* (JITCAR), Volume 10, N° 2. 2008

[7] Mohamed Amine Chatti, Ralf Klamma, Matthias Jarke, Ambjörn Naeve. “The Web 2.0 Driven SECI Model Based Learning Process”. *Proceedings of the 7th IEEE International Conference on Advanced Learning Technologies* (ICALT 2007), (pp. 780–782). IEEE Computer Society.

[8] Hugh Dubberly, Shelley Evenson. “Design as Learning - or “Knowledge Creation” - the SECI Model”. *Interactions magazine*, volume 18 Issue 1, January + February 2011, ACM New York, NY, USA. Doi:10.1145/1897239.1897256.

[9] G. Kruitbosch, F. Nack. “Broadcast Yourself on YouTube – Really?”. HCC’08, October 31, 2008, Vancouver, Canada.

[10] A. Nazir, S. Raza, C. Chuah. “Unveiling Facebook: A Measurement Study of Social Network Based Applications”. IMC’08, October 20–22, 2008, Vouliagmeni, Greece.

[11] P. Bellini, D. Cenni, A. Fuzier, P. Nesi, G. Pantaleo, M. Paolucci. “Metrics for best practice networks analysis”. *The 16th International Conference on Distributed Multimedia Systems*, Hyatt Lodge at McDonald's Campus, Oak Brook, Illinois, USA. October 2010.

[12] Gautam Das, Nick Koudas, Manos Papagelis. “Efficient Sampling of Information inSocial Networks”. CIKM/SSM 2008 Napa Valley, California, USA.

[13] European Collected Library of Artistic Performance. Platform available at: http://www.eclap.eu

[14] Drupal, Open Source CMS. Internet: http://drupal.org

Enhanced Fuzzy Multiple Regression Neural Model for Software Effort Estimation

Khaled Mohamed Shams D r. Amr Kamel
Faculty of Computers and Information
Cairo University

Faculty of Computers and Information
Cairo University

Abstract — **Quantitative methods have nowadays become very important tools in estimating the effort and cost in software development and support projects. Estimation accuracy is one of the most successful factors for a project's success; hence, the research has not stopped in improving the estimation models. Artificial neural networks (ANNs) are flexible computing frameworks that can be applied to a wide range of forecasting problems with a high degree of accuracy. However, ANNs need a quality training data in order to yield accurate results, data filtering and choosing the right ANN is important for the ANN to gain good results. In this paper, based on the basic concepts of ANNs and the use of Fuzzy regression as data filters a new hybrid method is proposed which yields more accurate results with non-quality data,**

Keywords-component; neuro-fuzzy, neural networks, fuzzy regression, software cost estimation, Back-propagation, Radial basis functions

I. INTRODUCTION

There are several effort estimation methods which are used to estimate the required amount of effort to successfully deliver a software project such as: COnstructive COst MOdel (COCOMO) [10], Software LIfe-cycle Model (SLIM), Experts' judgment, etc [1], [2], [3]. The mentioned methods are parametric methods, and the main conclusions were that these models perform poorly if they are not calibrated to project environment and technical characteristics [4], [5]. Parametric methods, often unable to capture the complex set of relationships between dependent and independent variables

(e.g. the effect of each variable in a model to the overall prediction made using the model), they are not flexible enough to adapt to a different environments, and they can't learn from their previous knowledge [13].

To evaluate relationship between dependent and independent variables, the Regression Analysis is one of the most widely used methods for this, and the Fuzzy Regression (FR) is a fuzzy type of classical regression analysis in which some elements of the model are represented by fuzzy numbers [6],[7]. The importance of the Fuzzy Regression, because of the lack of perfect fit of numeric data to the assumed linear format of the relationship between input data. The Fuzzy regression uses what is called membership function, which if correctly selected helps to minimize the deviations of the estimated output. The issue with Linear regression is due to its inability to model complex non linear relationships. To resolve the issue of complex non linear modeling the Artificial Neural Networks (ANNs) can be used. The ANNs are inspired by the way the human nerve system process information. There are many types of ANNs, but in this research only the feed-forward [20] and the Radial basis function network are considered. The ANNs consist of processing elements joined together to form the appropriate network with adjustable weighting functions for each network. Typically, there are three or more layers: an input layer, where data is feed to the network, an output layer which is the response to the given input, and one or more intermediate or "hidden" Layers. The operations of an ANNs involve two processes, the learning and the recall. Learning is the process of adjusting the connection weights in response to the stimuli of the input data. Recall is the process of accepting an input and producing a response determined by the learning of the network. There are a lot of learning techniques; The most popular algorithm for training feed-forward neural networks is the Back-propagation (BP) algorithm [8]. As the name suggests, the error computed from the output layer is back-propagated through the network, and the weights are modified according to their contribution to the error function. The BP uses different training algorithms, like Levenberg-Marquardt optimization, Bayesian regularization and others. In this research we have compared each of these algorithms, on the provided data. The BP works with local gradient search, and hence their implementation does not avoid reaching the global minimum, which affects negatively the learning process, in case of noisy data. In our research we have proposed a new model that resolves the limitations of the parametric methods by including the environmental characteristics and the limitation of the linear regression methods by considering complex relation. In this model we have introduced a novel technique to filter noisy data using the fuzzy regression, to improve the learning process for the ANNs. In this research we have also compared different learning algorithms and ANNs methodologies, and selected the most suitable algorithm for the provided data.

II. MODEL VALIDATION

For this research the Mean Magnitude Relative Error (MMRE) is used for estimation accuracy. MMRE is defined as: for n projects

$$MMRE = \frac{1}{n}\sum_{i=1}^{n}\left(\left| Estimated_i - Actual_i \right| / Actual_i \right)$$

MMRE is the preferred error measure of the software measurement researches [9]

III. PROPOSED MODEL

A. Problem formulation

1. Categorical data and environmental characteristics are either not considered in Parametric methods, or are subject to subjective qualitative measures
2. Filtering noisy data without affecting the data quality
3. Select the right ANN model, which is suitable for the data provided.

B. Model components

As shown in figure 1, the proposed model is composed of four building components, the Data Preparation, Categorical transformation, Fuzzy Regression - Data filter, and ANN component. In this section, a complete description of the building components, with clear illustrations on how this model with its building component will be able to address the challenges mentioned in "Problem Formulation" points – 1 through 3.

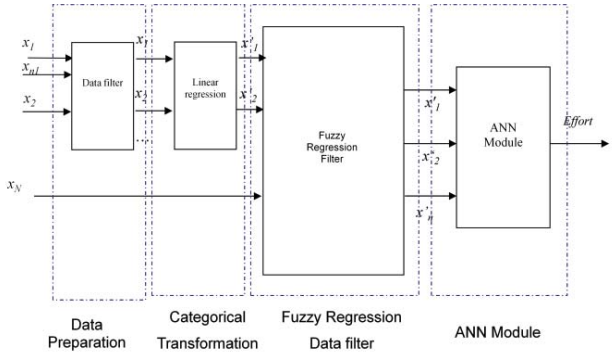


Figure 1: Block Diagram of Fuzzy regression ANN proposed model

1. Data Preparation – Building Component

In our research we have selected data with ISBSG-data [9] quality = "A", and removed missing data, and outliers. Outliers are either visually detected or which has higher *Project delivery rate* (PDR) than the mean PDR. PDR is defined as:

$$\text{PDR} = \frac{Hours}{FP}$$

Where the Hours represent the development effort time and FP represent the function point [17], which is a software sizing measure as the Source Line of Code.

The Mean PDR for Platform = PC is found to be = 3.92631 ~ 4, and for Mainframe = 8.2, and for Midrange =11.04. The mean PDR for the Midrange is h igher than the ratio for the Mainframe, which is not in-line with the industry standards.

The Mean PDR is also calculated for the rest of the attributes and the outliers are removed.

In our research we have used the ISBSG project attributes, as listed in [12]

2. Categorical Transformation – Building component

Projects fuzzy and Categorical attributes such as Development Platform (PC, Mainframe, Midrange), Development Type (new development, enhancement or re-development), cannot be ignored in the software estimation process, unless it is certain that there is no major impact on the effort estimation. It would be meaningless to use a variable in decimal numerical form as a regression predictor because the value of the numbers does not reflect its weight on the estimation process. As illustrated in [21] "Nonmetric variables can only be included in a regression analysis by creating dummy variables. A lot of statistical researches have presented several encoding and transformation techniques [23], [24], [25], in a way that the coding do not affect the regression process, thus the estimating process. In our research, we used binary encoding, and calculated the effect for each categorical attribute by solving the set linear regression equations. We name the attribute effect as categorical multiplier.

2.1 Calculate Categorical Multipliers:
Since the Development Effort Y, is function for both the categorical and numerical attributes, then Y can be presented as follows:

$$Y = f(C, N) \qquad (1)$$

Where:

$C:$ = Categorical attributes

$N:$ = Numerical attributes

Let's consider Y linearly relative to C_i and N_k then Y can be presented as follows:

$$Y = \sum_{Ci \in C} WiCi + \sum_{Ni \in N} WkNk \qquad (2)$$

For each categorical attribute C_i encode it a binary exclusive variables:
Example: the ISBSG-Development Platform, can take the following encoding values PC: [1 0 0], MF: [0 1 0], MR = [0 0 1]

$$Ci = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Using least square method linear regression, then we can the set of equations (2), to calculate the multipliers, using the ISBSG data. The values of W_k are set to 1, and N_k is the value of the numerical project attribute. Bellow tables represent the Multiplier values for different Categorical Values:

Development Platform	
Development Platform	**Multiplier**
MF	2028.83
MR	3156.40
PC	111.79

Table 1, Development Platform multipliers

From the table 1 above it is noticed, that effort to develop software Midrange projects are higher than the Mainframe and the PC, this is shown that through the high value of the multiplier for mainframe. In table 2, it shows that the new development projects take more time than development type projects

Development Type	
Development Type	**Multiplier**
Enhancement	1423.33
New Development	2829.84
New Utility	611.50
Purchased Package	-10553.44
Re-development	1812.74

Table 2, Development Type multipliers

Language Type	
Language Type	**Multiplier**
3GL	2136.85
4GL	1833.71
ApG	1373.17

Table 3, Development Type Platform multipliers

3. Fuzzy Regression Data Filter Building Component

If a variable y is linearly relative to other variables x_1, x_2,... x_N, the regression model of these variables is represented as: [15]

$$Y = b_0 + b_1 x_1 + b_2 x_2 \ldots + b_k x_k \quad (1)$$

Where:

$$S_{11} b_1 + S_{12} b_2 + \ldots + S_{1k} b_k = S_{1y} \quad (2)$$
$$S_{21} b_1 + S_{22} b_2 + \ldots + S_{1k} b_k = S_{2y}$$

-
-

$$S_{11} + S_{12} + \ldots + S_{1k} b_k = S_{1y}$$
$$S_{ij} = \sum u \; cu \; x_i x_j \; - \sum u \; x_i \sum u \; x_j$$

$$b_0 = \frac{\sum uy}{u} - b_1 \frac{\sum ux_1}{u} - b_2 \frac{\sum ux_2}{u} - \ldots - b_k \frac{\sum ux_k}{u}$$

where:

b_0, b_1, b_2, … are K +1 regression parameters, and "ui" is the membership function, and we will use the membership function as in [11]:

$$ui = \frac{1}{\sigma\sqrt{2\pi}} \; e^{-\frac{1}{2}\left(\frac{y_i-\mu}{\sigma}\right)2} \quad (3)$$

Where μ is average of sample points and σ is square root of variance math. The value of ui will be high if the data quality is low, and sample point is in line with the rest of the data points this mitigates challenge 1, described in problem formulation section.

Using Matlab R 2010b, the Fuzzy regression equation was formulated as follows:

Normalized work effort: -1298.1 + 0.354* [Application Type] + 0.156 [Business area Type] + 0.25 *[Development Platform] + 0.342 * [Development type] + 0.233 * [Organization Type] + 0.156 * [Primary Programming Language] – 0.05 * [Project Scope] + 2.2311*[Function points] (4)

From equation 4, it is found that the function point is multiplied by the highest number -"2.311", which is indicative for its high impact, while the project scope is the least importance. It has also concluded that fuzzy regression equation is data specific. i.e. Any small change in the data, due to filtering affect the equation enormously. It is a common method in software engineering to apply regression to transform the data into a scale where the measurements fit more closely to the normal distribution [13]. Transforming the data to exponential equation as in equation 5.

*ln (effort)=ln (a) + b*ln (x_1) + c*ln (x_2) + d*ln (x_3) +… for a multiple regression (5)*

Normalized work effort = -12.87+0.32917 * ln [Application Type] +0.15023 * ln[Business area Type] + 0.195* ln [Development Platform] + 0.094 ln [Development type] +0.73 * ln [Language type] + 0.0283 * ln[Organization Type] + 0.135 *ln [Primary Programming Language] + 0.466 * ln[Project Scope] + 0.765* ln[Function points] (6)

In this research we have used fuzzy regression, to filter noisy data as follows:

Since the fuzzy least squares method is re presented by the following expression [14]:

$$\sum_{i=1}^{n} u_i * (y_i - \hat{y}_i) = min \quad (7)$$

By filtering (removing) high values of u_i, the data quality will improve significantly as shown in the case studies section.

4. ANNs Building Component

Designing the appropriate ANN to achieve the desired mappings, and output is quit an art, since no sound procedure exists for determining either the network structure or the choice of the activation functions, but there are fit of purpose for each ANN design. For example the feed-forward has proven it efficiency in software estimation [18, 19], Radial basis function (RBF) has strong tolerance to input noise [15], while Probabilistic neural network (PNN) [16] has fast training speed as compared to BP. In our research we have examined the BP, using different training functions, and the RBF.

IV. CASES STUDY

As shown in figure 1, the proposed model is a combination of known techniques and/or algorithms that has been used previously in the software estimation techniques\models. In this section, comparisons between each of these techniques are done individually, and also a comparison with the full integrated proposal model is performed.

A. Fuzzy Regression

To study the effect of each the building component, we measured the MMRE after each stage, and compare it with the next stage. We use the following abbreviation in our study of the Fuzzy Regression:

FR: Fuzzy Regression module only
FR – WO: Fuzzy Regression after removing the outliers

Applying equation 4, on the ISBSG data, the MMRE is found to be equal MMRE = 1.3 for FR, and the MMRE = 0.83 for FR- WO, as shown in figure 4, which is the effect of the data preparation module- removing the outliers. Figure 2, plots the relation between Actual and the Predicated effort.

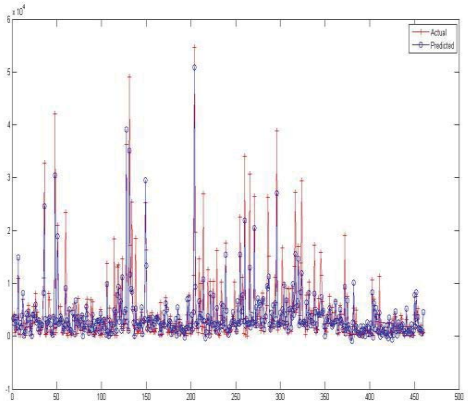


Figure 2: Actual vs. Predicted from Fuzzy Regression module

After applying equation 6, which is the logarithmic form of equation 4, the MMRE has improved to MMRE = 0.766, as shown in figure 4 – LN FR, but still that the fuzzy or the linear regression is not sufficient for the Software estimation process.

B. Feed-forward

In this section, different feed-forward learning techniques are applied, and compared with Fuzzy Regression Technique. Table 4, represents the feed-forward back propagation (BP) propagation.

Configuration:

Item	Value
No. of hidden layers	50
Training functions	Levenberg-Marquardt Back-Propagation, Bayesian Regulation BP
No. of Records (input)	640
No. of records after removing the outliers	460
epochs	200

Table 4, feed-forward Back-Propagation Configuration

Using Levenberg-Marquardt BP (LM BP), The MMRE has increased to MMRE = 1.4, without removing the outliers, and MMRE = 1.2 using Levenberg-Marquardt BP with removing the outliers (LM BP –WO). Applying the Bayesian Regulation BP (BR BP) which minimizes the combination of squared errors, the MMRE has improved to MMRE = 1.04 and by removing the outliers (BR BP WO), the MMRE has improved to MMRE = 0.85. Applying the fuzzy regression filter on the Bayesian Regulation BP, with removed outliers, the MMRE has improved to MMRE = 0.78, which still a relatively high value as shown in figure 3.

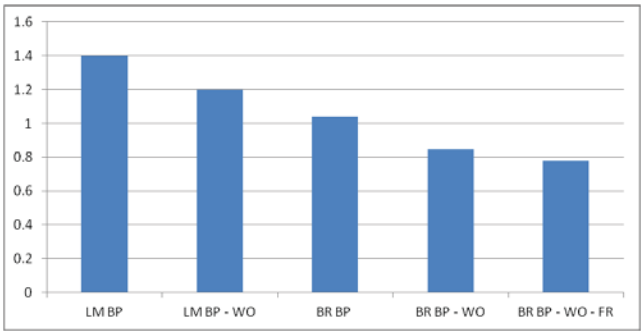


Figure 3: Comparison between different feed-forward BP training functions

The Bayesian regulation BP shows the best results relative to other BP learning functions. The issue is due that the back-propagation generally converge to any local minimum on the error surface [31].To overcome this issue, we tried the Radial Basis Function (RBF), which has not the issue of the Local minimum.

C. Radial Basis Function (RBF)

Table 5, represents the configuration for Radial basis functions. In this study the following abbreviations is used:

GRNN = Generalized regression Neural Network only
GRNN – WO = GRNN with removed outliers
GRNN- FR= Effect of using the Fuzzy regression with GRNN
GRNN-FT-WO= Effect of using the FR & data filer on the GRNN

GRNN

Configuration:

Item	Value
No. of hidden layers	2

Radial basis function type	Generalized regression neural networks (GRNN)
No. of data Records	640
epochs	5

Table 5, Radial basis Configuration

Applying the GRNN without removing the outliers, the MMRE has significantly improved to MMRE = 0.0125, and MMRE = 0.012 without the outliers, and MMRE = 0.008 with fuzzy Regression Filer, and 0.003 with both removing the outliers and the data filter as shown in Figure 4.

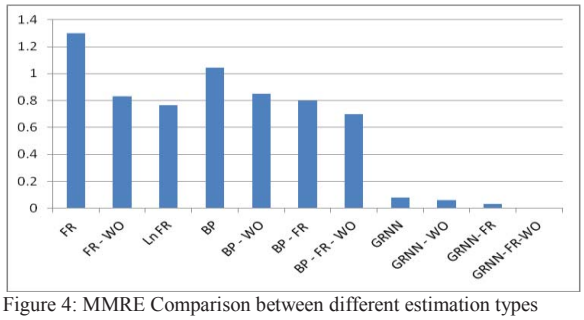


Figure 4: MMRE Comparison between different estimation types

CONCLUTION AND FUTURE WORK

As shown in Figure 3, the Fuzzy & linear regression cannot be considered alone, as good software estimation techniques for ISBSG data or data with complex relations, due to the non-linearity of the estimating variables. Removing the Outliers help to improve the quality of the estimation. Choosing the appropriate ANNs algorithm and training function is important and is dependent on the type of data. The feed-back back-propagation does not produce good results with noisy data, and improves when using data preparation, and fuzzy regression module, thus after improving the data-quality, but still does not give the required level of quality estimation. The Radial basis is the most suitable ANNs algorithm for the ISBSG data, and for noisy data. Also, choosing the right membership function, in the fuzzy regression has improved the data quality, and hence the quality of estimation. In our case the membership function used is the square root of variance math, which helps to filter the noisy data. Also the combined proposed model, gives the highest estimation quality, but it filters the data from 640, to 250 records.

As a continuation for this search, other ANNs algorithm need to be examined, and also comparison of between other estimation techniques that have been using the ISBSG data need to be considered.

REFERENCES

[1] Boehm, B. W., 1981. *Software Engineering Economics,* USA: Printice-Hall *(references)*

[2] Panlilio-Yap, N., 1992. "Software Estimation using the SLIM Tool", in Proceedings of the 1992 conference of the Centre for Advanced Studies on Collaborative research, IBM Press, pp. 439-475.

[3] Shepperd, M., Schofield, C. and Kitchenham, B., 1996. "Effort estimation using analogy", in *Proceedings of the 18th International Conference on Software Engineering*. Washington, DC, USA: IEEE Computer Society, pp. 170-178.

[4] Kitchenham, B.A., Taylor, N. R. Software project development cost estimation. The Journal of Sy stems and Software vol. 5 267- 278, (1985)

[5] S.D. Conte, H.E. Dunsmore, V.Y.Shen. Software Engineering Metrics and Models. The Benjamin/Cummings Publishing Company, Inc. 1986

[6] Nimet Yapıcı Pehlivan, Lean Thinking, International Journal of Lean Thinking Volume 1, Issue 1 (June 2010).

[7] Ishibuchi H., Tanaka H. Fuzzy Regression Analysis Using Neural Networks. Fuzzy Sets and Systems 1992; 50; 257-265.

[8] D.E. Rumelhart, J.L. McClelland, and the PDP Research Group, Parallel Distributed Processing. Cambridge, MA: MIT Press, 1986.

[9] *About ISBSG*, International Software Benchmarking Standards Group, 2004.

[10] B. Boehm, E. Horowitz, R. Madachy, D. Reifer, B. Clark, B Steece, A. Brown, S. Chulani, and C. Abts, *Software Cost Estimation with COCOMO II*, Prentice Hall, 2000.

[11] Venus Marza and Mir Ali Seyyedi, "Fuzzy Multiple Regression Model for Estimating Software Development Time" International Journal of Engineering Business Management, Vol. 1, No. 2 (2009), pp. 79-82

[12] Qin Liu, Wen Zhong Qin, Evaluation of preliminary data analysis framework- Software Quality Journal, 13, 91–115, 2005

[13] N. Fenton, S. Pfleeger, "Software Measurement: A rigorous and practical approach", International Thomson Computer Press, 1996.

[14] X.Gu, G.Song, L.Xiao, "Design of a Fuzzy Decision making Model and Its Application to Software Functional Size Measurement", IEEE, International Conference on Computational Intelligence for Modelling Control and Automation, and International Conference on Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'06), 2006.

[15] Hao Yu; Tiantian Xie; Paszczynski, S.; Wilamowski, *B.M."*Advantages of Radial Basis Function Networks for Dynamic System Design" - IEEE Transactions on Industrial Electronics, Dec 2011

[16] Ting-Chia Ou Cong-Hui Huang Chiung-Hsing Chen Chih-Ming Hong Kai-Hung Lu "Probabilistic neural network and Polynomial Fitting Approach used to determine radio field strength under power lines in radial network" Advanced Intelligent Mechatronics, 2009. AIM 2009. IEEE/ASME International

[17] Albrecht, A.J. and J.R. Gaffney, 'Software function, source lines of code, and development effort prediction: a software science validation', ***IEEE Trans.*** *on Softw.* ***Eng.,*** 9(6), pp639-648, 1983.

[18] Gerhard Wittig, Gavin Finnie "Estimating software development effort with connectionist models". Information and Software Technology (1997)

[19] Srinivasan, K. Fisher, D. "Machine learning approaches to estimating software development effort" IEEE Transactions on Software Engineering Feb 1995, 21 , Issue: 2 . Page(s): 126 - 137

Automating UML Sequence Diagram Generation by Treating it as a Planning Problem

Yaser Sulaiman Moataz Ahmed

Department of Information and Computer Science
King Fahd University of Petroleum and Minerals
Dhahran, Saudi Arabia
E-mail: {g200274900,moataz}@kfupm.edu.sa

Abstract

During requirement elicitation, the consistency of UML (Unified Modeling Language) use cases against the independently-developed class diagrams can be analyzed by trying to develop the sequence diagrams based on those models. But as the complexity of the system being modeled increases, generating the sequence diagrams manually becomes harder. Sequence diagram generation can be automated by treating it as a planning problem and solving it using an automated planning technique. Using such a technique requires expressing goals and actions with their preconditions and postconditions, which is indeed the case when the Design by Contract (DbC) approach is used in developing the models. Based on this similarity, we propose using DbC and planning techniques to automate sequence diagram generation. We illustrate the proposed technique with an online banking system example. With the increasing support for DbC in modeling tools and programming frameworks, the technique, once realized, should help in improving the software development process by enabling automatic consistency analysis of use cases against class diagrams through automatic sequence diagram generation.

1 Introduction

In UML (Unified Modeling Language), a *de facto* standard for object-oriented modeling, different models are developed to represent different views of the software system being modeled. In particular, use cases model a *functional* view by capturing the externally visible services of the system, while class diagrams model a *structural* view by expressing its static structure. Acting as a bridge that connects use cases and the class diagram, sequence diagrams represent the dynamic interactions within the system. Of course, this multi-view approach of UML can lead to inconsistencies between the different models, especially if they are developed independently, but detecting such inconsistencies is actually an objective of this approach. This can be done by cross-referencing the different models against each other and looking for system's aspects that were captured in some views but not in others. More specifically, the consistency of the use cases against the independently-developed class diagram can be analyzed by trying to create the sequence diagrams based on those models. But creating sequence diagrams manually is an error-prone process. This is supported by the results of the experiments by Yue, Briand, and Labiche [12]: trained fourth year undergraduate students failed to create 50% of the required messages, and out of the created messages, 25% were inconsistent with the reference diagrams.

If the use cases and the class diagrams are developed using Design by Contract (DbC), where the relationship between a routine and its callers is semantically specified using preconditions and postconditions, a striking similarity arises between creating a sequence diagram and solving a planning problem. That is, when DbC is used, and despite some differences in terminology, the same basic building blocks are present in both cases. This similarity looks promising and we found it too remarkable to be ignored.

The support for DbC in Computer-Aided Software Engineering (CASE) tools and programming frameworks is increasing. For example, Enterprise Architect 8 allows the user to specify preconditions and postconditions for individual methods in a class diagram, albeit as a somewhat simplistic supplement to the method's description. Also, support for code contracts is now a core feature of the .NET Framework 4 [1]. These observations suggest that more and more software developers and professionals are adopting DbC as they realize its practical benefits.

2 Related Work

It appears that none of the previous related work tackled the problem of generating sequence diagrams automatically as a planning one. Instead, previous approaches used case templates, restrictions on natural language, and transformation rules to automatically translate use cases to sequence diagrams. This section highlights some of the recent and most relevant previous work.

The proposed technique can be viewed as being at the intersection between software engineering and automated planning. A related research area is knowledge engineering for planning and scheduling [8]. An example of work in this area is that of Vaquero et al. [7], in which they report their research efforts to develop itSIMPLE, an integrated design environment for automated planning applications. They view the design process of a planning application project as a series of phases, such as requirements specification, model analysis, plan synthesis and analysis, where each phase requires a different knowledge representation. In itSIMPLE, requirements are first modeled using UML (Unified Modeling Language). Then, these UML models are analyzed using Petri nets. Finally, these models are automatically translated to a Planning Domain Definition Language (PDDL) representation and are presented to an automated planner.

Yue, Briand, and Labiche [11] present a systematic review that focuses on transforming textual requirements to analysis models in the context of model-driven development. They review 16 approaches and evaluate their capabilities, support for establishing traceability links, degree of automation, efficiency, and completeness. To do that, they design a conceptual framework and derive a set of evaluation criteria from it. Based on the limitations of the existing approaches, the authors suggest a pattern for future ones that consists of using reasonable restrictions on natural language, an automatic requirements preprocessing technique, and one intermediate model to transform a use case model into a UML model that compromises structural and behavioral aspects.

Yue, Briand, and Labiche [12] follow their own suggestion and propose a technique to automatically derive analysis models, including sequence diagrams, from use cases while maintaining traceability links. Requirements must first be defined manually using the Restricted Use Case Modeling (RUCM) approach, which the authors proposed in an earlier work [10]. The result of RUCM is a textual use case model (UCModel) that is expressed in a restricted natural language. aToucan transforms UCModel into an intermediate model (UCMeta), which is then transformed into the desired analysis model. At the same time, aToucan establishes traceability links between UCModel and the generated analysis model. An evaluation framework was devised to compare the sequence diagrams generated automatically by aToucan to ones generated manually by experts and undergraduate students. The empirical study shows that the automatically generated diagrams were very complete and consistent with the ones generated by experts and that they were more complete than the ones generated by students.

Liwu Li [2] presents a parser that translates a manually normalized use case to message records which may then be used to construct sequence diagrams. Li provides four guidelines for use case normalization and presents an automated teller machine use case as an example. Given a normalized use case, the parser identifies syntactic structures to deduce static information including classes, objects, attributes, and operations, and dynamic information including message sends. The user may need to manually instruct the parser on how to translate some sentences. This, coupled with the fact that the normalization step is manual, makes the approach semi-automatic.

Out of the works discussed above, there are two that appear to be quite similar to the proposed one. The first is that of Vaquero et al. [7]. Their work plays a central part in our proposed technique, but the difference between the two research efforts is subtle, yet crucial. Vaquero et al. are applying requirements engineering and knowledge engineering to planning, while we are applying automated planning to requirements engineering. In other words, for Vaquero et al., planning is the domain and the generated plans are the ultimate output; for us, planning is the tool and the generated plans are a step towards the ultimate outputs (sequence diagrams).

The second work that appear to be similar to the proposed one is that of Yue, Briand, and Labiche [12]. On further inspection though, one can see that in their approach, generating a class diagram is a prerequisite for generating the sequence diagram. This means that generating sequence diagrams must follow a linear path that starts with use cases and passes through class diagrams. With such a linear path, detecting inconsistencies by cross-referencing can not be carried out since any defect that is included in the use cases will be carried through to the subsequently generated models. Of course, the traceability links will help in tracing and fixing defects and inconsistencies, but they will not help in detecting them. In contrast, the proposed technique will generate sequence diagrams automatically from use cases and class diagrams that were developed *independently*. This will enable consistency analysis by cross-referencing the different models.

To summarize, the main difference between the proposed technique and the ones presented in previous related work is that, assuming the use cases and the class diagram were developed independently using Design by Contract (DbC), it solves the problem of automatic sequence diagram gener-

ation by tackling it as a planning problem.

3 Design by Contract

Bertrand Meyer [3] coined the term Design by Contract (DbC). In DbC, routines are semantically specified using preconditions and postconditions. Preconditions are the conditions that must be ensured before calling the routine if it is to function properly; postconditions specify the state that the routine guarantees yielding assuming that it was called with the preconditions satisfied. The preconditions and postconditions of a routine define a *contract* that binds it and its callers. In this contract, the routine is the supplier of a service and its callers are the clients. For the supplier and its clients, the contract defines obligations and benefits: the preconditions are benefits for the routine and obligations for the callers; the postconditions are obligations for the routine and benefits for the callers.

As an example, consider a square root routine that operates on and produce real numbers. This routine expects its input to be nonnegative. Adopting the notation Meyer uses in his book [3], in which he uses a **`require`** clause to express preconditions and an **`ensure`** clause to express postconditions, this constraint can be expressed as a precondition of the routine as follows:

```
sqrt(x: REAL): REAL
  require
    x >= 0
  do
    ...
  end
```

For another example, consider the `pop()` routine of a stack class that can store objects of type `T` and that has a limited capacity. The class developer may specify that this routine should not be called on an empty stack. This constraint constitutes the precondition of the routine. If the routine is called on a nonempty stack, it will remove the top element. By doing so, it will decrease the number of elements by one and ensure that the stack is not full even if it was so before the call. These guarantees constitute the postconditions of the routine. Adopting the notation of Meyer's again, these constraints and guarantees can be expressed as preconditions and postconditions as follows:

```
pop(): T
  require
    not empty
  do
    ...
  ensure
    not full
    count = old count - 1
  end
```

DbC can be applied to UML (Unified Modeling Language) models by using the Object Constraint Language (OCL) [9]. OCL is a specification language for UML, not an action language. It is mainly used to write queries to access model elements and their values and state constraints on model elements. UML model elements are annotated with OCL constraints to ensure their proper usage and validity of the whole model. An OCL constraint cannot change the value of a model element and hence is considered side-effect free. It can be used to express preconditions, postconditions, invariants, guard conditions and results of operations.

An OCL constraint typically consists of two parts: the context and a set of OCL expressions. As an OCL constraint highly depends upon which model element is constrained, this information is specified by the first part of the OCL constraint, that is, its context. An OCL context can either be a class, one of its attributes, or one of its operations, and it can be referenced in the constraint's body by the keyword `self`.

The second part of an OCL constraint consists of a set of OCL expressions, each of which consists of a type, name and body. The frequently used expression types include `inv`, which is used when the body contains a condition that must be met by all instances of a class; `pre`, which is used when the body contains a condition that must be true before executing an operation, i.e. a precondition; and `post`, which is used when the body contains a condition that states what should be true about the state of the system and the changes that occurred after executing an operation, i.e. a postcondition. Since OCL is a query language, it expects a result when querying a property or an operation of the class in context. OCL uses the "`.`" operator when it expects a single-valued result. Multiple expressions in an expression body can be combined by using the boolean operators `and`, `or`, `xor`, `not` and `implies`.

As an example, consider the `pop()` routine used above. Assuming that the stack class has the two boolean routines `isEmpty()` and `isFull()`, along with the integer attribute `count`, the preconditions and postconditions of the `pop()` routine can be expressed in OCL as follows:

```
context Stack::pop(): T
  pre: not self.isEmpty()
  post: not self.isFull() and
    self.count = self.count@pre - 1
```

4 Sequence Diagram Generation as a Planning Problem

As Figure 1 illustrates, solving a planning problem [5] means finding a sequence of actions that, starting from a particular state, will achieve a certain goal. To do that, each

Figure 1. Solving a planning problem consists of finding a sequence of actions $a_1, a_2, a_3, \ldots, a_g$ that, starting from an initial state S_0, will achieve a certain goal S_g. Each action a_i changes the state to an intermediate one S_i until finally a_g achieves the desired goal S_g.

Automated Planning		Design by Contract
Initial State	$\Leftrightarrow$	Use Case Preconditions
Goal	$\Leftrightarrow$	Use Case Postconditions
Actions	$\Leftrightarrow$	Methods
Action Preconditions	$\Leftrightarrow$	Method Preconditions
Action Postconditions	$\Leftrightarrow$	Method Postconditions

Table 1. Despite some differences in terminology, automated planning and Design by Contract have the same basic building blocks. This correspondence enables us to treat sequence diagram generation as a planning problem and to solve it accordingly.

action should be specified using its preconditions, which must hold before the action can be executed, and its effects.

When Design by Contract (DbC) is used to develop the use cases and the class diagram, generating the sequence diagrams based on them starts to look very similar to solving a planning problem. In both cases, and despite some differences in terminology, the same building blocks are present. For example, the initial state of a planning problem corresponds to the preconditions of a use case. The correspondence between automated planning and DbC, which Table 1 demonstrates further, opens the door to treating sequence diagram generation as a planning problem.

The proposed technique will inherit limitations from Design by Contract, formal methods, and planning techniques.

Despite its elegance and practical benefits, Design by Contract was, arguably, not widely adopted as it should have been. This can be attributed to two factors. The first is the paucity of skillful software developers who were willing to apply it in their own work. The second is the limited built-in support in many CASE tools and programming frameworks, or the lack thereof, which may have reduced the practical scalability of Design by Contract. Nevertheless, the support for Design by Contract in CASE tools and programming frameworks is on the rise: Enterprise Architect 8 allows the user to specify preconditions and postconditions for individual methods in a class diagram, and the .NET Framework 4 supports code contracts as a core feature [1].

The proposed approach requires the specifications to be expressed formally. Formal methods and specifications help in detecting errors and inconsistencies in informal requirement specifications, but they have limitations of their own. Ian Sommerville [6, pp. 218-219] identifies a number of factors that contribute to the lack of wide adoption of formal methods. One factor is the limited scope of formal methods: it is difficult to use them to specify user interfaces and interactions. Another factor is their limited scalability. The time and effort required to develop a formal specification of a system increase significantly as the size of the system increases. In other words, formal methods are really cost-effective when the cost of failure is really high.

5 Illustrative Example

To demonstrate the proposed technique, we use an online banking system as an example. To avoid overcomplicating the presentation of the example, which serves as a proof of concept, we will keep it simple and limit the discussion to the most relevant details.

Suppose that we want the system to enable a bank user to log in and out of the system, to activate or deactivate his bank account, and to withdraw or deposit money from his account. Suppose further that the user can withdraw a certain amount of money from his account only if he is logged-in, the account is active, and the amount is less than or equal to the account's balance. If all of these conditions are met, the given amount will be deducted from the balance.

Figure 2 shows one possible class diagram for such a system. The class model consists of two main classes: `UserProfile`, which holds information about the user and is responsible for user authentication, and `Account`, which holds information about the user's bank account and is responsible for account activation and financial transactions.

As a sample use case of the system, suppose that the user is initially logged-out of the system and has a deactivated bank account with $5000 in it. Suppose further that the user wants to withdraw $500 from the account and then log out from the system.

We used itSIMPLE [7], which is a knowledge engineering tool for designing planning domain models, to model the use case and the classes outlined above. We added the necessary preconditions and postconditions as Object Constraint Language (OCL) expressions to the methods of `UserProfile` and `Account`. For instance, Listing 1 shows the OCL specification of the `withdraw()` method of the `Account` class.

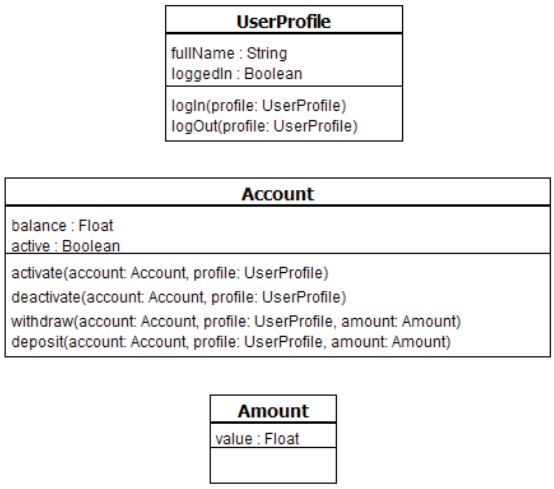


Figure 2. One possible class model for the bank example consists of two main classes: `UserProfile` and `Account`. In itSIMPLE, we had to add a simple wrapper class, `Amount`, to be able to specify the desired use case.

```
context Account::withdraw(account:
    Account, profile: UserProfile,
    amount: Amount)
  pre: profile.loggedIn = true and
     account.active = true and
     account.balance >= amount.value
  post: account.balance =
     account.balance@pre - amount.value
```

Listing 1. The specification of `withdraw` expressed in OCL.

```
(:action withdraw
 :parameters (?account - Account ?
    profile - UserProfile ?amount -
    Amount)
 :precondition
   (and
     (loggedIn ?profile)
     (active ?account)
     (>= (balance ?account) (value ?
        amount))
   )
 :effect
   (decrease (balance ?account) (value
      ?amount))
)
```

Listing 2. The specification of `withdraw` expressed in PDDL (as translated by itSIMPLE).

```
0: (LOGIN PROFILE) [1]
1: (ACTIVATE ACCOUNT PROFILE) [1]
2: (WITHDRAW ACCOUNT PROFILE AMOUNT)
   [1]
3: (LOGOUT PROFILE) [1]
```

Listing 3. The plan for the withdraw use case as reported by itSIMPLE.

After modeling the use case and the classes, we used itSIMPLE to translate the models along with their constraints into planning domain and problem descriptions expressed in the Planning Domain Definition Language (PDDL). Listing 2 shows the result of the translation that itSIMPLE performed on the specification given in Listing 1.

After obtaining the planning domain and problem corresponding to the modeled use case and classes, we used the planners bundled with itSIMPLE to generate the plan that solves the planning problem. The planners that were able to solve the problem generated the same plan, which consists of logging in, activating the account, withdrawing the specified amount, then finally logging out. Listing 3 depicts the plan as reported by itSIMPLE. This plan corresponds to a sequence diagram that realizes the withdraw use case. Figure 3 shows this sequence diagram.

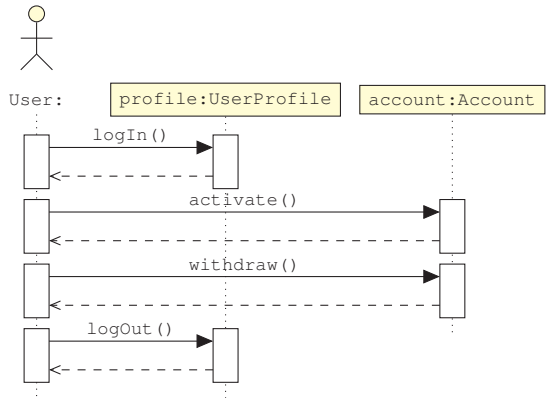


Figure 3. This sequence diagram realizes the withdraw use case. The diagram corresponds to the plan that was reported by it-SIMPLE.

6 Conclusions

With the increasing support for DbC in modeling tools and programming frameworks, the proposed technique, once realized, should help in improving the software development process by enabling automatic consistency analysis of use cases against class diagrams through automatic sequence diagram generation. We are currently working on using the proposed technique to automate the consistency analysis of use cases against class diagrams.

Our immediate next step is to validate the proposed approach. To do that, the sequence diagrams that were automatically generated must be compared to ones that were manually generated. This requires using an evaluation framework, which should consist of quality characteristics that can be used to compare sequence diagrams. An example of such an evaluation framework is the one used by Yue, Briand, and Labiche [12]. This framework uses quality characteristics such as sequence diagram completeness, boundary-control-entity consistency, and sequence diagram redundancy. These characteristics are defined based on simpler measures collected from the sequence diagrams, such as the number of messages, lifelines, and interaction uses in the given diagram.

Based on the results of the future empirical validation, we ultimately plan to implement the proposed technique as a plug-in to a Computer-Aided Software Engineering (CASE) tool. To ease the integration with existing CASE tools, we plan to use the XML Metadata Interchange (XMI) [4], which is an international standard for sharing metadata and models—especially UML models—using XML (Extensible Markup Language), as the format of the initial inputs and final outputs.

7 Acknowledgment

The authors would like to acknowledge the support provided by the Deanship of Scientific Research at King Fahd University of Petroleum and Minerals (KFUPM) under Research Grant IN111013.

References

[1] What's New in the .NET Framework 4. http://msdn.microsoft.com/en-us/library/ms171868.aspx. Accessed 7 May 2011.

[2] L. Li. Translating use cases to sequence diagrams. In *Proceedings ASE 2000. Fifteenth IEEE International Conference on Automated Software Engineering*, pages 293–296. IEEE Comput. Soc, 2000.

[3] B. Meyer. *Object-Oriented Software Construction*. Prentice Hall PRT, 2nd edition, 1997.

[4] Object Management Group. XML Metadata Interchange Specification Version 2.0.1, 2005.

[5] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Pearson Education International, 2nd edition, 2003.

[6] I. Sommerville. *Software Engineering*. Pearson Education Limited, 7th edition, 2004.

[7] T. S. Vaquero, J. R. Silva, F. Tonidandel, and J. C. Beck. itSIMPLE: Towards an Integrated Design System for Real Planning Applications. *The Knowledge Engineering Review*, pages 1–15, 2011.

[8] T. S. Vaquero, R. Silva, and J. C. Beck. A Brief Review of Tools and Methods for Knowledge Engineering for Planning & Scheduling. In R. Barták, S. Fratini, L. McCluskey, and T. S. Vaquero, editors, *Proceedings of the ICAPS2011 Workshop on Knowledge Engineering for Planning and Scheduling*, pages 7–14, Freiburg, Germany, 2011.

[9] J. Warmer and A. Kleppe. *The Object Constraint Language: Getting Your Models Ready for MDA*. Addison Wesley, 2nd edition, 2003.

[10] T. Yue, L. Briand, and Y. Labiche. A Use Case Modeling Approach to Facilitate the Transition towards Analysis Models: Concepts and Empirical Evaluation. In A. Schürr and B. Selic, editors, *Model Driven Engineering Languages and Systems*, volume 5795 of *Lecture Notes in Computer Science*, pages 484–498. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[11] T. Yue, L. C. Briand, and Y. Labiche. A systematic review of transformation approaches between user requirements and analysis models. *Requirements Engineering*, Aug. 2010.

[12] T. Yue, L. C. Briand, and Y. Labiche. Automatically Deriving UML Sequence Diagrams from Use Cases. Technical report, Simula Research Laboratory, 2010.

Knowledge Management to Support the Use of Agile Methodologies

Jordão E.[1] *and Furtado F.*[1,2]
[1.] C.E.S.A.R. – Recife Center for Advanced Studies and Systems
[2.] UFPE – Federal University of Pernambuco, CIn – Informatics Center
Recife, Brazil
{edson.jordao, felipe.furtado}@cesar.org.br

***Abstract*—There are evidences that the flexibility brought by the agile manifesto has caused problems to the knowledge construction and dissemination. As the manifesto doesn't emphasize documentation and record of events and facts, knowledge storage and dissemination in general are not performed, making its use and recovery too hard. The use of processes and mechanisms of knowledge management (KM) are essential for an organization that aims to meet time-to-market and keep itself competitive in the market. This article presents evidence of how the KM can be used in a context of agile development.**

Keywords-Knowledge Management; Scrum; Software Development Process;

I. INTRODUCTION

Software development, although not the end area in most companies, is intimately intertwined with a company's business activities, and generates added value to what matters most to organizations. It is also one of the areas that incurs the greatest costs in investment, which makes the area of information technology a very large asset and one that is essential for doing business. Moreover, for an organization to differentiate itself from its competitors, its key business rules need to be implemented in an agile way in its software so that the company establishes itself in the market ahead of others [1]. Control over costs and software development may well be achieved with the processes of Software Engineering, one of the objectives of which is to transform the creation of software systems from being done in an artistic, undisciplined and barely understandable way to doing so in an appropriately controlled, measured and predictable way by using procedures, methods and tools [2].

However, software engineering only took heed of agility after the agile manifesto became well known. The values and principles of this manifesto are focused more on individuals and their interactions, software that works, collaborating with the client and responding to changes. This does not mean that the manifesto did not give value to processes and tools, comprehensive documentation, contract negotiations and planning, but these are not as highly valued as the items first mentioned [3]. The agile manifesto came to prominence in the attempt to give more agility to and prioritize whatever generates most value when developing a piece of software.

However, there are indications that this agility caused some problems with regard to constructing knowledge. Since documentation is not emphasized, it can hardly be said that everything that is generated during the development process is recorded and stored, thus making it difficult to recover and spread knowledge. In order to reuse historically acquired knowledge, great effort and a significant amount of time is needed which could be avoided if knowledge were stored somewhere.

Therefore, to meet time-to-market and to give agility to the strategic development of the business rules of organizations so that knowledge generated is stored and later retrieved, well-defined processes of knowledge management (KM) should be used, thus making the exchange and spread of knowledge, an item of organizational culture.

This article has six sections. After this introduction, there is a review of the literature on the Management of Knowledge, which is followed by Sections on the Agile and Scrum Methodologies. Thereafter, a relationship between the use of knowledge management and Scrum is presented and how this can help and improve the use of agile methodology. Finally, conclusions are drawn from what has been discussed and proposals for future studies made.

The objective of this study is to explore the state of the art on the use of knowledge management (KM) in agile methodologies in order to solve the problems presented in this introduction.

II. KNOWLEDGE MANAGEMENT OVERVIEW

The most precious thing an organization can have is its knowledge. This asset is what projects it into a prosperous and long-lasting future, and into a dynamic market, in which technologies multiply and expand, and products become obsolete in the short term. Only companies that are prepared will become competitive and long-lasting organizations in an economy in which, according to Nonaka and Takeuchi [4], the only certainty is uncertainty.

According to Bose [5], this can be proven with some data taken from surveys as shown below:

1) A total of 80 per cent of the 500 largest companies have a KM team.

2) Texas Instruments has saved $1 billion since it began a KM program in the mid-1990s.

3) 95 per cent of CEOs interviewed at the World Economic Forum in Davos, Switzerland, in 2001, said that KM was essential for organizational success.

4) 91 per cent of the leaders of the Canadian market, surveyed by IpsosReid in 2001, believe that the KM practices have a direct impact on organizational effectiveness.

The production of new knowledge within organizations involves a process that amplifies the knowledge created by individuals and crystallizes it as part of the organization's knowledge network. The importance of individuals as to generating knowledge in the organization lies in the fact that only they can create knowledge. This knowledge has been classified by Polanyi as tacit and explicit [6]. The latter is characterized by knowledge that can be recorded, documented and distributed, while it is difficult to formalize and distribute tacit knowledge. The great difficulty of organizations is to amplify this knowledge given that they are dependent on individuals and therefore cannot on their own create knowledge. It is important that they create favorable and appropriate contexts and encourage knowledge-creating activities [4]. It is within this perspective that the authors, in [7], claim that the capture, learning and knowledge transfer processes in project settings are heavily based on standards, practices and social processes.

Due to the importance, wealth and magnitude of knowledge in organizations and, in addition, due to the uncertainties, inconstant behavior and economic dynamism of the market, it is of the utmost necessity for organizations to create processes for managing their knowledge [4].

These processes need to be aligned to the organization's culture, as well as to their business processes, because KM requires characteristics of the organizational context, such as processes, interpersonal connections, document repositories, as well as cultural and institutional norms to coordinate KM [8].

Organizations also need to balance what kind of knowledge they will be focusing on and based on this, the type can be categorized into one of four styles of KM defined by Choi and Lee [9]:

1) Passive: Little interest in KM. Knowledge is not managed in a systematic way.

2) Systems-oriented: a greater emphasis on codifying and reusing knowledge. This increases the ability to encode IT, and, therefore, diminishes the complexity of accessing and using knowledge.

3) People-oriented: emphasis is given to acquiring and sharing tacit knowledge, and interpersonal experiences. Knowledge usually originates through informal social networks. This means that knowledge cannot simply be retrieved from a database or repository.

4) Dynamic: emphasis is given both to explicit and tacit knowledge and does so in a dynamic fashion, similar to an organization in which communication is intensive. However this depends on a knowledge culture.

For Choi and Lee [9] the dynamic style shows better results in organizational performance. The authors state that the two types of knowledge should be integrated and interact with each other. Yet, the emphasis on each type of knowledge varies with the needs and culture of organizations.

Regardless of the style of knowledge management, KM can be understood as a process that involves the creation, capture, storage, transfer and application of knowledge [10]. There are several definitions for each of these stages that are part of the KM process. In 2004, Bose introduced the stages that make up a cyclical process in knowledge management [5]:

1) Creating Knowledge: knowledge comes first and foremost from collaborators' experience and prior knowledge. Knowledge is created to the extent that people find new ways of doing things. Sometimes it is necessary to bring in a piece of external knowledge when this is not found in the organization.

2) Capturing Knowledge: Knowledge created is stored in a database or repository.

3) Refining Knowledge: context is added to knowledge, and thus can be easily reused. At this stage, tacit knowledge, usually from human experience, is captured, processed and refined along with explicit knowledge.

4) Storing Knowledge: This stage includes the codification of tacit and explicit knowledge, after both forms have been refined. Thus both can be used later.

5) Managing Knowledge: knowledge must be kept current, so it should be reviewed systematically to check if it remains valid and accurate.

6) Disseminating Knowledge: knowledge is available, so everyone in the organization can easily access it anytime and anywhere. New technologies are normally used to facilitate the dissemination of knowledge.

Some of the various benefits of using a knowledge management strategy are: to reduce the loss of the intellectual capital of collaborators who leave the organization; to reduce the costs of developing new products; and to increase productivity by making knowledge easily accessible to all collaborators [5].

III. AGILE METHODOLOGIES

According to Fowler, agile methodologies are collections of software methodologies, which follow the same principle of development processes, in which conceptual structures are created to design a piece of efficient and predictable software. The concept of agile methods is different from that of traditional methodologies and differs in having shorter iterations, thus simplifying the process, is less focused on documentation and is also characterized as an antidote to bureaucracy [11].

For Fowler, agile methodologies should not be analyzed from the standpoint of their slenderness, but in their adaptive

nature and their tendency to put people first. To achieve this agility there must be effectiveness and efficiency, thus creating a balance between having no processes and having a lot of them, by providing just enough process to obtain a reasonable return [11].

And for Highsmith [12] the concepts and assumptions behind the defined and empirical processes are fundamentally and irredeemably different. For him, agile practices of software development such as short iterations, continuous tests, self-managing teams, constant collaboration and frequent re-planning, based on current reality, are all geared to understanding software development as an empirical process.

Agile development is not defined by a small set of practices and techniques, but rather by the strategic ability to create and respond to changes, to balance flexibility and structure, to extract creativity and innovation from a development team, and by the ability to lead organizations through an uncertain and unstable market [12].

The proof that agile methodologies are not a failure is the fact that over half of respondents in the annual survey which is conducted by VersionOne, the last being in 2011, responded that they have been using use agile practices for two years and a third took agile methods with them to other companies. Almost two thirds of respondents said that more than half of the projects in their organizations use agile methods and that their companies have adopted agile practices and have three or more teams for this [13].

While concern about agile scalability, its regulatory compliance and its lack of documentation arose, there were fewer barriers from management, given that implementation is agile. Furthermore, 64% of initial agile supporters are in the management layer. And of the main reasons for adopting agile practices, the one that tops the charts with 22%, is accelerating the time-to-market [13].

Despite the small increase in management support for the use of agile methods, the survey results suggest that the greatest barrier to increasing the adoption of agile methodologies seems not to be lack of knowledge of the methodology, but the internal culture of organizations. Only 13% of respondents in large organizations (more than 500 collaborators), said that almost all of their projects are agile. In these large companies, respondents said that lack of management support (27%) and "general resistance to change" (26%) were the main barriers to adopting an agile approach. On the other hand, among small businesses, the vast majority (almost 75%) of projects use agile methods and only 10% of respondents cited lack of management support or "general resistance to change" as a problem as to adoption [14].

As to agile practices, research has shown that quick daily meetings (78%), iteration planning (74%) and unit testing (74%) are continuous practices. The most noticeable trend this year was the increased use of the principles of Kanban (24% vs. 18% in 2010). By far, Scrum/ XP Scrum continue as the front-runners and so are the most widely used agile methodologies (76%) [13].

This year there was a change in the benefits achieved from agile methods in companies. In 2010, respondents said that productivity was the greatest benefit (74%), but in 2011, productivity improved for only 55% of the respondents. In 2011, respondents said that the best results brought about by agile practices were project visibility (77%) and the ability to manage the client's changing priorities (84%). Reducing costs and the ability to m anage distributed teams were also mentioned as benefits. Overall, three quarters of respondents said that at least half of their agile projects have been successful [13].

In 2011 research showed that there was a small increase in the number of companies that currently do not use agile methodologies, but are interested in using them in the future (17% this year against 13% last year). Of those who already use agile methods, a third will continue to use them and only 3% said they do not plan to continue using them [13].

The most common types of tools being used (or which it is planned to use) include: Acceptance tools (20%), tools for managing releases (16%) and tools for continuous integration (16%). Interestingly, the tool that is being used least today is also the one that it is most desired to use in the future. This is the tool for managing ideas [13].

In terms of specific, agile software tools, most use Office productivity standards such as Excel, for example, followed by specialized tools such as VersionOne. The use of bug trackers has also been heavily used this year (14%) [13].

IV. SCRUM METHODOLOGY

According to Schwaber [14], Scrum is a methodology for improving the management and maintenance of an existing system or a production prototype. In Scrum, each release is planned based on the following variables:

1) The client's requirements: how the current system needs to be improved.
2) Time pressure: the time needed to gain a competitive advantage.
3) Competition: What is the competition, and what is needed to be better than them.
4) Quality: What is the required quality, given the above variables.
5) Vision: what changes are needed at this stage to satisfy in full the vision of the system.
6) Resources: the personnel and financial resources that are available.

These variables form the original plan of a project to improve software. However, these variables may change during the project. A successful development methodology must take these variables and their evolutionary character into account [14].

But the success of Scrum depends not only on these variables, but also on how the organization adapts to Scrum. Many organizations have decided to use Scrum because of the vast success that many companies are having or because they have seen a p roject succeed which has used the agile methodology. But, according to Cohn [15], many organizations

have tried to use it and have not succeeded for the following reasons:

1) A successful change is not entirely top-down or bottom up: change within an organization needs to be done from the bottom up with the support of management. There should be a point of equilibrium at which the change should begin. More than wanting to use an agile development process, it is necessary to change the thinking of the entire organization, and this ranges from those in marketing and the commercial sectors to those working in the development area.

2) The final state is unpredictable: To think that the use of a methodology such as Scrum, fits seamlessly into one's organization, in the way it is laid down in writing, is wrong. It will always be necessary to adapt the methodology to one's organization. To do so there is a need to make changes slowly, to test the changes and to identify the improvements and advances needed.

3) Scrum casts a wide net: The changes made by Scrum within the organization will affect not only the software development department but also various areas of the organization, besides how contracts are undertaken and how commitments are made to the scope and allocation in negotiations. Therefore it is much easier to meet with resistance when implementing Scrum.

4) Scrum is very different: another point of great concern in implementing Scrum is to adapt the team so they can use it. Like any other agile methodology, Scrum is very different from what members of a development team are accustomed to.

5) The change is faster than ever: A lot of change in a short space of time makes people stressed and overloaded. Then there are the changes in technologies, tools, languages and platforms so it is not surprising if people get stressed and do not accept the changes caused by Scrum.

6) The best practices are dangerous: Whenever the best and most correct way to do something is found, the organization adopts it and it is shared by others as "best practice". For some types of work, collecting and reusing best practices helps a lot in making the effort to make changes. However, in the transition to Scrum, adopting best practices can be dangerous. This is because the best practices can lead us to relaxing which interrupts the effort to make continuous improvements which is essential to Scrum.

Despite the difficulties mentioned above, the advantages and the results brought by Scrum are much greater and this can be shown in surveys already conducted and some of those cited earlier. Cohn [15] conducted research by making a rigorous comparison of 26 agile development projects and a database of 7,500 development projects, most of which were traditional ones.

In addition, Cohn [15] used several published academic articles, an online survey of more than three thousand people and also used his extensive experience. The result of this research and this study was that there are six reasons that lead people and organizations to make the transition to Scrum: (1) increased productivity and lower costs, (2) staff feel greater engagement and job satisfaction, (3) the time-to-market is faster, (4) higher quality, (5) increased stakeholder satisfaction, and (6) what we are doing no longer works.

In his comparison between agile projects and traditional projects, Mah [16] found that agile projects are 16% more productive, an increase he considered was statistically significant. The fact that a functional product is delivered faster to the customer makes staff feel more motivated and engages them even more in the project, which makes productivity even greater. As they receive something that works sooner, clients succeed in meeting their time-to-market target and positioning themselves more strongly in the market and thus serving their clients' needs more quickly and efficiently.

If they add to these points, those of delivering faster and being more productive, the team can deliver a product of higher quality because it can organize itself to achieve continuous improvement and deliver a product that meets the needs of stakeholders. And by having their needs met more quickly, achieving time-to-market efficiently, clients and other stakeholders become more satisfied, thus further contributing to the projects [15].

Scrum has a set of management rules and practices that should be adopted if a project that uses the methodology is to be successful. For Schwaber and Beedle [17], Scrum is actually not a methodology but rather a framework. That is, it does not detail how to do something, but what to do. Its practices will be highlighted and how they work, in the sub-sections below.

A. Product Backlog

This is defined as all the tasks and activities needed to get the final product developed. It is the starting point for a Scrum project. The definition of backlog items can be made by any person or sector involved in the project. For any item to be added, the entire team should agree to it. A backlog item should have a description, an estimate in hours, a person-in – charge and a priority [17].

B. Sprint

This is where the items defined in the product backlog are performed. This should have a maximum duration of 30 days. To start a sprint, it is necessary to choose within the items within the product backlog which have the highest priority should be chosen and those that can be developed in the time period during which the sprint will take place, which thus creates the sprint backlog. No processes are defined within a sprint, but while it is taking place, some practices must be used, such as daily meetings to monitor existing activities, integrating software parts, development tests, and formal testing cycles [17]. Abrahamsson [18] suggests that there should be traditional phases within a sprint such as analysis, design, implementation, testing and delivery.

C. Sprint Planning Meeting

Each sprint starts with a meeting that aims to analyze the product backlog items in order to prioritize them for development and as the final result of the meeting to obtain the sprint backlog [17]. At this meeting the product owner explains each item prioritized for the sprint so that the team may obtain a full understanding of what should be done and that adds most value to this. At the end of this meeting, the team can meet to discuss what was planned and items in the sprint backlog.

D. Daily Scrum Meeting

Daily, preferably at the same time early in the day, the team gathers for a quick meeting where each member should participate basically by saying what was done the day before, what should be done today and if there is any obstacle to an activity being performed and the reason for this. This meeting should not be used for technical discussions, which should be held during the sprint. [19] The author also points out that this meeting serves to give visibility to the progress of the sprint and of the work which has still to come.

E. Sprint Review

At the end of each sprint, the team gathers and presents the evolution of the development to the client in an informal manner. In general, a demonstration is given of the new features and functionalities developed. Participants examine the outcome and decide on future activities. At this meeting, new items can be added to the product backlog [17].

Scrum also defines, according to Schwaber and Beedle [17], four roles with very different purposes and tasks during the process and its practices. These are: (1) the product owner, (2) the scrum master, (3) the scrum team, and (4) the client.

V. KNOWLEDGE MANAGEMENT IN SCRUM

How is the knowledge generated during the phases and practices of the Scrum to be managed? In Scrum, with its agility, the creation, transformation and transfer of knowledge are dynamic which makes it difficult to store and subsequently disseminate this knowledge.

The teams, which are normally smaller, need to exchange information constantly to generate knowledge and make it independent of experts. This requires leaders to encourage the team to learn, to design it so that it is learning-oriented and to provide means so that knowledge is transferred constantly [15].

Cohn [15] suggests that teams should use tools as a means to enhance and enable the exchange of knowledge, as well as for storing and disseminating knowledge. For companies to succeed, teams should have concrete ways to share what they have learned not only with each other but also with the rest of the organization. There are several ways of doing this such as using many Scrum forums for internal communication. The daily scrum meetings disseminate information among team members, and possibly among some additional participants. Usually the Sprint reviews are sent to a wider readership, especially if stakeholders and members of other teams take part in them.

In large companies the scrum of scrums enables teams to share information with representatives from all the other scrum teams. Tools are also used in the teams to help them spread knowledge. Wikis and large graphics, besides the noticeboard, give a good view of the current state of the sprint and the project among team members and to anyone that sees them [15].

Besides using tools, it is natural for team members to talk informally and in a non-planned way with members of other teams in order to exchange information and experiences. Often large Scrum projects and departments form communities of practice, in which groups of like-minded people or with similar skills can meet regularly to discuss and share not only common problems, but also and above all, the solutions found. Communities of practice are an excellent means of disseminating and sharing knowledge between the teams, and consequently throughout the organization [15, 8].

It is also necessary, according to Cohn [15], eliminate the waste of knowledge. This waste is characterized by the dispersion that sometimes affects the team, when it is halted in order to pass on its status when urgent activities are being concluded or when some member is held up several times. The transfer of knowledge can also become a waste if it is not well structured and prepared.

As the dynamic style shows better results in terms of performance and efficiency in knowledge management, attention must be paid to both tacit and explicit knowledge, without giving preference to either one [9]. The tacit knowledge that is generated throughout the Scrum process should be transformed into explicit knowledge such that it is stored by using the appropriate tools so that it is disseminated to all in the team and the organization [15].

VI. CONCLUSION AND FUTURE WORK

The use of the agile Scrum methodology made development more efficient by allowing products to be developed in an iterative way by prioritizing what gives most value to the client. On receiving their product faster and with a better quality, clients had the possibility of meeting the time-to-market target, thus positioning themselves in a competitive and long-lasting way in the market.

The development teams underwent a process of adaptation, in which the culture of traditional development culture gave way to the culture of agile development, thus changing their habits and practices. Communication between team members became much more intense and the production of knowledge stronger.

The practices of knowledge management as well as the adoption of its practices enable organizations to position themselves competitively in the market by aligning the development of their products to their business strategies. Thus, it is seen that by combining the use of agile practices with KM practices makes the performance of the organization much more efficient.

As to future studies, Scrum practices should be mapped with KM practices, thus creating a guideline so that KM might support the use of the Scrum agile methodology on a defined

and structured way, thus enabling agile teams to use Scrum to improve the process and make it more efficient.

VII. REFERENCES

[1] S. C. Vieira, "Um modelo baseado em inteligência artificial para a gestão do conhecimento aplicado ao processo de desenvolvimento de software", p.146, Universidade de Brasília. Brasília, 2008. (portuguese only)

[2] R. S. Pressman, Engenharia de Software, 5th ed., São Paulo: Makron Books, 2002, pp.3-33.

[3] K. Beck, *et al*., Manifesto for a gile software development, http://agilemanifesto.org/, 2001. Last accessed on 21/02/2012.

[4] H. Takeuchi, I. Nonaka, Gestão do Conhecimento. 1st ed., Porto Alegre: Bookman, 2008, p.320.

[5] R. Bose, "Knowledge management metrics". Industrial Management & Data Systems, vol. 104, no. 6, pp. 457-468, 2004.

[6] A. Aurum, J. Ward, "Knowledge management in software engineering – describing the process", Australian Software Engineering Conference. IEEE, 2004.

[7] N. Brookes, N. Burns, A. Dainty, S. Morton, "Social processes, patterns and practices and project knowledge management: A theoretical framework and an empirical investigation", International Journal of Project Management. UK: Elsevier, 2006, pp.474-482.

[8] E. Wenger, "Knowledge management as a doughnut: Shaping your knowledge strategy through communities of practice", London: Ivey Management Services, 2004, in press.

[9] B. Choi, H. Lee, "An Empirical Investigation of KM Styles and their Effect on Corporate Performance". Information & Management, vol. 40, pp. 403-417, 2003

[10] K. C. L audon, J. P. Laudon, "Managing Knowledge for the Digital Firm," in Management Information Systems: Managing the Digital Firm, 8th ed., Ed. Pearson Education, 2004, ch. 10.

[11] M. Fowler, The new methodology, http://martinfowler.com/articles/newMethodology.html, 2005. Last accessed on 26/02/2012.

[12] J. Highsmith, "What is agile software development?". The Journal of Defense Software Engineering, 2002.

[13] VERSIONONE, The state of a gile development: Sixth annual survey. Posted as a downloadable PDF in the Library of W hite Papers on the VersionOne website. http://www.versionone.com/pdf/2011_State_of_Agile_Development_Survey_Results.pdf, 2011. Last accessed on 27/02/2012.

[14] K. Schwaber, Agile project management with scrum, Redmond: Microsoft Press, 2004

[15] M. Cohn, Succeeding with Agile: Software Development Using Scrum, 1st ed., Addison-Wesley Professional, 2010, p.496.

[16] M. Mah, "How agile projects measure up, and what this means to you". Cutter Consortium Agile Product & Project Management Executive Report, Vol. 9, No. 9, 2008.

[17] K. Schwaber, M. Beedle, Agile software development with scrum, 1st ed., New Jersey: Prentice Hall, 2002, p.158.

[18] P. Abrahamsson, *et al*., "Agile software development methods: reviews and analysis". Espoo: VTT Publications, 2002.

[19] J. Sutherland, K. Schwaber, "The scrum papers: nut, bolts, and origins of an agile framework". Boston: Open View Labs, 2010, p.219.

Proceedings
International Workshop on
Distant Education Technology
(DET 2012)

An educational game to learn type 1 diabetes management

Pierpaolo Di Bitonto[1], Teresa Roselli[1], Veronica Rossano[1], Elda Frezza[2] and Elvira Piccinno[2]
[1] Department of Computer Science, University of Bari
[2] Department of Endocrinology and Diabetes in Childhood and Adolescence, "Giovanni XXIII" Hospital
Bari, Italy
{dibitonto, roselli, rossano}@di.uniba.it

Abstract — Virtual environments have proved to be successful for educational purposes in different fields. They are particularly useful when learner experience is necessary but impossible in a real context, like learning type 1 diabetes management, for example. When a diagnosis is made, in fact, the young patient is asked to learn, in a short time, how to control blood sugar levels and adopt a healthy life style. The paper describes a virtual environment in which young users are trained to control the balance between energy/physical activity, one of the main issues in diabetes care. Moreover, a pilot study aimed at measuring the users' learning achievement reveals that the active learning approach is successful even if users have poor motivation.

Keywords: virtual game, active learning, design model

I. INTRODUCTION

Active learning is one of the most effective approaches to learning. There are numerous studies in pedagogy literature that provide experimental evidence of students' deeper learning in active learning situations. In a report by the Association for the Study of Higher Education (ASHE), Bonwell and Eison [1] discussed different methodologies for promoting active learning. Simulation in a virtual environment and role play are the most widely adopted methods used to engage students in active learning. The wide use of Second Life, in fact, has given simulation and the virtual world a stronger impact in the educational process. The use of the virtual world and role play allows immersive learning where the learner experiences as closely as possible the emotions and the style of the real world. This kind of learning is especially successful in fields where the learner experience is difficult or dangerous (i.e. chemical experiments), expensive (i.e. physical experiments), or where the learner experience is necessary but not always possible in the real context, in order to train the learner to become a surgeon, pilot, lawyer, and so forth.

The research described in this paper aimed at designing and building a virtual world to train young people to manage type I diabetes. In medical contexts, education and training of patients are essential for those diseases that can be kept under control by adopting a healthy lifestyle, recognizing early symptoms and intervening promptly. In particular, those issues are important when patients are extremely young, as in type 1 diabetes, also known as juvenile onset diabetes.

In order to make learning and training processes more attractive, practitioners' studies and researches are, in recent years, focused on how new technologies can be implemented to improve young patients' self management skills. The growing number of edutainment on type 1 diabetes [2, 3, 4, 5, 6, 7, 9] is evidence of this trend.

The main purpose of this paper is to present a virtual environment in which young users are trained to control the balance between energy/physical activity, one of the main issues in diabetes care. The virtual environment has been developed in collaboration with the medical team, Elda Frezza, Elvira Piccinno and Cataldo Torelli, of the Department of Endocrinology and Diabetes in Childhood and Adolescence, at the "Giovanni XXIII" Hospital in Bari, Italy.

The paper is organized as follows: section I discusses the domain issues that are the basis of the research work; section II discusses the technological issues faced; section III presents the virtual environment and section IV illustrates the pilot study aimed at measuring users' learning achievement. Finally some conclusions and future work are discussed.

II. DOMAIN ISSUES

When a diagnosis of type 1 diabetes is made, the patient is asked to undergo periodical check-ups, in order to control the general health conditions, and to learn that the most important person in managing diabetes is her/himself. For this reason, the young patient and her/his relatives have to learn, in a short time, the basic steps of diabetes management:

- how to monitor blood sugar at home;
- how to recognize and treat low blood sugar (hypoglycemia);
- how to recognize and treat high blood sugar (hyperglycemia);
- how to plan a diabetic meal;
- how to give insulin;
- how to check blood glucose and urine ketones;
- how to adjust insulin and food during physical activities;
- how to handle sick days;

and so forth.

All this information, or at least the most important, has to be acquired during the (first) hospitalization, but it is difficult both for young patients and for their families. In addition, their young age and their refusal to be seen as "children with special needs" make the process even more difficult. In this context, new technologies can help to make the learning process more engaging. Moreover, the learning process must not end during hospitalization but is a continuous process throughout the patient's life and daily activities.

III. TECHNOLOGICAL ISSUES

To apply the active learning approach, that is usually successful in stimulating interest in young users, a virtual environment was designed and developed.

The design process of the virtual environment was split up into three dimensions: the domain content, the space/environment, and the interaction. The lack of formal models to design virtual worlds for education prompted us to use a hybrid approach that integrates existing models with ad hoc procedures.

As regards the domain content, since the final aim of the virtual environment was very similar to an educational hypermedia, two existing design hypermedia models were combined. Firstly, Allessi and Trollip's guidelines [8] were used to define the scope of the environment and to analyze final users. Then the model of domain content was defined, according to Allessi and Trollip's model and the RMM (Relationship Management Model) [10]. Both of them allow the relevant domain concepts to be individuated and structured using the hypermedial structure.

For the space/environment dimension a model of physical environments and furnishings was introduced in the design phase. In particular, the model requires a map to be drawn for each environment, and for each object in the environment a formal description is required. The information about the object is related to: the domain notion to be represented, the real object, the virtual object, the arrangement in space, and the functionalities that the virtual object should have.

Finally, the interactions should be described. Since in the environment there are different types of actors, three types of interactions have been defined: object-object (O-O), object-avatar (O-A) and avatar-avatar (A-A). For each of them, it is necessary to describe the actors involved, the direction of interactions, the event that is responsible for the start of interactions (i.e. user click), and the effects of the interactions. The interactions are formally defined using the sequence diagram, usually used in software engineering to model interactions among processes.

In order to develop the designed virtual environment, OpenSimulator [11], an open source multi-platform and multi-user 3D application server, was used.

IV. THE VIRTUAL ENVIRONMENT: DIABETLAND

The virtual world developed, called Diabetland, aims at training young people to manage type I diabetes. The main idea is to use the game approach in order to make the learning and training process more attractive and intuitive for children aged between 9-14 years. The challenge of the game is to allow users to understand what physical activities they can do during play, accordng to their blood sugar levels and energy store.

The user has to guide the avatar through different levels in the virtual world. In each level, the avatar has to face different missions, defined in order to allow specific issues on diabetes to be acquired. The issues have been defined by the medical team, that selected the most important information for the final users of the game.

In particular, the levels are:

- terminology - to present the basic concepts of diabetes;
- diet - to train users to recognize healthy foods for the diabetic diet;
- energy/physical activity balance - to train users to monitor their blood sugar levels during physical activities.

A. *Model of domain content*

As stated above, the virtual environment has been designed and developed in collaboration with the medical team of the Department of Endocrinology and Diabetes in Childhood and Adolescence of the "Giovanni XXIII" Hospital in Bari. The doctors defined all the domain concepts to be implemented. The task analysis, proposed in [8], allows, firstly, definition of the complex concepts to be represented in the game and, then, their subdivision into the entry-level skills, which represent the concepts that the user has to acquire to reach the learning objectives of the educational game. In figure 1, the defined task analysis for the three levels is depicted.

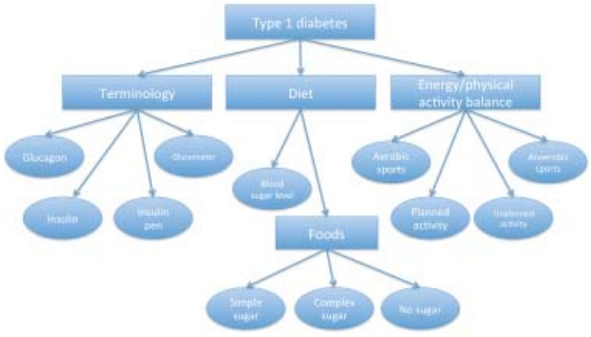


Figure 1. Task analysis [8]

B. *Model of physical environments and furnishings*

For the space/environment dimension, a model of the physical environments and description of the furnishings in them are required. DiabetLand is set in a home; in each room a mission is presented and in order to proceed through the rooms the system requires the avatar to accomplish each mission.

The steps needed to design a virtual world are the same needed for a real world. In both cases, it is necessary to define

how to arrange outdoor spaces using streets, buildings, or gardens and how to furnish the rooms inside a house.

For these reasons, in DiabetLand a map was drawn of each environment and a detailed description of all the furnishings in it was defined. Figure 2 depicts the map of second level, the kitchen.

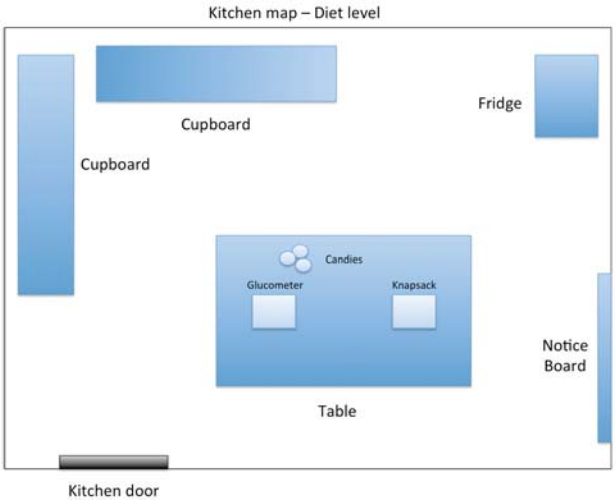


Figure 2. The kitchen map

The map is accompanied by the list of the main furnishings in the kitchen:

- the kitchen door;
- the cupboards and the fridge containing the foods;
- the candies (visible only if the blood sugar level is 60);
- the table;
- the knapsack;
- the glucometer;
- the notice board.

For each of them, a detailed description is given to allow the developer to understand how each object should be represented in the virtual world, what are its functionalities (if any), and how it should interact with other objects or with the avatar.

For example, the description of the glucometer is the following:

- the domain notion: the blood sugar level should be controlled periodically;
- the real object: the instrument to measure the blood sugar level;
- the virtual object: the glucometer;
- its arrangement in space: on the table near the knapsack;
- the functionalities: when clicked they show a blood sugar level;
- feedback: random value of blood sugar level (the available values are: 60, 90, 150, 220).

C. *Model of interaction*

As already stated, the model of interaction has been subdivided into three different layers: object-object (O-O), object-avatar (O-A) and avatar-avatar (A-A). In this first prototype no avatar-avatar interaction has been developed.

As regards the O-O interaction, a communication protocol has been designed for each object, according to which each object can exchange messages with another object. Usually the communication is based on the message-passing approach of the object-oriented paradigm.

In the diet level, for example, the candies, initially set as invisible, interact with the glucometer and the knapsack. When the click message is received by the glucometer a blood sugar level is chosen among the available values. If the value chosen is the lowest one, 60, the glucometer sends a message to the candies that become visible on the table. At the same time, the candies send a message to the knapsack which cannot receive foods until the avatar has eaten the candies and so corrected the blood sugar level.

D. *Example of use*

When the game starts, the avatar is outside the house. The first mission is proposed in the living room: learning diabetes terminology. The user has to discover the stuff about diabetes in a sideboard; only when all the things are discovered can the game continue, and the avatar can go into the kitchen.

Figure 3. The kitchen. On the table are the glucometer, some candies, and the knapsack to be filled

The second mission proposed takes place in the kitchen (figure 3): diabetes diet. In this level the user is asked to put in the knapsack those foods necessary to go outdoors and do physical activity. But, the first action required is to check the blood sugar level using the glucometer. On the basis of the value found, the user has to make the right choice among the proposed foods. The level aims at allowing users to acquire the ability to recognize the foods that contain simple sugars, complex sugars or neither.

The final mission takes place outside the house: energy/physical activity balance (figure 4). On the basis of the blood sugar level registered using the glucometer and the foods in the knapsack, the user has to pick the right physical activity, choosing among aerobic and anaerobic sports. The notice boards are used to give information about how much energy is required by the different type of activities, and about the actions needed to face planned physical activity and the actions needed for unplanned activities.

Figure 4. Level of energy/physical activity balance. The notice board on the right side explains the mission for the avatar, and the notice board on the left side gives information about different sports and if they are aerobic or anaerobic.

In this level if the user has a low level of blood sugar and/or has put the wrong foods in the knapsack the system gives the avatar a negative feedback. For example, if the blood sugar level was 80 mg/dl (the lowest acceptable value), in the knapsack the avatar has only protein (ham, frankfurters, and so on) and the avatar wants to play soccer, the system will remind him "WATCH OUT! You cannot play soccer because there are no sugary foods in your knapsack".

V. PILOT STUDY

The pilot study was carried out for different purposes: to see whether the virtual environment could improve users' knowledge about diabetes and to verify if users appreciate the active learning approach.

A. Sample

This study involved 17 individuals. The users, aged 9-14 years, were all patients at the Department of Endocrinology and Diabetes in Childhood and Adolescence of the "Giovanni XXIII" Hospital in Bari.

B. Procedure

First of all, the doctors and the authors illustrated the aim of the meeting and the virtual environment, giving the users only the instructions necessary to start the game. Then, the users underwent a pre-test to assess their prior knowledge of diabetes. Using different computers, each user played the game several times. In this way, the user had to handle different situations, since different blood sugar levels are proposed for each interaction. Afterwards, all users were given a post-test.

C. Results and Discussion

Comparing pre and post-tests supplied to the sample, the pilot test results show that the users improved their knowledge about all three levels.

In particular, as regards the terminology level the test asked them to describe the basic concepts of diabetes: insulin, insulin pen, glucagon, glucometer. In this level it is important to note that all the subjects gave the right answers, underlining that all users had learned the basic information about diabetes.

For the diet level, the test required users to discriminate the difference among simple and complex sugars and to recognise among a set of foods those containing simple sugars, those containing complex sugars and those containing neither of them. In the pre-test, 9 out of 17 gave the right answers, 3 gave partially correct answers and 5 gave wrong answers. In the post test, instead, 7 gave the right answers, 7 partially correct answers, 2 the wrong answers and 1 no answer at all. The differences are minimum, but a deeper analysis of the results shows that the users who gave wrong answers were the youngest. In one case, in fact, a child was continuously trying to ask his father, who stated that his son was still not able to independently manage diabetes in general and meals in particular.

The last part of the tests was designed to measure the ability to balance energy/physical activity. Figure 5 summarises the results. In this case, there is an interesting difference between the pre-test and post-test. The learning gain is probably due to the active learning approach, in which the avatar has to choose the physical activity that can be done on the basis of her/his blood sugar level and the foods in the knapsack. In addition, in this level different notice boards have been placed to describe the relevant information about physical activities. The users who gave the wrong answers were the same as in the previous questions.

It is important to notice that the pilot study was carried out with a small sample, composed of patients in the hospital for the periodic check-up. None of them had been warned about the learning session and some of them, particularly the older ones, were annoyed at first, because they considered it a waste of time since they felt they were already experts. Instead, a brief interview after the session showed that the less motivated players were the most enthusiastic. One of them exclaimed: "If only we'd had it at the beginning, the learning sessions with the doctors would have been more fun". This confirms that the active learning approach is successful when used with younger age groups, even groups with little initial motivation.

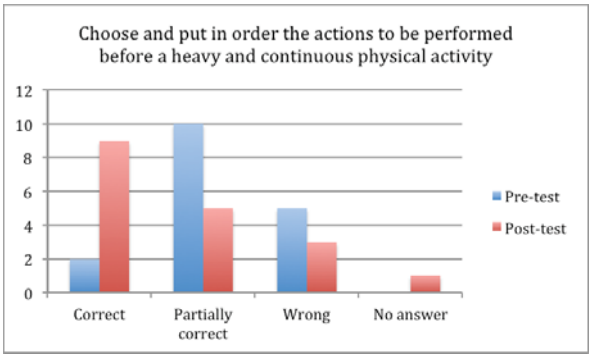


Figure 5. Pre-test and Post-test results comparison

VI. CONCLUSIONS AND FUTURE WORKS

The paper describes a virtual environment designed to train young people to manage type 1 diabetes. In this kind of disease, learning and practicing a healthy life style is essential. Young patients have to learn basic information about diabetes in order to prevent and treat low blood sugar (hypoglycemia) and high blood sugar (hyperglycemia). Since for young children, being individuals with special needs can make the learning and training process very difficult, it is necessary to think up an expedient to attract their interest. New technologies can be helpful in this. In this context, a game in a virtual environment, named DiabetLand, has been designed and developed. The main aim of the game is to support young patients, teaching them that it is possible for them to play any sport, even with diabetes, but it is necessary for them to control their blood sugar level, and have foods by them that could help during physical activities. To design the educational virtual environment, different existing models for designing educational hypermedia have been used and modified.

The pilot study reveals that users improved their knowledge about diet and the energy/physical activity balance, and that they much appreciated the environment as a learning tool.

The next step will be to expand the virtual world, implementing other levels. At the moment, the game is being used in the waiting room at the Hospital by those patients who are waiting to see the doctor. For this reason, further data will shortly be available to evaluate the levels implemented and to collect ideas about the new levels to be developed.

ACKNOWLEDGMENT

We would to thank Vito Rossano, the student who designed and developed DiabetLand prototype in his bachelor thesis in Informatics and Multimedia Communication and the young patients and their parents that participate to the pilot study.

REFERENCES

[1] Bonwell, C., Eison, J., "Active Learning: Creating Excitement in the Classroom", AEHE-ERIC Higher Education Report No.1. Washington, D.C.: Jossey-Bass.

[2] N. Aoki, S. Ohta, H. Masuda, T. Naito, T. Sawai, K. Nishida, T. Okada, M. Oishi, Y. Iwasawa, K. Toyomasu, K. Hira, and T. Fukui, "Edutainment tools for initial education of type-1 diabetes mellitus: initial diabetes education with fun". In M. Fieschi, E. Coiera, Y. Jack Li (Eds), Medinfo 2004, Amsterdam: IOS Press, pp. 855–859.

[3] S. J. Brown, D. A. Lieberman, B. A. Germeny, Y. C. Fan, D. M. Wilson, and D. J. Pasta, "Educational video game for juvenile diabetes: results of a controlled trial", Informatics for Health and Social Care, 1997, vol. 22(1), pp. 77-89

[4] N. Aoki, S. Ohta, T. Okada, M. Oishi, and T. Fukui, "INSULOT: a cellular phone-based edutainment learning tool for children with type 1 diabetes", Diabetes Care, March 2005, Vol. 28(3), pp.760.

[5] J. DeShazo, L. Harris, A. Turner, and W.J. Pratt, "Designing and remotely testing mobile diabetes video games", Telemed Telecare, 2010, Vol. 16(7), pp. 378-382.

[6] J. DeShazo, L. Harris, and W.J. Pratt, "Effective Intervention or Child's Play? A Review of Video Games for Diabetes Education Diabetes", Technology & Therapeutics, October 2010, Vol. 12(10), pp. 815-822.

[7] Panarea digital, idiabetes iPhone app, http://www.recupera.panareadigital.com/projects/managing-type-1-diabetes/

[8] S. Alessi, S. Trollip, Multimedia for Learning: Methods and Developmen http://itunes.apple.com/it/app/eurotouch-did-diario-interattivo/id422056479?mt=8

[9] Me.Te.Da. srl, EuroTouch DID - Il Diario Interattivo per il Diabete, iPhone app

[10] T. Isakowitz, E. A. Stohr, and P. Balasubramanian, "RMM: A Methodology for Structured Hypermedia Design", Communications of the ACM, August 1995, Vol. 38 (8), pp. 34-44.

[11] OpenSimulator, http://opensimulator.org/wiki/Main_Page

Building Wider Team Cooperation Projects from Lessons Learned in Open Communities of Practice

Paolo Maresca
Dipartimento di Informatica e Sistemistica
Universita' di Napoli Federico II, Napoli, Italy
paomares@unina.it

Angela Guercio
Department of Computer Science
Kent State Univ. at Stark
Canton, OH
aguercio@kent.edu

Lidia Stanganelli
Dipartimento di Informatica, Sistemistica e Telematica
Università di Genova
Genova, Italy
lidia.stanganelli@unige.it

Abstract *- Communities of Practice (CoP) are a cornerstone in the design and innovation process of knowledge acquisition and knowledge transfer. This paper presents an empirical evaluation of an experimental Open Community of Practice (OCoP), designed to deliver a programming course. The empirical model fosters cooperation between communities which share similar learning objectives. The results show that knowledge transfer is highly increased when a formation paradigm is used together with a set of cooperation tools. A scalability model has been extracted from the results of this experimental community. The model is currently being used as a launching platform to start an open collaborative learning environment to connect two communities of practice, one located in Italy and the other in the United States.*

Keywords: ***Community of Practice, collaborative learning, distance learning.***

1. INTRODUCTION

Communities of Practice (CoP) with their knowledge transfer are vital to boost the economy and the life of small and mid-range companies, especially in times of recession. In a similar way, communities of practice in academic institutions with their research activity are of great value to our society.

Each Community of Practice has its own mission. It is identified by its knowledge, its research area, its achievements, and its members [11].

When two CoPs share an interest in cooperating, there is a high potential of knowledge transfer, but this potential remains limited if the communities are not able to create an effective flow of knowledge between the members of the communities. While it is realistic to expect that such a flow may encounter initial resistance at the beginning of the cooperation, it should become more and more fluid with time, therefore producing an effective knowledge transfer between the two communities.

When we look at existing Open Communities of Practices (OCoPs), we observe that the majority of them spontaneously cooperate by using social networking tools. When they cooperate in an effective manner, their results are of great interest. For example, consider the successful results produced in the area of open source software by such communities and the extensive production of software development that these communities have either generated or stimulated.

In this paper we consider an academic OCoP shaped around a programming course and we design a process that is able to bring closer together communities with similar objectives. In this process, the teacher, in quality of the community moderator, moves from the classic role of knowledge transfer to the role of "coach" who motivates, stimulates, and sows ideas, in order to attract the interest of its members, engage their abilities, and facilitate the flow of knowledge on the net. Moreover, the teacher should support the natural and spontaneous attraction that such communities reveal while reaching common goals. Innovative teaching tools must be used to support such a dynamic community. The experimental community described in this paper is part of the "Enforcing Team Cooperation Project" (ETC project) [6] which involves seven Italian universities with the common goal of educating and training wide groups of students in the area of software engineering and software development [3, 4]. The community provides a contextual collaboration, embeds collaborative systems, and uses the IBM Jazz collaborative application

development environment [1] to implement this experience. The results analyzed in this paper come from the experimental community built at the Faculty of Engineering of the University of Naples, Italy. The study of this community is used to identify a scalable model for the formation and growth of wider international academic communities that have joint common goals.

In Section 2 we describe the ETC project and in Section 3 we introduce the experimental open community derived from the ETC project. In Section 4 we discuss the technical details, the results, and the lessons learned from the experimental community. Finally, in Section 5 we draw our conclusions and describe our future research.

2. THE ETC PROJECT

The goal of the ETC project is to prepare a generation of students who are able to best compete in the future job market by incorporating in everyday life the use of modern software engineering tools, by enforcing student collaboration and by giving them a modern working experience.

The objectives of the ETC community are:

1. To strengthen the teaching of the software engineering disciplines with industrial tools of high level and high quality.
2. To generate opportunities for the students to use modern technologies for the design, development, testing and quality control of software products.
3. To integrate several of the communication paradigms available on Internet (i.e. chat, collaboration, web 2.0, social web, mashups) in the academic learning environment and to use them as a modern platform for the management of software development and knowledge transfer.
4. To offer real opportunities to every student to play one or more than one of the multiple roles of the software engineer (i.e. analyst, designer, developer, tester, etc.) in an educational project.

In order to achieve the above objectives, a set of tools that enable the students to interact with each other beyond the boundaries of the single course or university is required. In fact, we believe that when such collaboration moves from within the course outside its boundaries, knowledge transfer is enhanced, assets are produced [8], and a collaborative interaction among students is encouraged. In this open collaboration, there is the creation of a real software team, which potentially may include all the heterogeneous engineering roles, epitomized by the students of the various courses participating in the project.

The ETC is an incubator for cooperative learning communities and it generates an ecosystem of innovative ideas for teaching and learning in an academic environment. An ETC community was designed at the Faculty of Engineering of the University of Naples and empowered the students with learning tools, such as the IBM Rational software platform. With the help of this tool of cooperation, students have acquired knowledge in a modern fashion by acting like a distributed team. They have cooperated with other students in the same course as well as with students in different courses and in different universities. Consequently, in the spirit of this experience we can say that in such communities the participating universities and the students overcome traditional physical limitations, cooperate in a distributed fashion, share their assets and their experiences, greatly improve their tutoring skills, and foster a general mutual collaboration.

3. BUILDING AN ETC ON RATIONAL JAZZ

Many challenges have been encountered during the development of the experimental community. When cooperative learning is implemented, it is important to understand how knowledge management occurs [10] and which teaching resources must be used to guarantee the success of the community [9].

In this section we describe the experimental ETC community built on top of Rational Jazz. The benefits derived from this approach are:

1. Enrich the students' curriculum of studies by working on a modern interconnected and instrumented platform and change the way students study in general computer sciences and software engineering courses.

2. Stimulate the participants and drastically improve their skills.
3. Foster learning cooperation between universities.
4. Generate better grades for students and a better teaching environment for instructors.

The implementation of this community has required the activation and the monitoring of a set of processes such as:

- Assigning work items to students
- Controlling the progress and students' improvements
- Building and managing teams
- Designing and assigning distributed team work
- Stimulating and moderating collaboration between different courses and universities
- Empowering the courses with industrial tools

The experimental ETC community has been created at the University of Naples for two courses: Software Engineering and Programming II. The Open Unified Process, the iterative software development process, has been chosen for the course of Programming, while the SCRUM agile process with 4 sprints has been chosen for Software Engineering.

The solution was instrumented by allowing the students to interact with the knowledge of the academic project via smart phones (i.e. Android, Apple iOS, Symbian). The ETC experimental community was implemented on the IBM Rational Jazz platform (fig. 1) in a distributed environment upon which the ETC project connects seven physically distributed universities in Italy. The IBM Rational modules chosen for the implementation of the experimental community include: Team Concert v. 2.0.0.2, Requirements Composer v. 2.0.0.4, Quality Manager v. 2.0.1, Asset Manager, the RequisitePro, and IBM DB2 UDB for Windows. An IBM System X 3650 powered the solution. The virtual environment was generated by the creation of a virtual software engineering laboratory that could be accessed in a distance learning fashion. The communication between teachers and students was performed within Jazz. The system supported the teachers with tools for the monitoring of students' workload and of students' progress. The experimental application created on top of the Jazz platform was innovative and permitted for the first time the connection of a large number of students.

This way of working in teams and of communicating in a broader environment on distributed projects is similar to the experience required in a real job. Therefore, it is vital for students who are preparing for the job market. In addition, this experiment of learning cooperation challenges the way our universities provide, manage, and transfer knowledge, and encourages them to indentify and introduce new ways to educate.

From the results observed, we believe that this solution can easily grow and connect more universities in different countries by reusing the processes created in this experimental community and by leveraging the lessons learned from this experience.

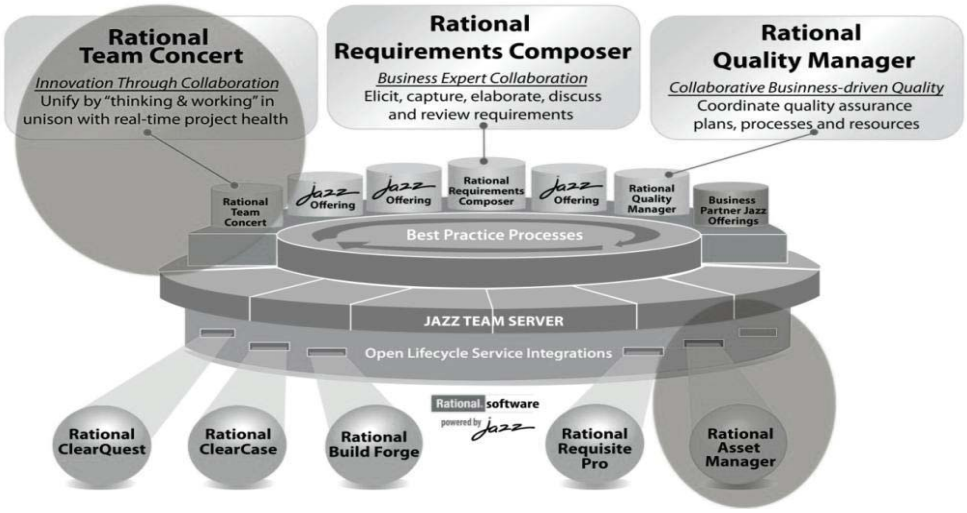


Figure 1 – The ETC platform on Rational Jazz

4. LEARNING FROM THE ETC EXPERIENCE

The experimental ETC community was set in place in 2011. Since its launch, the students have produced interesting assets such as:

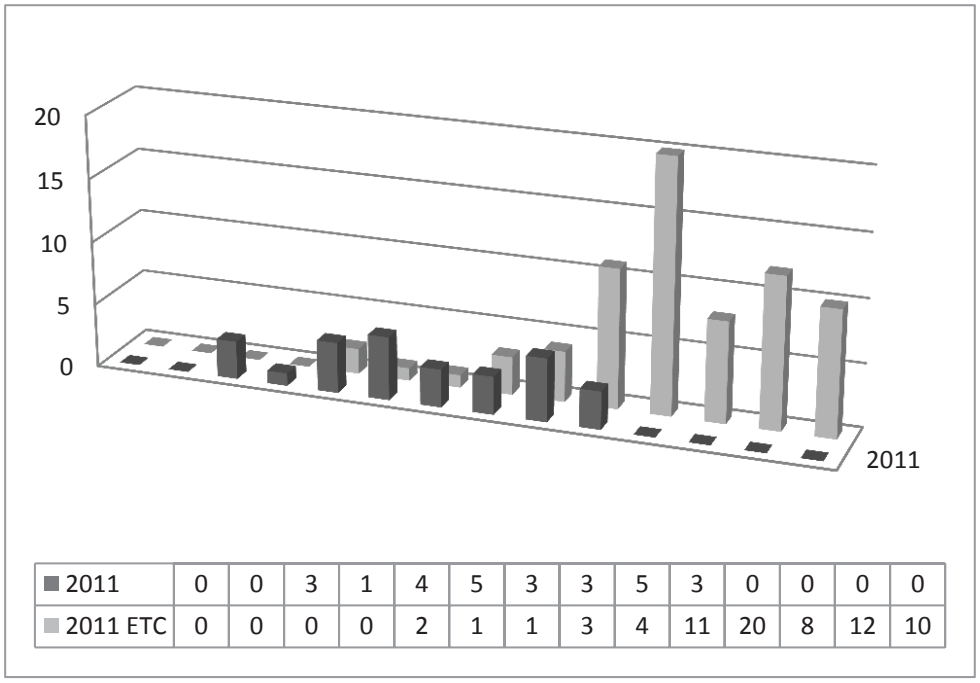
- tutorials
- quick start guides
- virtual client machines to help new students
- new Eclipse plug-ins
- Smartphone apps for Android, Apple iOS, and Symbian, for the use of the Rational Asset Manager in mobility.

These products are the first tangible measure that the platform provided an effective environment to support the users, to foster the growth of knowledge, the success of students and of teachers, and to build skills that appeal to modern companies.

The virtual software development laboratory for Programming II had 187 students divided in two groups: one group participated in the experimentation, the other did not. The course was composed of 9 credits of theory and 3 credits of laboratory. The creation of a virtual laboratory was dictated by the high number of students which was too large to manage in a single laboratory (the physical lab capacity was of 60 students). Without the virtual laboratory, the course would have required 3 repetitions of the same laboratory, taking away almost 1/3 of the hours from the theoretical part of the course. In the experimentation, each student had access to a local repository by using the Eclipse platform to connect with the Rational Team Concert [2, 5]. A set of 23 homework assignments, performed in 4 distinct iterations was created. The performance of the ETC students participating in the project is shown in figure 2 and 3.

Fig. 2 shows the results of the data collected for the ETC group and the non-ETC group. The graph shows a visible distribution of the students' grades on the high part of the interval denoting a higher success of the students.

In fig. 3 the mean value of the ETC 2011 grade is compared to the mean of the grades of the non-ETC 2011 students and the years 2009 and 2008. We observe that in 2010 the course was not held so data are not available. Also in 2008 and 2009 a different paradigm of formation was applied. The graph says that as soon as a formation paradigm was eliminated (i.e. non-ETC 2011) the performance of the students dropped dramatically. The performance of the ETC students was better than the one of the students of the previous years and shows the success of this experiment.



2011	0	0	3	1	4	5	3	3	5	3	0	0	0	0
2011 ETC	0	0	0	0	2	1	1	3	4	11	20	8	12	10

Figure 2 – Results of the 2011 ETC experimental lab

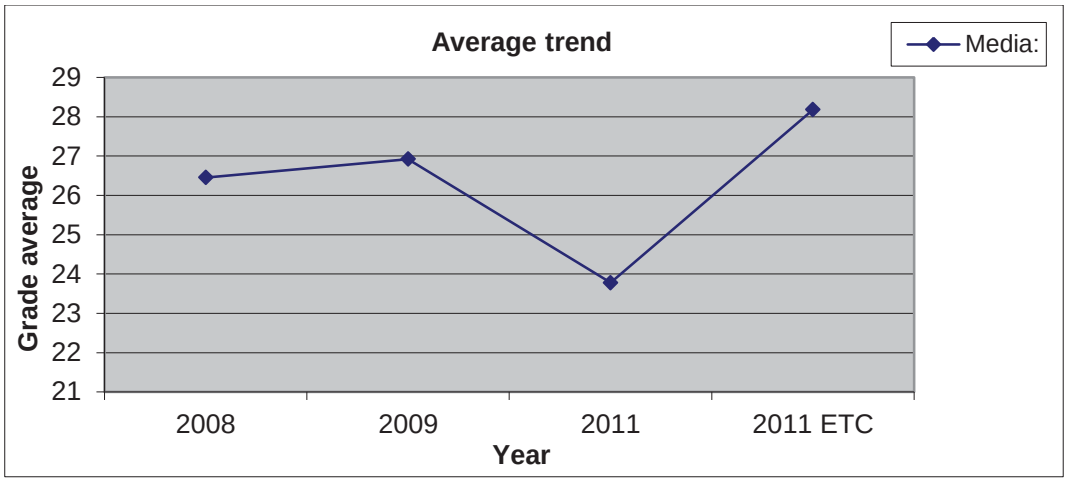


Figure 3 – Average Trend 2008-2011

While the learning curve of the platform was a little steep at the beginning of the experiment, it was overcome with the help of a mini tutorial. By the end of the first semester of the activation of the project, the percentage of students who successfully completed the course increased by 25% with respect to the same semester in the academic year 2009. The number of students who completed the course with a grade of 27 or more (i.e. >90%) increased by 20% with respect to the year 2009. The number of students who completed the course with a score of 30 (i.e. 100%) increased by 28%, and the number of students who obtained the "cum laude" increased from 1 to 10 with respect to the year 2009.

The quality of the product produced by the students strongly improved. Two sophomore students who worked on a software development project published their results in an journal article presented at a national conference. This shows how the project attracts and stimulates the students and opens endless research projects opportunities. The cost in terms of time and management of the laboratory was acceptable. The largest cost was spent for the installation of the server and the relative processes for the creation of the laboratory. Currently the platform already supports the managements of a number of students' artifacts that reaches the order of thousands. We understand that when the number increases a scalable approach is required. A scalability model has been tested by moving the ETC to the cloud. The ETC-HUB (the ETC in the cloud) via Jazz-HUB (http://jazz.net/hub) is ready to open the ETC processes to a broader set of communities [11]. The use of the cloud reduces the set-up and maintenance costs moving the attention of the community to learning goals. In this scenario we are currently developing a new experimental open community that connects two international universities.

5. CONCLUSIONS AND FUTURE WORK

In this paper we describe the results obtained in an experimental open community of practice designed to develop a course of Programming II at the University of Naples. The results obtained from this community have been described and they are very encouraging. Prof. Maresca, who was guiding the community, was awarded the IBM Faculty 2011 award for his experiment.

Many lessons have been learned during the set-up, implementation and development of the community. Leveraging this experience we are currently extending this experimental open community to connect two international universities with the intention of analyzing the behavior of the participants in an area that transcends cultural and linguistic differences. The advantages of an extended community is to provide the virtualization of an environment, such as a course laboratory, that favors students' distance learning, that provides study flexibility, and that minimizes the set-up and maintenance time for the instructor. Moreover, the health of open learning communities is strongly

connected with the level of cooperation that flows between them. However, it is hard to measure if communities are moving apart or if they are getting closer while operating the knowledge management organization [10]. We can only draw conclusions very late in the process by observing their past activities of cooperation and the results of such cooperation. This inability of measuring the degree of cooperation while in progress impedes the ability of intervening while there is still time to revitalize the community and therefore it is the reason for the death of several promising communities. In general the aspects of the fluid process that brings communities together, that pushes them apart, and that brings them back together again is not well understood. We believe that this is due to the dynamic complexity in which communities operate from their birth to their death. The measure of this dynamicity requires the identification of a set of objective attributes that involves the measurement of the activities of the members, their ability to be reachable, their ability to transfer knowledge and tutor other participants, their evolution in terms of skill acquisition, and their asset production. In future work we would like to analyze this problem and provide some helping measures.

REFERENCES

[1] G. Booch, A. Brown, "Collaborative Development Environments", *Advances in Computers*, vol. 59, Academic Press, Aug. 2003, pp. 1-27.

[2] L. Cheng, S. Hupfer, S. Ross, J. Patterson, "Jazzing up Eclipse with Collaborative Tools," *Procs. of 2003 OOPSLA Workshop on Eclipse Tech. Exchange*, Anaheim, CA, 200, pp. 45-49.

[3] Coccoli, M., Maresca, P., Stanganelli, L., "Enforcing Team Cooperation: an Example of Computer Supported Collaborative Learning in Software Engineering", *Procs. of 16th Intl. Workshop on Distance Education Technologies (DET2010)*, Oak Brook, IL, pp. 189-192.

[4] M. Coccoli, P. Maresca, L. Stanganelli, "Computer Supported Collaborative Learning in Software Engineering", *IEEE EDUCON Education Engineering Conf. 2011*, Amman, Giordania, 2011, pp. 990-995, 10.1109/EDUCON.2011.5773267.

[5] R. Frost, "Jazz and the Eclipse Way of Collaboration," *IEEE Software*, vol. 24, no. 6, Nov./Dec. 2007, pp. 114-117.

[6] G. Galli, F. Gorga, P. Maresca, C. Milani, "Enforcing Team Cooperation Using Rational Software Tools: Merging Universities and IBM Effort Together", *Procs. of 5th Eclipse-IT2010*, http://2010.eclipse-it.org/proceedings/

[7] G. Galli, F. Gorga, P. Maresca, C. Milani, "Towards Cloud Learning with ETC and Jazzhub", *Procs. of 6th Eclipse-IT 2011*, Milano, Italy, http://2011.eclipse-it.org/home/?page_id=424

[8] F. Gorga, P. Maresca, "Persistence of the Memory: an Introduction to the Asset and Intellectual Capital Management. Application into the ETC project". *Procs. of 6th Eclipse-IT 2011*, Milano, Italy. http://2011.eclipseit.org/home/?page_id=424

[9] P. Maresca, G. Scarfogliero, L. Stanganelli, "Enhancing Team Cooperation Through Building Innovative Teaching Resources: The ETC_DOC Project", *Procs. of 5th Eclipse-IT 2010* http://2010.eclipseit.org/proceedings/

[10] A. Rhamchand, S. L. Pan, "The Co-Evolution of Communities of Practice and Knowledge Management in Organizations", *SIGMIS Database*, vol. 43, Issue 1, Feb. 2012. pp. 8-23.

[11] E. Wenger, *Communities of Practice: Learning, Meaning, and Identity*, Cambridge University Press, 1998.

Information systems and e-Learning platforms: from loose coupling to deep integration

Luigi Colazzo, Andrea Molinari
Department of Computer and Management Sciences
University of Trento, Trento, Italy
{ luigi.colazzo, andrea.molinari}@unitn.it

Abstract—**The paper discusses the results of our almost 15-years' experience in designing, developing, implementing and using e-Learning system (LMS), with a specific approach in mind. This approach is, in some sense, "against the current" of standardization and "normalization" of LMSs, in our opinion too flattened over these pre-defined, pre-design software platforms. This approach implies the creation (from scratch) of a virtual communities' platform definitely oriented towards the complete integration with (and dependence from) various services offered by the information institution that hosts educational activities. This means not necessarily a traditional educational institution, like schools or universities, but rather any organization where educational (and probably more important, collaboration) activities are supported by ICT-based platforms. The paper presents the main points of contact between the information system of the organization and the services provided by the virtual communities' platform, presenting their characteristics and the main reasons why a virtual communities approach, less related to specific classroom-based activities, is better suited for this integration.**

Keywords: information systems; e-learning; lifelong learning; collaboration

1. INTRODUCTION

This work presents our experience in e-learning settings, particularly related with the idea of an e-learning platform as a component of the Information System of any organization. Specifically, we present our software platform, developed along the years and now used by several organizations in our region for supporting collaboration and educational activities of large organizations. When such a platform enters into organizations, its effects are immediately visible:

- needs for integration with sub-systems existing in the organization: just to mention the simplest ones, integration with the single-sign-on system implemented in the company;
- overlapping of some functionalities of LMS/Virtual communities' platform with pre-existing functionalities in the information system of the organization. Examples: document repository, mailing distribution, virtual room management, forum, etc.;
- Competition with possible new systems entering in the organization, mainly due to the web 2.0 functionalities that nowadays most of the companies intend to implement, and that normally any (serious) LMS is able to supply:
- partially overlapping and competition with some functionalities already present, somewhere in some software. These are the most insidious aspect, because none of the systems (LMS and other information systems) are able to satisfy the specific needs, but all of them are able in some way to supply part of the functionalities needed. The typical example we founded in our experience is the support to document sharing for groups of people without having to mount some network disk for file sharing, normally not appreciated by system administrators, and most of the time not accessible via web. In this case, virtual communities are better candidates, as the on-the-fly creation of a virtual community with a set of services available for the members is a perfect solution for many of these situations, not necessarily related with educational activities.

The last example is what mainly led us, in 1998, to build a new system with virtual communities as the center of our approach. At the time, Moodle™ or similar LMSs did not exist or were not accessible to most of the people, and other solutions were particularly costly, proprietary or not available. In our vision, a virtual community is an aggregation space for participants, thus supporting cooperative activities among users instead of just learning activities.

As previously stated, our platform has been created to be adapted and connected to the information system of the organization. Considering e-learning and collaboration platforms as external bodies, relegated to secondary roles inside the information system, is in our opinion losing an excellent opportunity to improve collaboration and open innovation inside an organization. Integrating eLearning systems with existing information systems is not an easy task[1], mainly due to some resistance and ostracism against learning applications that are seen as not relevant for the organization by the ICT departments. Other difficulties come from the technical side, due to the diversity in these systems. The major difficulty we found, however, is the problem of considering the life-long aspects related to e-learning: continuous learning is a paradigm that just recently has become clear to organizations, therefore putting on the table the strict relationships

between the learning paths of people and their presence inside an organization and its information system.
In this work we will describe the evolution of our virtual community platform towards being a component of any organization's information system. We will show why didactic and cooperative activities in an educational environment require coherent architectural choices in the software platform, in order to facilitate the integration of the learning platform into the existent information system.
The paper is organized as follows: in the next section we will describe the rationale behind our decision of developing a customized version of our "Communities online" platform. The third section will briefly introduce the platform itself, while in the fourth section we will discuss the dilemma between coexistence and deep integration with the Information system's services.

2. VIRTUAL COMMUNITIES AND INFORMATION SYSTEMS: COEXISTENCE OR DEEP INTEGRATION?

"On Line Communities" is a dynamic web application that has been created to support blended learning activities by the University of Trento. Specifically, our activities have been oriented towards three main different settings:

- students of the faculties of the University of Trento: at the moment, we have approximately 16.000 users enrolled in 5.000 different virtual communities;
- public employees of our provincial administration (Autonomous Province of Trento - P.A.T.- Italy). The name of the project, "L3- lifelong learning", is clearly identifying target and objectives of the initiative. In this case, we are talking about potentially 20.000 users. The system has been recently opened to all the public institutions of our province, so at the moment we have approximately 5.000 users, and this number will surely increase in the very near future;
- employees of private organizations with various needs of collaboration services, sharing, social interaction, cooperation etc. All these activities are heavily related, but not necessarily *only* related with educational activities.

These use-cases, and especially the last one, have clearly demonstrated, in our opinion, the need of a tight integration between our virtual communities platform and the information systems of the respective organizations. This integration is not simply an exchange of data between two different databases, but it is a more complex, personalized set of services that should share and keep aligned different sets of information. This is not a dissertation about software engineering tools and architectures to implement these quite obvious needs: SaaS, SOA, web services, SOAP and so on are not the discussion point here. In the "L3" system we realized that the simple re-design of a traditional e-learning system was too limited in respect to the needs, mainly:

- sharing and collaboration services that do not imply external or complex tools: for example, direct access with my credential from everywhere in a single-sign-on perspective;
- users' active participation in a more "social" sense, following not only the hype of web 2.0, but effectively allowing the user to collaborate with more sofisticate tools rather than the usual e-mail message with attachments;
- deep integration with the information systems of the organization, i.e., a full set of services that transparently allow the users to see educational and collaboration activities like any other service available on my desktop: for example, substituting the traditional corporate shared disk resources with repositories organization by project and virtual community the user is participating to.

We would like to demonstrate that having a virtual communities' platform with a different approach to people aggregation is an advantage in respect to traditional LMS. Finally, having built the software platform from scratch is probably the most interesting fact in the perspective we are presenting. The availability of an open source software that has not been built from zero by the organization is not enough to be effective in complex and pervasive integrations with the information system. In contrast with other LMSs, such as Moodle and Blackboard, the system is not based on the metaphor of "course" but that of "virtual community". "On Line Communities" enables to create, in a generalized way, different types of virtual spaces called Communities, to aggregate them into larger organizational structures and to endow them with an educational and collaborative environment. This approach is slightly different respect to the proposal of some other authors, such as Beamish [2], which considers the virtual community as a group of people who communicate using computer mediated communication tools. Participant of these communities are physically in different places but however they can exchange information on common interest in a communitarian way. Virtual communities are considered by Rheingold [3], the original creator of the definition of VC, as emerging social phenomena. We are more influenced by the ideas of Jones [4],[5] where the technological structure of VCs (named virtual settlement) is conceptually separated from the community itself. Our role in this case is a double role: designers and developers of the technological infrastructure, but on the other hand we are teachers and administrators of the platform and of several communities. The separation of the technological capabilities of the platform from how customers use it has been very clear in our design.
The concept of VC is very contiguous with what is commonly defined Web2.0 [6] [10]. We partially integrated some elements coming from web 2.0 and new e-learning approaches, but on the other side we

considered and implemented inside the platform some new concepts to develop collaboration activities, the "virtual community" as a container of activities and collaboration, instead of the strict "classroom" concept. As mentioned, our best test bed has been, among others, the "L3-lifelong learning" project, conducted since 2007, an interesting opportunity to understand the applicability of distance learning and virtual communities to the public sector [6], due to the number of subjects involved, the temporal extension of the project, the extension of topics faced during distance learning sessions and the involvement of the organization in all the activities of distance learning. Our experiments presented interesting perspectives for virtual community systems:

1) organizations need virtual places with collaboration and web 2.0 tools available for employees and partners

2) traditional LMS, like Moodle™, are not suitable for a relevant number of these needs, mainly because they are oriented to e-learning, and their pillars are metaphors like "classroom", "class", "course", rather than other ideas more oriented to collaboration, like "community", "project", "workgroup", "team", "secretary", "board", "office", "department". A classroom is, of course, a community, where collaboration is oriented towards a specific target, i.e., training, but the opposite is not necessarily true. A community, instead, can be many other things than a classroom: a virtual collaboration space for a faculty council, a board of directors, a recreational group, a research group, a temporary association that needs to share some services [13].

3) social networks, if used in learning settings, are not suitable for large and rigid organizations like public administrations, because of social networks' business models, where numbers and objectives are not allowing small, "private" communities where collaboration can take place in a private, reserved, personalized space. The main objective is "the more we are, the more we will interact": this is not exactly what a public institution wants in most settings.

3. ONLINE COMMUNITIES: A BRIEF PRESENTATION

Business models are typical of any product / service that can aspire to success, and software systems have the same needs like any other product. In the case of educational activities and Learning Management systems, the model below pertains to how a certain system hypothesizes that the teaching activities of an educational organization are performed. Therefore the use of a software platform imposes limits on how the task for which the LMS is designed will be conducted.

According to our interpretation, a Virtual Community is not the result of a process of social networking. In fact, it is a virtual space shared by groups of people who have a common goal. A community's virtual space can be simple or complex; for example it can contain further virtual communities, thus establishing a "father-child" relationship. The (virtual) community can be an open space accessible to anyone, but also a restricted space reserved only to people authorized by the community administrator. The users can have different roles with rights and duties which vary according to the virtual community. The system provides users of a community with a range of on-demand services that can be activated and used in accordance with the permissions granted and the roles assigned.

This structure has allowed us to shape the organizational structure of an educational institution easily. For example, the communities of the teaching courses are components of larger communities called "Degree courses", which are in turn part of the "Faculty" community. On the other hand the community "Faculty" also includes the community "Faculty council", restricted to the only members of the council, as well as other heterogeneous communities, such as the "chess circle" or "first year students" community, promoted by the Faculty Board to help new students settle down rapidly. Every community exists within a scope and with one or more goals, and according to these goals, different services provided by the platform can be activated. *On Line Communities* includes many services: the system is able to offer services such as:

- "traditional" asynchronous/synchronous services;
- Lifelong Learning and training on the job services (tutorship, training on demand, contextual search tools,FAQ, etc.);
- Integration with external information systems, such as students' service office , register of the lectures;
- offline courses, recorded, digitized and provided to communities of users with the opportunity of synchronizing the videos with slides, podcasts, webcasts, SCORM modules, etc.;
- self-assessment tools, questionnaires, surveys, opinion polls;
- statistical analysis of users' behavior, collecting in a data warehouse the actions performed by users.

Over the last few years the system has evolved into a platform for professional training orientated to lifelong learning outside academia. [11]. This idea originates from the fact that the context of use should determine the services and instruments which the software should offer, not the other way round, as often happens. Another reason, debated in this paper, deals with the relation between LMS and more general information systems of institutions. E-Learning platforms seem to act in a restricted circle made up of only teachers, tutors and students. This idea is actually deceptive and the LMS should be designed to be a wider and more complex module of information systems.

The portal has the possibility of including a community into another, creating a tree of nested communities. It is also possible to aggregate the communities into a transversal super-community, so creating a mesh of connections composed by the different links among

communities from different branches of the tree. With these types of mechanisms, it could be possible to model many types of collaborative structures, even of a high complexity. For example, hierarchical structures like Faculties, Didactic Paths, Master Degrees, Courses, Work Groups related to a course, all the hierarchical relationships between these communities can be easily represented. Most of all, this hierarchical structure allows to map the organizational chart into a tree of communities able to represent each unit in the company, facilitating the replication of dependency relationships inside the organization onto a structure of virtual collaboration spaces.

4. INTEGRATING VIRTUAL COMMUNITIES AND INFORMATION SYSTEM SERVICES

While adopting the VC platform, users of "Online communities" almost immediately asked for a strict integration with the rest of the Information system. We found, and are still finding, many resistances to this process of integration: different scenarios whose power position has been "attacked" by the generalization of the metaphor of "virtual communities" and its already-available, ready-to-use services:

- pre-existing applications colliding with "online communities" services (Document management systems, internal blogs, forum, time management tools, shared agendas…..)
- competitors already present with their specific offers
- off-the-shelf solutions already selected and therefore threatened by our multi-purpose, all-in-one solution,
- previous projects still not concluded

More than this, it's been quite easy to integrate our platform with pre-existing datasources, using web services created on our side, or on the provider side to exchange data between the different platforms. Here follows a simple list of some of the connections of the VC system with the rest of the information system. Some of them have already been created, some of them are on the way, some of them have been (not mysteriously) declined due to the above pressures or legitimate choices, others are in the future developments of the integration. We have classified them in several categories of integration services:

- authentication and authorization services;
- participants records alignment and exchange;
- attendance certifications;
- unified communication services (voip, videoconferencing, shared desktop, shared whiteboards, instant messaging) ;
- time management(agenda, doodle, task management, project management) ;
- ticketing, tutor and teacher requests;
- questionnaires / polls /self-evaluation tools;
- attendance records (for hr departments) ;
- accounting and erp integration;
- automatic membership and/or enrolment to specific communities/ courses;
- payment management;
- public tenders and concourses;
- social media from e-learning to open innovation.

Due to size limitations, in this paper we will present just some of these items, with the indication of issues found and respective solutions.

A first element is the integration with the organization's authentication and authorization system. Single-sign-on is an immediate need of integration between the educational system and the rest of the information system, but here we find a second problem of integration, i.e., the authorization schema of the employee. What are users' permissions? Which courses can be attended? What materials can be seen by this person depending on her permissions? Who can participate to which course / community? To these and many other issues, any e-learning platform should provide it, but one thing is by far more difficult in traditional "classroom-based" platforms. Virtual communities, as said, have hierarchical relationship among them, while e-learning platforms don't. It has been straightforward to implement a direct mapping between communities and the organization chart of the institution, thus providing a structure of communities that replicate the hierarchical relationships and permissions of the organization. This is not only related with permissions of communities that are part of a certain branch, but also (and particularly) with the objects inside these communities, specifically documents, forum posts, FAQ and wikis. Last but not least, a deep integration is fundamental for another reason, i.e., the instantaneous need of alignment between the original directory of the information system and educational system. For example, an unauthorized access to some material by an employee that recently changed workplace or department is not so relevant for pure educational activities, but could be very important when sensitive information are stored in this repository. When the LMS is used as a document management system for e-learning activities, permissions on documents are dependent not only on the role, but could also refer to the position of the person inside the organization. Virtual communities allow mapping the organization schema on the organization of communities, inheriting all the needed permissions for any organization member, and being immediately updated on changes in workplaces, job position and roles inside the company.

Another field of integration is the communication services available in the system. The interaction with e-learning platform is very often characterized by the presence of a tutor, a crucial person for the success of any e-learning initiative. How can we book a virtual interaction with a tutor? How can we interact with the institution if we have a problem? In our experiences of managing considerable numbers of participants like in our university (16.000) or in PAT (20.000), the

helpdesk created around and e-learning initiative is very important, but most of all, it is important in public settings. We are not necessarily dealing with high-level support (from the teacher, for example), but most of the time what is needed is a pure helpdesk-like service, that takes care about simple problems that prevent the user to use the learning resources. What we have done is using virtual communities created specifically for help-desk, and developing a ticketing system where the user can book a slot with a "solution provider", could this be a specialist for hardware problem, a process tutor or even the teacher herself.

The last example regards a fundamental integration for institutions that manage courses for relatively high number of people. When an online edition of a course is offered to large numbers of users, together with some activities managed by the LMS for the educational part, many other tasks are needed before the beginning of the course delivery. These activities are very often effort-driven activities for the administration of the institution. Among the others:

- enrolling people to the course;
- providing information about the course to potential subscribers;
- providing different and specialized communication to the different actors (potential subscribers, teachers, tutors, administrative staff, etc.);
- create the virtual community for the course;
- upload documents regarding that course;
- organize time of subscribers (meetings, videoconferences, appointments, shared agenda etc.);
- organize a forum for discussions before the beginning of the course.

The above items are just examples of needs of (deep) integration between different components of the information system and a (flexible) platform that could delivery educational contents (but not only). More than this, the integration should be transparent to the user and should not be limited to the pure educational aspects. All these features have been implemented in Online Communities with a very limited effort, as we took advantage of the approach based on a community as a set of services available to users with a specific role to play inside the community itself.

5. CONCLUSIONS

The paper proposes a strong statement in favor of a greater attention for software developers, instructional designers and e-learning managers to the problem of integrating educational platforms with the information system of any institution involved in educational activities. In our opinion, a different metaphor is needed between typical LMS, bonded to the "classroom" metaphor, and social networks, where the main objectives of participants are different from learning. The paper describes the peculiarities of a "built-from-scratch" virtual community system, where some features are specifically devoted to collaboration (thus overcoming main problems found in web 2.0 tools), and other features are totally new and only feasible thanks to specific characteristics of our platform, like inheritance, polymorphism, permissions and roles granularity.

Our experiences in the university context and later (the last five years) also addressing some public administrations and private institutions, allow us to make some comments on the most useful services for our users. Even in the public administration there is an increasing need to create a personal, professional (or learning) space available to its employees, based on a deep integration with the rest of the services that normally employees have at their disposal, not just to store the teaching materials used in class, but also to manage their time and their activities, communicate with the colleagues, carrying out projects or homework.

REFERENCES

[1] Albarrak A. (2010), An eLearning Approach to Integrate Hospital Information Systems in Medical Education, International Journal of the Computer, the Internet and Management, Vol. 18 No. SP1, December, 2010

[2] Beamish, N., Armistead, C., Watkinson, M., Armfield, G., "The deployment of e-learning in UK/European corporate organisations", European Business Journal, 14 (3), pp. 342–50, 2002.

[3] Rheingold H., (1993) The Virtual Community, 1993 (italian translation Comunità virtuali, Sperling & Kupfer, Milano, 1994).

[4] Jones S. G. (1995), *Cybersociety*, Sage, London.

[5] Jones S. G. (1997), *Virtual Culture*, Sage, London.

[6] O'Reilly T. (2005), *What is Web 2.0?*, OReillynet.com. Retrieved from http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html. on October 2008.

[7] IDEA, e-Learning in Local Government: A briefing paper, Improvement and Development Agency, 2002.

[8] Colazzo L., Molinari A., Villa N. (2008), *From e-learning to 'co-learning': the role of virtual communities*, in M. Kendall, B. Samways (a cura di), Learning to Live in the Knowledge Society: IFIP 20th World Computer Congress, IFIP TC 3 ED-L2L Conference, September 7-10, 2008, Milano, Italy, NY: Springer Verlag, 2008, p. 329-338. - (IFIP International Federation for Information Processing; 281). - ISBN: 9780387097282. - DOI: 10.1007/978-0-387-09729-9. - ISBN: 978-0-387-09728-2.

[9] Murphy, K.; Mahoney, S.; Chen, C. Y.; Mendoza-Diaz, N.; Yang, X., Constructivist Model of Mentoring, Coaching, and Facilitating Online Discussions. Distance Education, 26 (3), 341–366, 2005.

[10] Villa N., Colazzo L., Conte F., Molinari A. (2007), *Real communities vs. Virtual communities: structural adaptation of a Learning Management System*, in Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2007, Chesapeake (VA): AACE, 2007, p. 2416-2423. Atti di: E-Learn 2007 Conference, Quebec City, 15th-19th October 2007.

[11] Anderson C. (2006), *The long Tail: how endless choice is creating unlimited demand.* Random House Business Books, N.Y.

[12] Strother, J., "An assessment of the effectiveness of e-learning in corporate training programmes", International Review of Research in Open and Distance Learning, 3 (1), pp. 1–17, 2002.

[13] Bates, A.W., Bates, T., Technology, E-learning and Distance Education, Routledgefalmer Studies In Distance Education, 2005

Revisiting design of learning and mobility services

Bruno Apolloni
Dipartimento Scienze Informazione
Università degli Studi di Milano
Milan, Italy
Bruno.Apolloni@unimi.it

Raffaella Folgieri
Dip. Economia, Management e Metodi quantitativi
Università degli Studi di Milano
Milan, Italy
Raffaella.Folgieri@unimi.it

Giorgio Valle
Dipartimento Scienze Informazione
Università degli Studi di Milano
Milan, Italy
Giorgio.Valle@unimi.it

***Abstract*—**Three Business case studies in University Laboratories on learning and mobility allow experimenting new multidisciplinary expertise, multimedia data management, analysis, distributed collaboration. Software and Hardware engineering practices, cloud technology-based, benefit from information acquired from the Web, increasing abilities in rapid prototyping, in satisfying the needs of customers, validation of pricing policies, marketing strategies and the creation of new professional skills in line with recent trends.
As an alternative to standard software engineering programming, the cloud approach allows to specify the projects' concepts by reaching stakeholders of needed ICT components to implement a prototype relying on SW components available elsewhere. These sample cases refer to: 1) Program for Recovery Insufficient Grades in High-School (PRISC), focuses on the learner and on how proficiency assessed through the grades could be improved; 2) Social Mobility (SM) to reduce dependence on owned cars to satisfy mobility needs exploiting the empty seats available in most cars in metropolitan cities with air pollution problems; 3) Innovative Technology Learning for Humanities at University level (I-DEAS) to create an innovative learning and teaching environment focused on a creative contamination among Humanities and Technology disciplines.

I. INTRODUCTION: CLOUD ENGINEERING

Discussions about Cloud computing on scientific publications and on daily press as well, represent an important indicator from the standpoint of a University environment, for the implications on business innovation and job opportunities for the young graduates. A win-win situation is established leveraging on cloud technology techniques to acquire on the Web what is needed for rapid prototyping of business model concepts, while reducing time-to-market to satisfy the needs of customers.

Technical aspects are widely covered in many International Conferences and research works [1]. Compared to the client-server innovation in the past decades in computing resource allocation, the focus of Cloud Computing now moves to acquiring what is needed over Internet and over mobile devices without any time and location constraints. Cloud computing encompasses any subscription-based or pay-per-use service that, in real time over the Internet, extends IT's existing capabilities. [2] Users and University students, the latter particularly interesting for our learning and teaching aims, are discovering new approach to Business and new emerging interdisciplinary skills needed for successful entry in Job Market, matching competencies and Job demand.

Schematically, Cloud based services require to open a browser, address a provider, open an application, and start using it [3]. The benefit is particularly significant for professionals and small, medium enterprises which benefit of IT services never previously available, enabling them to compete in innovation with much bigger companies because of lower budgets needed and of the possibility offered by the Web in taking advantage from collective intelligence.

Cloud computing present a variety of services, from full-blown applications to storage services: *SaaS*, *Utility computing*, *Platform as a service*, *Service commerce platforms, etc.* and often it also assumed to be a form of green computing (in fact, there are many studies to substantiate the environmental effects of Cloud Computing). Companies are served by a variety of data centres whose resources are very under-used, where the server utilization level is less than 20% by many estimates [4].

This paper illustrates how this cloud approach has engineered two applications related to:

- *education* where users are in one case high-school student with bad grades, and in the other case University students from non-technical University Courses needing technological skills to introduce innovation in traditionally far-from-technology study subjects.

- *mobility* where users are those leaving their house with a car that requires fuel and parking space.

In section II, III and IV we present respectively the vision, state-of-the-art, and the outcomes of projects PRISC, SM, and I-DEAS. In the last section we draw our considerations and conclusions.

II. PROGRAM FOR RECOVERY INSUFFICIENT GRADES IN HIGH-SCHOOL (PRISC)

High School Education is a major concern for Governments since economy globalization and

technological innovation in industrial processes require staff and managers able to anticipate unpredictable market changes. While lower government budgets and economy recession force to prototype new educational formats where ICT and social networking make education more effective.

To deliver a better overall education in High-Schools requires a co-ordinated action on the stakeholders: teachers, students, textbook publishers, school administration board, parents, laboratories and their equipment. [5]

TABLE I

Service	*Provider*	*project*
Cloud platform	Heroku, Amazon EC2, IBM Cloud	PRISC- Sustainable Mobility, I-DEAS
Database	Cloud Database, MongoHQ, Cloudant	PRISC - Sustainable Mobility, I-DEAS
Video encoding	Panda Stream, Zencoder	PRISC, I-DEAS
CDN service	Amazon AS3, Amazon CloudFront, IBM Cloud	PRISC, I-DEAS
Queue system	Amazon SQS, IBM queue	PRISC, I-DEAS
Push Notification	Pusher, PubNub	Sustainable Mobility, I-DEAS
Payments	Paypal, Google Checkout, Amazon Payments	Sustainable Mobility
Mail and SMS notification	Sendgrid, Cloudmailin, Moonshado	Sustainable Mobility
Search engine	Websolr, Elasticsearch	Sustainable Mobility
Social authentication	Facebook, Twitter	Sustainable Mobility, I-DEAS

PRISC strategy focuses on the learner and on how proficiency is assessed through grades. Learners interested in PRISC action are both those with bad grades and those with excellent grades in the last 4 years before either looking for a job or entering the University. The problem addressed is how a student with bad grades in specific subjects succeeds in recovering them in order to proceed to the next higher level class. In this initial phase, PRISC focuses on a few subjects: Mathematics, Physics, Computer Science.

At present bad grade recovery happens either with additional classroom-based teaching organized by the School free of charge to students (in Italy these teachers are paid order of 50 euro/hour) or by private action where parents look for teachers or University students) for private tutoring usually paid cash with no receipts. In Italy this process brings to one of the worldwide most expensive School systems ranking Italy as 37th Country out of the first 40 main Industrial States in the world in performances proposed by OCSE and PISA.

The overall objectives of PRISC project, while taking care of the students' recovery of bad grades when in High School, are reached by the following actions:

1) enhancing topic understanding focuses on exercises related to text books adopted for the school year;
2) exercises and tests are managed in an interactive WEB environment on screens of mobile telephones or PC;
3) a students' social network driven by students with excellent grades supports this extracurricular learning activity;
4) all exercises and explanations are indexed and stored in a central database;
5) assessment of better students' study performance happens in the same school environment where a bad grade was received at the beginning;
6) this "social learning community" is an obvious target for exploitation of recent digital social networks like: LinkedIn, Facebook, Plaxo, etc [7] established for other purposes;
7) the same happens with telecommunication technologies, such as video-telephone group calls among group of students and teachers.

All the procedures are managed on the cloud by definition of their characteristic data (see Table 1). The resulting system is inherently distributed (users do not know where data and procedures are stored), where the central management interacts with a set of web-services provided by the cloud interface.

III. THE SM (SUSTAINABLE MOBILITY)

The SM strategy focuses on exploiting the empty seats available in most cars in metropolitan areas with air pollution problems [8]. Its implementation assembles system components acquired on the Web leveraging on a cloud technology approach for

- short range contacts for prompt satisfaction,
- GPS ride surveillance.

Universal awareness of the necessity for ITC interventions that combine respect for people's mobility needs with protection of health and the environment -- by avoiding that the traffic becomes a major source of pollution, energy consumption and everyday personal stress -- dates at least thirty years ago.

Meanwhile, many ride sharing systems have been experienced both in Italy and abroad. An example of a generalist service (carpooling, carsharing, ridesharing) operating at national level, as well as internationally, is RoadSharing (http://www.roadsharing.com/it/). The typical drawback is that transactions are not real-time. On the contrary, the organization Jungo (http://www.jungo.it/) offers a hitchhiking type of extemporaneity -- thus meeting in view of the rider with the driver, with the additional benefit of security at the cost of a somewhat cumbersome procedure (annual subscription, membership card with photo, sms message at the beginning of transition, etc.) at a fixed rate (20 cents + 10 cents per kilometer fixed). A more technological and less procedurally heavy experimentation is underway in the city of Seattle (http://go520.avego.com/st-pilot/) along the route 520 which runs through much of the city and connects it to the suburbs. Transactions are done via smartphones and security is ensured through the phone's GPS. However, the rides are booked and carried out between fixed points of gathering and landing along the route, and the business model is based on the burden of the transition.

It is common opinion that the ride-sharing problem has to

be faced within a global vision, indeed, as witnessed by the born of planning and coordination institutions, like the Ridesharing Institute (http://ridesharinginstitute.wikispaces.com/). However, studies like "Data on users of the car pooling and ride sharing system" (http://www.echoaction.net/pdf_files/Final_Report_Mobility.pdf) promoted by the European community in the frame of Echo Action, show that we are still in the early stages. Moreover, a long record of dynamic time sharing systems which have been implemented but are not in operation may be found on the webpage http://dynamicridesharing.org/projects.php. Besides a valuable list of factor possibly inhibiting the success of these systems, the analysis quoted in http://dynamicridesharing.org/inhibitors.php holds: *The feeling amongst practitioners is that schemes that are truly dynamic succeed or fail on the critical mass issue. While other inhibitors (such as incentives and attitudes) are important, critical mass is just that: critical.*

The Sustainable Mobility Project aims to reduce the use of private cars through a community of Social Mobility which meets the mobility needs of its members. The system of Social Mobility is implemented as a framework of "Internet of Things" (IoT) in which the object of transactions are passages (rides) provided by car drivers (OFT) to people who request them (RDT). These players (ATT) are members of a social community and are terminal nodes of an ecosystem that provides web services for an optimal management of ride transactions with respect to different criteria: i) travel cost, ii) ride safety, iii) fast procedures, iv) impact on traffic congestion and pollution.

The typical transaction between community members implies the following steps:

1. a RDT member, say ID_0, who needs to go from place A to place B casts a call to the cloud through its mobile terminal looking for someone who wants to give him/her a ride (with fee or for free) within 10 minutes.
2. OFT members ID_1, .., ID_n, who are in the neighborhood of ID_0 position, declare their willingness to cover all or part of the route.
3. The part of the ecosystem living in the cloud proposes a number of ID_i (or ID_i sequences), on the base of their trustworthiness and on the options selected by ID_0 (with fee or for free, complete or stepwise, urgent or with no hurry, and so on).
4. ID_0 chooses one of the proposed ID* (or a sequence of ID*s).
5. The ecosystem assigns the chosen ID* to the route, it assists the meeting of ID_0 and ID*, follows via gps the development of the transaction, updates accordingly the credibility and credit of the involved players as well as the database of rides and other auxiliary data.

On the social front, we are going to implement in daily practice a virtual community designed for the transport of its members. It focuses on implementing a more efficient urban mobility system with decreasing costs (for those who leaves at home their vehicle) and positive environmental impact (for the reduction of used vehicles). The system being built both summarizes past experiences that have been tested in Italy since the mid-90s and offers different solutions with respect to current ride sharing systems: impromptu requests, immediacy of their fulfillment, basically for free services.

In terms of technology, the solution we are implementing meets the widespread availability of geo-information technologies owned by a wide segment of the population and people's general familiarity with social networks. With the spreading of low cost smartphones (about 100 €), a wide hybrid network has come to life. It is a network with wireless terminal connections, able to convey always and everywhere information related to the geographical position of the terminals themselves, with a less than ten meters precision. In such networks the information is automatically linked to the "id" of the person who generates it, with a wealth of details (e.g., photos and other ancillary information) that s/he decides to share. Likewise, the IoT in which this network is immersed, is able to provide an abundance of boundary conditions, which go from simple atmospheric conditions to a whole series of data as usual equipment of smart cities.

Methodologically, the project consists of a front-end with data that enables the transactions, and a back-end that handles the sociological, economic and transport related issues, in order to optimize the transactions.

The main ingredients of our paradigm are:

1. almost instantaneous connection of the user with the cloud to: i) place ride calls, ii) communicate his/her identity (in terms of ID, automatic profile, and photos) and GPS coordinates, iii) negotiate ride offers and conclude transactions. The client side of these functionalities will reside in a smartphone app or surrogates,
2. time/space locality of the transaction. A call concerns a ride to start in a short time (say, 10 minutes). Only drivers who are close enough to the ride requester (say, 2 kilometers away in a urban area) may offer a ride. A call expires in any case (satisfied or not) once the above time has elapsed.
3. a sophisticated platform in the cloud with two main functions:
   a) suitable sorting of the ride offers on the requester screen. This will depend on a proper combination of various factors, such as driver distance, his/her credits within a recommendation system, cost of the ride, etc.
   b) transaction safe. Identities and photos of transaction actors plus tracking of the ride from beginning to (agreed upon) end will underpin an intelligent ride monitoring.
4. a wise balance between individual and agglomerative solutions. On the basis of both ride histories collected by the platform database and other statistics on the field, as a result of deep researches on modeling and planning the overall urban and suburban multimodal transportation

plan on specific test-sites, and with the commitment to keep the user at the center of any decisional process,

1. The socio-economical conditions will be identified that determine the choice among the different person transportation modalities and the economic-systemic factors fixing the convenience borders between them,
2. Proper politics will be studied to promote our ride sharing in all modalities, to lower congestion and pollution charges,
3. Technics of multimodal resource planning and optimization will be employed to rethink on-demand transportation plans in light of the real-time information on demand-offer and satisfaction feedback made available by our webservice system,
4. A business model will be devised to render our innovative human-centered approach to the personal mobility that is feasible for the community and rewarding the private and public institutions supporting it without direct charge on the single transactions.
5. an equally friendly and sophisticated approach to the services offered by the modern Internet infrastructure, where the users are called to improve the rational aspects of their moves, getting a companion improvement to their style of life and cultural habits
6. an involvement of user in co-designing the system within the open-source philosophy, so that any person may contribute to the improvement of the functionalities of the system on the basis of his/her personal ICT skills.

IV. THE I-DEAS PROJECT (ICT-HUMANITIES CREATIVE LAB)

The PRISC and the SM experiences on Cloud Computing inspired the I-DEAS project, which challenge was to fill the gap between humanistic and technological skill in the context of the Faculty of Economics, Political and Sociological Studies at University of Milan, shaping new study course profiles, so that at the end of their University careers graduates may be not only literate in sociological, economic, business and political disciplines – but also in technology, creativity and innovation. The project argues that the considered students can develop creativity and innovative approaches by merging humanistic and technological skills.

The project started by data and suggestions collected by the Stella and by the Excelsior investigations, performed respectively by CILEA and by UnionCamere, highlighting occupational data, graduates skills and Market needs. This information has been integrated with those obtained by a survey submitted to a sample of Industries and Companies, asked to explain current and future probable skills needed by our students to enter efficiently in the Job Market. All the collected suggestions confirmed a strong orientation towards the integration of humanistic and technological skills and the need to cover the current gap between humanistic and scientific disciplines, with particular focus on ICT.

The current situation in Humanistic University courses shows a general difficulty in matching technology and Humanities during University studies. Within the I-DEAS project, the attention has been focused on looking for a solution to involve students in practical technological experiences [14] that at the same time do not distract them from the main objectives of their studies. Our point of view focuses on the fact that the Humanities are traditionally kept away from technology studies, with the consequent risk to lose not only the technological skills, but also not to encourage an exchange and interoperability among diverse disciplines – an encounter that would instead create a wealth of new ideas, progress and innovation in all fields.

Teaching experiences matured in previous e-learning projects carried at our University [15, 16], demonstrated that students are much more motivated if involved in research experiments combining practice with learning, not only limiting to the final thesis discussion, but contributing with their ideas and points of view during their entire university careers. This is how they can develop multidisciplinary skills, whose results translate into a new cooperative learning approach that merges the above mentioned issues, and that drives the students' curiosity in discovering new disciplines.

With the aim of giving a real Job market orientation to the "I-DEAS" project, it has been investigated the possibility to involve leading technological industries in facing the challenge to set up a creative ICT laboratory for our faculty students, merging different disciplines and innovative technologies. The term "creative" have been adopted to indicate a new different approach to ICT and technology, designed specifically having in mind humanities and their approach to disciplines. We do not believe in substantial differences in way-to-face disciplines and subjects adopted by scientific and non-scientific skilled people. On the contrary, we believe that, if we could reach the root of reasoning mechanisms, we could find a similar way to proceed, based on similar problem solving approaches but just on different expressive grammars used to explain the same concepts, developed as a consequence of different cultural references acquired in school experiences. On this basis they have been conducted some experiments in teaching different disciplines to students, starting by similar concepts but applying them on different case studies. The first steps in the described direction consisted in joining in the same class students from Political, Sociological and Economics students with students attending the Virtual Reality course from IT, Digital Communication and Music Technology courses. The aim was to confuse them, submitting them some interdisciplinary problems to solve with problem solving and decision making approaches applying a sort of "brainstorming" in which each "kind" of student could show to the others the own methods acquired in different disciplines. Surprising students (not the authors), students

discovered to have many common knowledge, even if different competencies, in methodologies field-related.

In the so created cross-disciplinary course, open to all students, different techniques and methodologies to apply the problem solving approach have been taught. Disciplines goes from the representation of a problem in different fields, such as biology, economics and information technology, to semiotics and languages, always explained in a theoretical and practical approach, showing to students how to apply the learned lesson to their own disciplines and what information technology tools could be used to translate the theory into practice. Students are always asked for active participation and direct involvement during the lessons, sharing humanistic and scientific knowledge. The aims are to destroy previous preconception, to demonstrate them that the concept of creativity is cross-disciplinary and not necessarily linked to the Humanities. In this context, the Cloud Computing approach has been essential. The students soon begun to enjoy and get surprised by their ability and by the possibilities offered by joining forces among different disciplines and by the potential of the Web and the Cloud computing. Students passed from the initial "silent brainstorming" to really active and participative involvement in lessons and lab activities. The following step has been, as anticipated, to invite students to engage themselves in engineering understanding of technology. To this aim it has been evident the need to create a permanent cross-disciplinary engineering lab, based on the Cloud Computing approach.

The idea has been presented as first to IBM University Relations in the form of a Country Project, which was approved in September 2011, winning the competition and being granted the initial resources that enable start the lab's activities thanks to the technological platform implemented by IBM cloud computing methodology and tools, leveraging on the complimentary IBM Academic Initiative (https://www.ibm.com/developerworks/university/academic initiative/) resources.

The primary objective of this lab is therefore that of spreading the knowledge of ICT and Web Science (http://webscience.org/home.html) among the students [17], thus encouraging new ideas, business research, cross-sectional expertise, connection with the business world and excellence in both academic achievements and technical knowledge acquisition. Also, being based on cloud computing, the aim is facilitating the exchange of knowledge, opinions, ideas and cultures throughout contacts with other Universities all over the world (i.e. study sessions and cooperation projects remotely conducted).

The Cloud Computing platform, currently in use, will even more help to create a stimulating, amusing and enjoyable environment through which students, but also Researchers and Professors involved in the project, will exchange ideas, will start projects and will create new job profiles and business opportunities by involving their peers all around the world. A physical meeting room has been equipped, consisting of advanced hardware and software communication tools, providing also a virtual space accessible from the lab (and home). The physical space consists of a creative lab located in a meeting room equipped with VoIP, video conference tools, touch screen monitors, cooperative work environment and wiki.

The following step concerns the decision to introduce all students to programming languages through a graphical user environment, LabView™ (by National Instruments) suited for both programming and engineering. This choice has been done to invite students to mix their competencies, trying to create new IT or engineering solutions in their fields of study, using their own knowledge and creativity and demonstrating them that also students from Humanities studies could create, with the right tool, innovative technological solutions for purposes matching new-era needs in various disciplines. For this last reason, we also kept in contact with NI (National Instruments), a leading industry whose open graphical programming software (LabView™) and modular hardware could allow also to non-technicians to design and test electronic devices to create new solutions in communication tools, gaming systems or medical devices. National Instruments enthusiastically joined the initiative and currently students can develop ICT and engineering solution using the LabView™ graphical platform.

Figure 1 shows the logical organization of the "I-DEAS on the cloud" lab and the corresponding technical architecture.

The so conceived lab provides the students with a solid, updated and versatile ICT education, which is the key to their future careers. At the same time, it offers the knowledge of the right instruments for cross-cultural communication, creativity and learning. Of course, in line with tradition, students are encouraged to gain highly-developed skills in criticism, necessary to perform solid and innovative work in their studies.

In conclusion, the "I-DEAS on the cloud" project aims at creating an experimental site, considering that students wish to create technical, and often engineering, solutions and to apply those solutions to social, cognitive or marketing studies. Thanks to the involvement of Industries and both Professors and Researcher expert in marketing, business, cognitive science and ICT disciplines, the lab aims at stimulating the growth of students' know how to master technologies and be carriers of change. Moreover, the lab will support the development of interdisciplinary research, basic and applied or empirical and the development of empirical computational thinking.

This project represents an opportunity to foster our faculty and in general Academic research community relationships with IBM, NI (and other future Companies), both on a local level and internationally, increasing future development and granting our students the opportunity to also get in contact with innovator industries participating in the initiative.

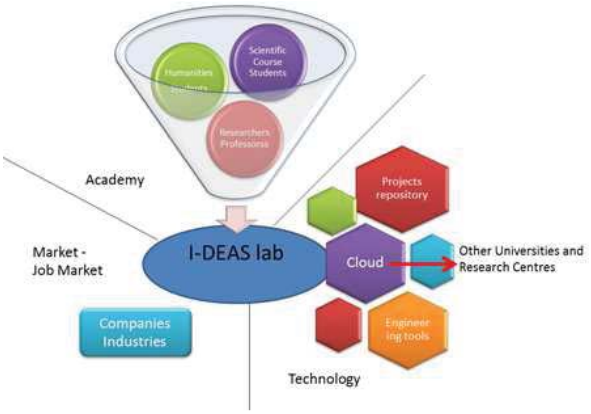


Figure 1: O rganization and technical architecture of the "I-DEAS on the cloud" lab.

V. CONCLUSIONS

In the ICT epoch many services must be reconsidered and redesigned to remain at the top of quality and efficiency. The instances we prospect in this paper frames new services within meaningful advances in the culture, with special emphasis on:

a) The user is asked for conceptual engagements to satisfy his/her needs. It is not a matter, for instance, of simply reaching the bus stop or his/her own car. She/he must pay attention to issue his/her request to the community, consider the offers and perform ancillary sagacity. On the one hand, they are tasks familiar to the user in other contexts, for instance booking a flight via Web. On the other hand, the user tends to face some additional complication, provided the revenue is tangible. In doing so, she/he enhances his/her intellectual ability.

b) The creative capacities of more ICT-smart people are promoted, since they find wide room to create geo-located applications.

Other rooms are open in the people psychology where the confidence perception re the environment of services into which they are immersed and the reliability of their delivery mate with the inclination to start and maintain shared behaviors.

Moreover, the new services promote job creation, due to the fact that the questioned services have a server located in the cloud and must be adapted to the local needs. Cloud computing is a paradigm of transversal services that are delivered by platform integrating servers and apps which guarantee long life cycles of the solutions made available by the tec hnology. The so structured services constitute proprietary informatics ecosystems competing with one another, so that the user may select the most convenient one. This is a source of notably job opportunities re both the client side, since many apps with increasing performance to be installed on the user terminal may appear on the marked, and the server side. As mentioned earlier, the web platform of these services is burdensome, as it requires tuning as a function of local condition and ride safety monitoring. Rather than a drawback, this constitutes a benefit since it requires new jobs to handle not only the mentioned tasks but also a wide range of spinoff services, such as advertising, database, etc., which makes the system economically feasible. In this respect, the cloud features allow young entrepreneurs starting their activities without a lot of capital. Indeed the cloud services philosophy is that no costly client hardware is necessary and users pay for what they consume.

REFERENCES

[1] "IEEE International Conference on Cloud Computing (CLOUD)".Thecloudcomputing.org http://www.thecloudcomputing.org.

[2] Luis M. Vaquero, Luis Rodero-Merino, Juan Caceres, Maik Lindner "A Break in the Clouds: Towards a Cloud Definition" ACM SIGCOMM, Volume 39, Number 1, January 2009

[3] Jörn Altmann, Björn Rupp and Pravin Varaiya : Internet demand under different pricing schemes, pp 9 - 14, ACM Conference of Electronic Commerce, 1999

[4] R. Smith, "Power companies consider options for energy sources," The Wall Street J., p. A.10, Oct. 29 2007.

[5] G. Valle, & F. Epifania, (2008). An Interactive VideoJournal to complement Learning & Training. In Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2008 (pp.1164-1169). Chesapeake, VA: AACE.

[6] F. Epifania, (2007). An experience in distance learning for university and secondary school. In G. Richards (Ed.), Proceedings of World Conference on E- Learning in Corporate, Government, Healthcare, and Higher Education 2007 (pp. 6786-6791). Chesapeake, VA: AACE.

[7] F. Fovet, "Impact of the use of Facebook amongst students of high school age with Social, Emotional and Behavioural Difficulties (SEBD)" 2009 Proceedings - Frontiers in Education Conference, FIE , art. no. 5350786

[8] http://www.erideshare.com/

[9] A. L. Bristowa; A. D. Pearmana; J. D. Shiresa. An assessment of advanced transport telematics evaluation. Transport Reviews, 17, 3 , 177 – 205, 1997.

[10] G. Elofson, P. M. Beranek, and P. Thomas. An intelligent agent community approach to knowledge sharing. Decision Support Systems, 20(1):83–98, 1997.

[11] D. D. Brockmann, L. Hufnagel, and T. Geisel. The scaling laws of human travel. Nature, 439:462–465, 2006.

[12] B. Apolloni, S. Bassis, D. Malchiodi, and P. Witold. The Puzzle of Granular Computing, volume 138 of Studies in Computational Intelligence. Springer Verlag, 2008.

[13] B. Apolloni, S. Bassis, and A. Zippo. Processing of information microgranules within an individuals' society. Human-centric Information Processing Through Granular Modelling, pages 233–264, 2009.

[14] B. Taraghi,, M. Ebner, G.Till, H. Mühlburger, 2009. Personal Learning Environment a Conceptual Study. In Conference ICL2009, 997-1006.

[15] M.A. Kazancigil, G. VALLE, R. Folgieri, 2009. Information technologies incentivizing scholastic fruition, Proceedings of Society for Information Technology and Teacher Education International Conference, EdITLib: AACE.

[16] R. Folgieri, G. Valle, 2004. Continuous Learning Relaxing Time & Location Constraints In a Regular Curricula Degree Offered by University of Milan. World Conference on Educational Multimedia, Hypermedia and Telecommunications, 5166-5171.

[17] M. Croitoru, S. Bazan, S. Cerri, H. Davis, R. Folgieri, C. Jonquet, F. Scharffe, S. Staab, M. Vafopoulos, S.White, 2011. WSCD: Negotiating the Web Science Curriculum Development through Shared Educational Artefacts. ACM WebSci '11, 3rd International Conference on Web Science, Koblenz, Germany.

Blended learning in Continuing Medical Education

Francesco Epifania
Dip, Informatica
Università degli Studi di Milano
Dip. Elettronica e Informazione
Politecnico di Milano
Milan, Italy
Francesco.epifania@di.unimi.it

***Abstract*—** The aim of this project is to realize a new multimedia interactive web system for digital learning that, using a different and innovative concept, can spread and aggregate the knowledge using different crucial component: 1) Professional knowledge: experts create contents to be disseminated over communication channels 2) Multichannel 3) Social web. This Interactive Multimedia Web System for Digital Learning will be tested and evaluated in permanent training course as in medicine (ECM) whit 10000 users. In these years, Pharmacy Profession has become the base with most nationwide spread for the professional updating of Italian pharmacist. From 2010, Medical Evidence will also offer its courses in Spain since the training schedule of Pharmacy Profession has been accredited nationwide by the Permanent Training Commission of the Health National System of Spain. These courses are executed remotely, requiring the only presence at the last lesson of the course. Completing the course, participants will receive credits for attending permanent medical training.

Keywords: - E-learning 2.0 - Digital learning – ECM

I. INTRODUCTION AND PROJECT MOTIVATION

Medical Evidence is the division of Marketing & Telematica which from 1994 elaborates high profile professional updating and training programs, spreading the scientific content of great importance aimed at the different public-health categories in Italy. Furthermore, Medical Evidence publishes from 1998 the digital magazine called Pharmacy Profession whose content comes from the monographic courses of Drug, approved by the American College of Pharmaceutical Education.

The 6.500 plus pharmacists registered to the Pharmacy Profession Training 2009 confirm the success of this training program that, for 12 years from now, secures a constant upgrade in the area. The number of registered participants, that represents the 13,5% of Italian professionals, increases for more than 4 years and, considering the number of pre-registrations during the last finals sessions (more than the 55% of the participants), is yet set to increase next year. However it is not about a trend limited to the pharmacists category: in 2009, Medical Evidence has organized more than 11 FAD and FAD Blended courses that have involved altogether more than 10.000 professionals.

This trend is shown in next table, and underlines the increasing success of this FAD approach.

One important assessment element is the provenance of the registrations, not limited to just the Region of Lombardy, but distributed through all the national territory. This is also due to the low cost of the Medical Evidence training programs (210 € and in only one case 250 €), that have been conceived to allow the dissemination of information and scientific know-how to an increasingly large public. Even more relevant is the fact that the 95,4% of the registered participants has completed the training program obtaining the expected credits (35 or 40), thanks to the keen assistance of the Contact Center and the Organizational Administration ECM, ensuring continuous coaching and giving attention and support to participants in difficult with the lessons.

Main goal of this project is to realize a new multimedia interactive web system for digital learning [1] that, using a different and innovative concept, can spread and aggregate the knowledge using different crucial component: 1) Professional knowledge: experts create contents to be disseminated over communication channels 2) Multichannel: users benefit by access in digital form Web, ITV, smart-phone, podcast, iPad 3) Social web: once a community is created, it is possible that users participate in a discussion with the expert. This discussion could generate new knowledge with educational impact for other users . What described here above could be realize developing a new generation of Social Learning Management System ; this is an evolution of existing the Leaning Management System (LMS). The initial phase of the project will investigate state-of-the art on scientific literature and will provide a benchmark for a first prototype applications. This phase will establish the outcomes of the LMS. The second phase will produce the statements of the key requirement as follow: 1) Efficient contents' management 2) Dissemination over communications channels 3) Support to users' communities 4) Create e-learning recommender system [4]. The third phase will consist of the development of system and will include the test using a community of selected users. In the last phase will be done a accurate evaluation using a scientific base about efficacy of prototypical and of real experimentation. The multimedia interactive system for Continuing Medical Educationit's available at the links www.farmasimposio.it ; http://www.cardio14.it/;www.nefro14.it .

II. LEANING MANAGEMENT SYSTEM

The aim of this project is to realize a multimedia interactive system for Continuing Medical Education, using open source systems. This system allows learners to enrol in several courses upon payment to achieve ECM credits and a certificate recognized by the Ministry of Health.

The purpose has been achieved creating a Learning Management System (LMS), using web-oriented languages (PHP, SQL, JavaScript, HTML and CSS) and customizing the e-learning open source "Moodle" platform, in the future the system could be transferred in the cloud [2]. LMS allows teachers to create and manage on-line courses with huge possibilities of interaction with the students. Teachers and experts can upload training aid. This material is spread with several technological channel (web, mobile, smartphone, Iphone, tablet, Ipad…) Users can view their results, achieved answering to quizzes, and check their reached ECM credits to receive the course certificate.

We followed the reported methodology and we divided the work in four stages. [5]. Several modules have been used and created to customize the existing platform. The most important are: site navigation, courses' structuring, news, recent news, next events, planner, events' reservation, certificate, quiz, feedback, course's evaluation. Users' learning is checked by quizzes and surveys. The student will be able to enroll to a "Finale session" after completing the whole training course and finally to download his own certificate created by the system after completing the session.

III. THE FINAL SESSIONS OF THE COURSE

In 2009 all the final sessions of the course have been held in the months of October and November in 26 provincial capitals for a total of 100 sessions (about 40 sessions more than those of the last year) at the disposal of the participants.

The convenience of the access to the course, the usefulness of the content of the daily clinical practice and the minimal request of resources for the achievement of a substantial number of ECM credits render this undertaking a very successful one, answering the need for an increasingly big number of pharmacists and proving this formula more effective than the traditional residential courses.

*The plus of the pharmacy professions training program: 1)*Qualified scientific content of international relevance, validated by the American council of Pharmaceutical Education and, for Italy, by the Scientific Board of Medical Evidence.2)35 ECM credits with only one compulsory presence of 4 hours in the classroom.3)Optimization of the training time with online updates on the desired time and place 4) Reduced registration fees.

Report of the final sessions of Pharmacy profession 2009

The following table show the final report based on a sample of 2042 participants from 25 course sessions. The questionnaires on the evaluation of the event filled in by the participants at the end of the course, confirm the success of the global effort, highlighting that the significant majority of the registered participants has considered the didactic model effective, evaluating very positively both the quality and the relevance as well as the excellent organization.

EVENT					EVALUATION
Overall Relevance	Irrelevant 1	Little relevant 8	Quite relevant 271	Relevant 1229	Very relevant 542	2042
Quality	Poor 0	Mediocre 12	Satisfying 183	Good 1344	Excellent 503	2042
Effectiveness	Ineffective 1	Partially effective 23	Quite effective 428	Effective 1198	Very effective 393	2042
Quality of the organization	Poor 1	Mediocre 15	Satisfying 209	Good 1127	Excellent 691	2042

Table 2: final evaluation report

Moreover this year, the final session in the classroom has received the general approval of the participants that have welcomed the use of cartoons in the presentation of the didactic material rendering it not hard at all. Medical Evidence organizes ECM courses for diverse medical and sanitary specializations, such as: Oculists, Resuscitators and Anaesthetists, Nurses, General Medicine Doctors, Ophthalmologists, Veterinarians.

IV. CONTENTS OF THE WEB PLATFORM

The Editorial Plan is made up of scientific material, carefully selected and drafted in the form of monographic lessons by the Scientific Board of Medical Evidence, that is an international interdisciplinary team of specialists coming from important Centres of Excellence or Universities for each discipline. A great attention is given to the quality of the didactic material. In fact, in many cases, the didactic material is selected as sources internationally authoritative and scientifically independent such as Drug Topics, a specialized magazine published in the United States by the Advanstar Communication Society and it is approved in its original version by the American Council on Pharmaceutical Education.

Accessing with the personal user name (login) and password to the Internet platform, each pharmacist is able to:1)Look up or download the lessons contained on the updating monographs of Pharmacy Profession (PDF files) 2) Fill in the questionnaires related to the online lessons (multiple choices questions for each subject). 3) Find out the tests' results, redoing each questionnaire until getting at least 80% of the correct answers.4) Visualize the personal record summarizing the obtained credits up to that moment. 5) Book the (compulsory) participation in the final session, choosing the available offices for the final session. [3]

V. TRAINING SCHEDULE

Concerning the schedule of the Training Plan, the Provider has organized 10 courses of 12 (83,3%). One of the courses, Optician Profession, has been in any case completed although it has been excluded of the compulsory ECM during the progress of the course, with ordinance of the Department of Health, in other words removed from the definitive events.

Course Target	Overall N. of Participants	Participants who *have* acquired the credits	Participants who *have not* acquired the credits
Pharmacy Profession	5.355	5.085	269
Oculist Profession	1.132	1.053	79
Veterinarian Mission	197	186	11
ATI14 – Updating in Anaesthesia and Intensive Therapy	587	570	16
CARDIO14 – Updating in Cardiology	144	134	10
NEFRO14 – Updating in Nephrology	41	36	5
Management of partners in pharmacy	797	346	451
Everyday wellness	1.504	1.112	392

Table 3: Number of participants for each event, with or without credits

Gathering of statistical data for the improvement of the training offer:

Pharmacy Profession: The course has been structured in two phases: 1) Online preparatory session with 12 lessons 2) Final session for the examination. A total of 69 editions of the final session has been organized, in which 5.085 Pharmacists have participated from 5.355 professionals previously registered (95,0%). During the final session the Customer Satisfaction questionnaire has been provided, expected according to the normative, the questionnaire is kept at the Provider office. The summary of the results has delivered an indication equal to the 85,1% of evaluations assessed as "Good" or "Excellent". The Tutors that have organized each final session have obtained from the participants the 91,8% of evaluations with the "Good" or "Excellent" indication.

Oculist Profession: The course has been structured in two phases: 1) Online preparatory session with 12 lessons 2) Final session for the examination. A total of 12 editions of the final session has been organized, in which 1.053 Oculists have participated from 1.132 professionals previously registered (90,2%). The summary of the results of the Customer Satisfaction questionnaire has delivered an indication equal to the 88,4% of evaluations assessed as "Good" or "Excellent". The offices for the Final Session have been connected through video-conference to ease the interaction between the participants and the Teaching Staff. The Teaching Staff that has organized each final session have obtained from the participants the 91,7% of evaluations with the "Good" or "Excellent" indication.

Veterinarian Mission: The course has been structured in two phases: 1) Online preparatory session with 12 lessons 2) Final session for the examination. A total of 9 editions of the final session has been organized, in which 186 Veterinarians have participated from 197 professionals previously registered (89,3%). During the final session the Customer Satisfaction questionnaire has been provided, expected according to the normative, the questionnaire is kept at the Provider office. The summary of the results has delivered an indication equal to the 79,6% of evaluations assessed as "Good" or "Excellent". The offices for the Final Session have been connected through video-conference to ease the interaction between the participants and the Teaching Staff. The Teaching Staff that has organized each final session have obtained from the participants the 88,6% of evaluations with the "Good" or "Excellent" indication.

ATI14 – Updating in Anaesthesia and Intensive Therapy:The course has been structured in two phases: 1) Online preparatory session with 14 lessons 2) Final session for the examination. A total of 10 editions of the final session has been organized, in which 570 Doctors have participated from 587 professionals previously registered (97,1%). During the final session the Customer Satisfaction questionnaire has been provided, expected according to the normative, the questionnaire is kept at the Provider office. The summary of the results has delivered an indication equal to the 79,4% of evaluations assessed as "Good" or "Excellent". The offices for the Final Session have been connected through video-conference to ease the interaction between the participants and the Teaching Staff. The Teaching Staff that has organized each final session have obtained from the participants the 85,3% of evaluations with the "Good" or "Excellent" indication.

*CARDIO14 – Updating in Cardiology:*The course has been structured in two phases: 1) Online preparatory session with 4 lessons 2) Final session for examination A total of 5 editions of the final session have been organized, in which 134 Doctors have participated from 144 professionals previously registered (93,1%). The summary of the results of the Customer Satisfaction questionnaire has delivered an indication equal to the 89,5% of evaluations assessed as "Good" or "Excellent". The offices for the Final Session have been connected through video-conference to ease the interaction between the participants and the Teaching Staff. The Teaching Staff that has organized each final session have obtained from the participants the 92,4% of evaluations with the "Good" or "Excellent" indication.

*NEFRO14 – Updating in Nephrology:*The course has been structured in two phases: 1) Online preparatory session with 4 lessons 2) Final session for the examination. A total of 5 editions of the final session has been organized, in which 36 Doctors have participated from 41 professionals previously registered (87,8%). The summary of the results of the Customer Satisfaction questionnaire has delivered an indication equal to the 90,8% of evaluations assessed as "Good" or "Excellent". The offices for the Final Session have been connected through video-conference to ease the interaction between the participants and the Teaching Staff. The Teaching Staff that has organized each final session have

obtained from the participants the 90,4% of evaluations with the "Good" or "Excellent" indication.

Management of partners in pharmacy: The course has been delivered in e-learning modality, with the aid of multimedia lessons in self-learning and support from the forum for the participants. At the end of the course an examination questionnaire about the quality of the course has been provided, consisting of 17 direct questions and 3 open questions for a total of 20 questions. The adopted system of examination of the direct questions is a Likert scale of 4 or 5 modalities. The answers can be put together in four groups: G.1 Accessibility and usability of the platform; G.2 Support (help desk, tutorship and level of communication); G.3 Effectiveness of the methodology (Level of involvement of the course and pertinence of the subjects, level of difficulty); G.4 Quality of the didactics (evaluation of the content of the course, also related to one's expectations).

For all groups the percentage of positive and very positive answers has been over the 95% with a global percentage of satisfaction of the 97%. It has not been proposed the question if the presentation of the course subjects has been balanced and objective, no participant has answered negatively because the theme of the course does not concern medical subjects but administrative. From the open questions it has been appreciated the didactic modality. According to the participants it resulted flexible and adaptable to their time for understanding and learning. Moreover, it has been appreciated the expositive clarity and the representation of knowledge, judged direct and effective. The results of the satisfaction questionnaire confirm that the methodology of planning, delivery and support are absolutely well-balanced.

Everyday wellness: The course has been delivered in e-learning modality, with the presentation of downloadable didactic material and a forum involving participants. At the end of the course an examination questionnaire about the quality of the course has been provided, consisting of 17 direct questions and 3 open questions for a total of 17 questions. The adopted system of examination of the direct questions is a Likert scale of 4 or 5 modalities. The answers can be put together in three groups: G.1 Accessibility and usability of the platform; G.2 Support (help desk, tutorship and level of communication); G.3 Quality of the didactics (evaluation of the content of the course, also regarding one's expectations)

For all groups the percentage of positive and very positive answers has been over the 97% with a global percentage of satisfaction of the 98,5%. To the explicit question if the presentation of the course subjects has been balanced and objective, no participant has answered negatively.

VI. CONCLUSIONS

The 6.500 plus pharmacists registered to the Pharmacy Profession Training 2009 confirm the success of this training program that, for 12 years from now, secures a constant upgrade in the area. The number of registered participants, that represents the 13,5% of Italian professionals, continues to grow for more than 4 years and, considering the number of pre-registrations during the last finals sessions (more than the 55% of the participants), is yet set to increase next year. However it is not about a trend limited to the pharmacists category: in 2009, Medical Evidence has organized more than 11 FAD and FAD Blended courses that have involved altogether more than 10.000 professionals. The registered results gave excellent scores both in terms of increasing number of participants and in participants' performance (learning results).

The evaluation surveys of the platform given to the testers show a positive response. We conducted several usability tests and the system appears efficient and effective. Thanks to its simple structure, there hasn't been found any complication in the pages navigation. In fact, we improved some pages, where the learner could be disoriented, in order to make the system easy and pleasant.

New modules have been introduced to make the system utilization easier, allowing the user to have all the necessary instruments to achieve his purposes. Furthermore, downloading the lessons documentation is an easy and quick process. Thanks to the multimedia contents (audio, video and synchronized slides), users live a learning experience similar to a videoconference lesson and users have the advantage to read and study a certain lesson every time he wants. Learners can directly ask teachers explanations using a messaging system as e-mail, or better, an integrated forum through which learners can share impressions and opinions with other users in the course.

The success of the initiatives encourages new frontiers in investigating effective learning to improve quality and long-lasting memory. The FAD structure and the learning sessions organization allow to experiment new techniques for long-lasting learning, such as those based on priming and new advanced learning methodologies [6, 7].

REFERENCES

[1] Francesco Epifania. *Design And Development Of Multimedia Interactive Systems For Digital Learning*. In: International Conference on Distributed Multimedia Systems (DMS 2011). Italy, (pp. 238- 241).

[2] G. Valle , B. Apolloni, F. Epifania. *Cloud Design for Learning And Mobility Services*, 2011 . (pp181- 184). In: In proceedings International Conference on Distributed Multimedia Systems (DMS 2011). Italy,

[3] Nicoletta Di Blas, Paolo Paolini, Aldo Torrebruno: *Innovative Technologies and Education - Two Successful Case-studies*. CSEDU (2) 2010: 341-348

[4] Paolo Cremonesi, Francesco Epifania, and Franca Garzotto. 2012. User profiling vs. accuracy in recommender system user experience. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (AVI '12), Genny Tortora, Stefano Levialdi, and Maurizio Tucci (Eds.). ACM, New York, NY, USA, 717-720. DOI=10.1145/2254556.2254692 http://doi.acm.org/10.1145/2254556.2254692

[5] Francesco Epifania. *An experience in distance learning for university and secondary school.* In G. Richards (Ed.), Proceedings of World Conference on E- Learning in Corporate, Government, Healthcare, and Higher Education 2007 (pp. 6786-6791). Chesapeake, VA: AACE.

[6] Wiggs, C. L., Martin, A.. (1998). Properties and Mechanisms of Perceptual Priming, Current Opinion in Neurobiology, 8(2): 227-233.

[7] Banzi, A., Folgieri, R. (2012). *Preliminary ressults on priming-based tools to enhance learning in museums of fine arts.* In Prooceding of Electronic Imaging & the Visual Arts : EVA 2012 Florence, 142-147, Firenze University Press.

Pages 165 to 168 are removed from the proceedings due to author’s nonpayment of the registration fee.

Proceedings
International Workshop on
Visual Languages & Computing
(VLC 2012)

DEViL3D – A Generator Framework for Three-Dimensional Visual Languages*

Jan Wolter
University of Paderborn
Department of Computer Science
Fürstenallee 11, 33102 Paderborn, Germany
jan.wolter@uni-paderborn.de

***Abstract*—Visual languages are beneficial particularly for domain-specific applications, since they support graphical metaphors of the domain. The development of graphical editors for such languages can be simplified by using generator frameworks. Up to now the majority of visual languages are two-dimensional, but there are domains which are much better described by three-dimensional language constructs. The use of three-dimensional representations is well known in the area of (scientific) visualizations, games, or movies. Our approach is to use 3D graphics for visual languages that make use of 3D relationships, and to develop a generator framework to simplify the implementation of 3D languages. Our system DEViL3D accomplishes this task and encapsulates special knowledge necessary to implement 3D editors for such languages. The language designer does not need to know about implementation of 3D graphics and interaction with them, because our framework supports this for each language implementation automatically. This paper introduces previous approaches in the area of 3D languages, describes the specification process to get a 3D language implementation using our generator framework, and especially illustrates the 3D specific features of these implementations automatically derived without further effort.**

***Index Terms*—automated generation, visual languages, visual programming, three-dimensional representations.**

I. Introduction

Visual languages play an important role in computer science today and are omnipresent in tools for modeling software and specific application domains. Examples of successful visual languages are LabVIEW [1]—used in industrial automation and instrument control—and the well-known UML, which is used to model object-oriented software systems. Both languages use two-dimensional representations, e.g., boxes and lines between these, in order to visualize dataflow or dependencies. But in some cases, the 2D representation of UML diagrams is not efficient enough and can be extended to 3D, e.g., to overcome the problem of intersecting edges in sequence diagrams, or to use the third dimension to focus on special classes of interest in class diagrams [2].

The first idea of three-dimensional visual languages goes back to a publication of Glinert [3] in 1987. He reveals that visual languages, so far, make use of at most 2.5-dimensional representations, which means the use of multiple, interconnected planar diagrams. Glinert shows how 2.5D representations can be extended to 3D and proposes projects in the area of visual languages that use the third dimension and therefore differ from classical visual language approaches.

A good example for a three-dimensional visual language is the description of molecular models. These models consist of atoms that are visualized as balls, and bonds between them that are represented by sticks. The arrangement of the atoms in the 3D space is the result of the electron cloud repulsion. Hence, the molecular language can be regarded as a representative of the class of application-specific languages that describe inherent three-dimensional graphical models.

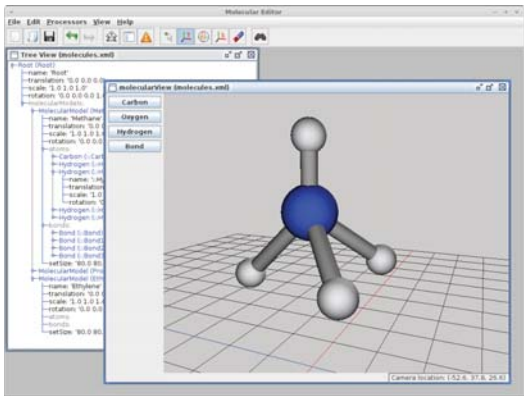


Fig. 1. Editor for molecular models generated with DEViL3D.

With our generator framework DEViL3D (*D*evelopment *E*nvironment for *Vi*sual *3D L*anguages), we have already generated an editor for molecular models (see Fig. 1). It allows to insert different atoms (we provide a small snippet of the periodic table, which knows 118 elements) and connect them by bonds. The editor offers techniques to manipulate language constructs and navigate inside the 3D scene.

As mentioned before, many visual languages were successfully developed for domain-specific applications. Such languages are often specifically tailored to the graphical notations of that scope. The development of a language specific implementation is justified, only if the effort is appropriately small. Therefore, effective generator systems are useful. We have experience in developing the generator framework DEViL [4] that generates structure editors for two-dimensional visual

*This work is supported by the German Research Foundation (Deutsche Forschungsgemeinschaft – DFG); contract no. KA 537/6-1.

languages from high level specifications. The system is characterized by methods, which effectively simplify the creation of specifications. These include in particular so called *visual patterns*, layout mechanisms, and an editor to specify graphical primitives. The DEViL system has been used successfully to create visual languages in cooperation with nameable companies like Volkswagen, Bosch [5], and SagemOrga [6].

In comparison to the development of 2D languages, new methods and techniques are needed, to develop editors for 3D languages. Knowledge about implementation of 3D graphics and interaction with them is required. Hence, a well usable generator system for 3D language implementations is needed, which encapsulates these methods and make them available for each generated editor, without further effort for the language designer. Some basic methods to develop visual languages can be inherited from the DEViL system.

Compared to editors for 2D languages, there are the following tasks that are new for 3D language editors: The navigation inside the 3D canvas, the insertion of new constructs into the 3D world, and the interaction with them. Furthermore, the layout of three-dimensional programs needs new methods. This paper focuses on these techniques, which are encapsulated in DEViL3D (see Section II). The general specification process to generate 3D language editors is described in Section III. In the section about related work, we give background information of some already existing three-dimensional languages and present a classification of 3D visualizations that are also applicable to 3D languages. Since the development of DEViL3D is work in progress, we give an outlook on future work in Section V.

II. Fundamental techniques for 3D languages

For the acceptance of three-dimensional languages, it is essential that there are editors that make the construction of 3D programs as easy as possible. This challenge will become clear, if we compare editors that allow the creation of 2D and 3D representations. They can be distinguished by the *degree of freedom* (DOF) that describes the possibilities of object placement in space. The 2D space has three DOFs: translation along the two axes and rotation around the neutral point. By contrast, the DOFs in three-dimensional space are twice as many: translation and rotation along all three axes. In order to cope with this increase in complexity, we need new techniques for navigation, interaction, and layout.

A. Navigation

Navigation is the planning and execution of travel trough the three-dimensional space with the aim to *explore* a scene without an explicit target, or to *search* a particular target. Each 3D canvas in editors generated by DEViL3D comprises a first-person-view camera to support such tasks.

To enable navigation tasks (and in general, all interaction tasks), the user needs a device to "communicate" with the editor. Input devices are just physical tools, which are also governed by the degree of freedom they have. The editors that are generated by DEViL3D support both, classical 2D devices and special 3D devices that cover all DOFs. Our aim is that the editors are available for users that do not own special 3D devices. Hence, it is possible to navigate in the scene, by using a classical 2D mouse and a keyboard: moving the mouse rotates the camera and pressing special keys let move the camera forwards, backwards, upwards, or downwards. It is also possible to use 3D input devices. A 3D mouse, which captures all six DOFs, can be used without any further device to navigate inside the scene. A prototypical support for the Microsoft Kinect motion detection device is also included. The Kinect device recognizes the users hand gestures and the camera moves according to this motion.

With this navigation support, the user is able to inspect the three-dimensional program from different perspectives. If some language objects, the user is interested in, are hidden by others, the camera can be used to navigate to these objects and avoid this occultation problem.

B. Interaction

One of the most important features of a structure editor is its ability to prevent the user from constructing syntactically incorrect programs. To do so, our editors provide special techniques that follows the well known *direct manipulation* paradigm by Shneiderman [7]. Interaction tasks can be distinguished into the *insertion* and the *manipulation* of objects.

The insertion of language constructs is triggered by so called *insertion contexts* that highlight positions, where new constructs can be inserted. Each 3D view provides besides the 3D canvas in the middle, buttons on the left-hand side. If the user clicks on such a button, all appropriate three-dimensional insertion contexts appear inside the 3D canvas. The nature of the insertion context is determined by the way, how the language object is organized and layouted in the scene. For example, objects can be part of a three-dimensional set, or a list.

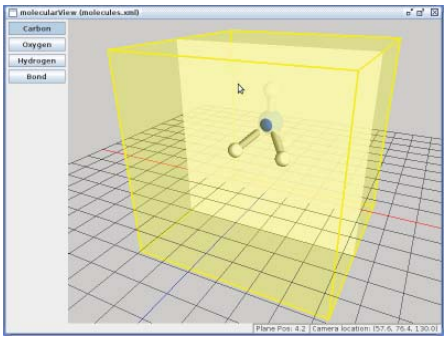
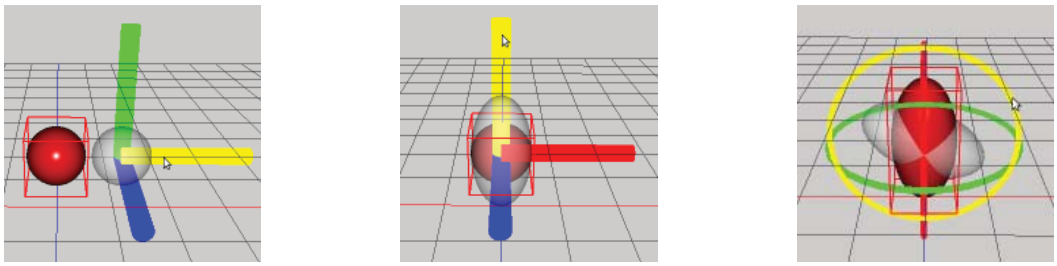


Fig. 2. Three-dimensional insertion context.

Let us consider the task to insert an object into a 3D set, e.g., to insert an atom to construct a molecule. The insertion context appropriate for this task is shown in Fig. 2. With this, the user is able to define a position, where the new language construct should appear. Similar to the navigation task, we want to make it possible for users to construct the program with a 2D mouse. Hence, the cubic insertion context includes a plane to ensure a three-dimensional position. This plane can



(a) Translate (b) Scale (c) Rotate

Fig. 3. Gizmos are used to manipulate language objects.

be moved along the z-axis, while the right mouse button is pressed. The 3D position of the object can be determined after a mouse click on the plane. This position is comprised of the x- and y-position, allocatable by the two-dimensional position of the mouse cursor, and the z-position of the plane.

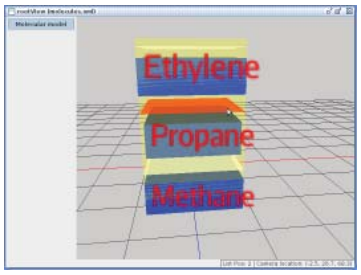


Fig. 4. Insertion context for list elements.

An insertion context for an object that is organized as an element of a list (along one of the three dimensions) can be seen in Fig. 4. The area, where the list is located, is visualized as a cube with planes inside that indicate the insertion position of the list. The actual selected plane is highlighted orange and after a click the new element appears exactly on this position.

After the user inserts an object into the graphical representation, an update (according to the *model-view-controller* paradigm) of the view is triggered that recomputes the graphical representation. This comprises, e.g., list elements which have to move up after inserting a new element, or special layout algorithms which ensure further requirements.

After the insertion of an object, the user is able to modify the object. To do so, the editor provides widgets [8] (called *gizmos*) that can be applied to objects and enable the user to translate, scale, or rotate an object. The gizmos to translate and scale are depicted by three cylinders and the gizmo to rotate objects is depicted by three tori (see Fig. 3). Each dimension can be manipulated separately by pressing the right mouse button and will be colored yellow. While moving the mouse, a so called *ghost object* will be visible. This object is used to give the user feedback of the temporary state, while the mouse button is still pressed. On releasing the mouse button, the object gets the new position, size, or orientation.

C. Layout

The way, how language constructs will be layouted, can influence the usability of the editor. For example, the question how much effort is required to insert new objects into the scene can be influenced by the layout. This question refers to the dimension *viscosity* of Green's and Petre's famous *cognitive dimensions framework* [9]. In our framework, the layout of the visual representation is encapsulated through the utilization of *visual patterns*. Examples for visual patterns are three-dimensional sets, lists, line connections, or graphs. Depending on the chosen pattern, the language constructs are layouted differently.

For objects that are organized within a three-dimensional set, a special algorithm ensures the mutual *non-interpenetration* of all objects (analog to *non-overlapping* in 2D). The algorithm searches for the leftmost interpenetration of two language constructs and identifies the x-, y-, or z-dimension where the overlapping is minimal and moves the language construct along this dimension to fix the interpenetration. This will be done until all interpenetrations are resolved. The concrete position for each element of a list is also computed by a mechanism encapsulated in the list pattern. The layout of nested representations will be computed by a hierarchical procedure: First, the size requirements from inside to outside are computed. In a second step, the final positions are propagated from outside to inside.

We are working on a pattern that is used to layout three-dimensional graphs. To do so, we are using an algorithm that arranges the graph in an aesthetical way. Eades shows [10] that this can be achieved with *spring embedding algorithms*, which layout the graph according to two important criteria: all the edge lengths ought to be about the same and the layout should display as much symmetry as possible.

III. The DeviL3D framework

This section gives a brief overview on the concrete steps necessary to generate a structure editor for a 3D language with DEViL3D. A language specification addresses three different aspects: the *abstract structure*, the *visual representation*, and *code generation* (see Fig. 5). From these parts, DEViL3D generates a structure editor, which includes techniques to construct and manipulate three-dimensional programs (as described in the previous section). This approach is inherited

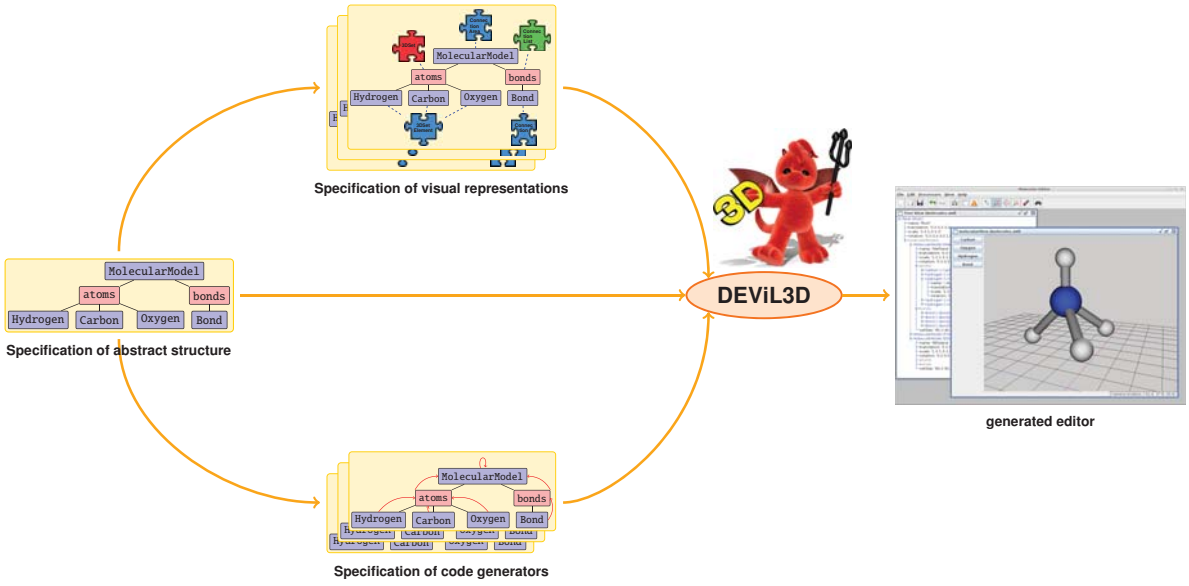


Fig. 5. Specification aspects in DEViL3D.

from the predecessor system DEViL and is provably beneficial to develop structure editors efficiently. The concepts that are adapted from DEViL are described only briefly here. The reader is referred to publications on that system: [4], [11].

The abstract structure describes the language constructs and the way how they are connected, without defining a concrete representation. This is done with a specifically tailored textual domain specific language, which is strongly related to object oriented programming languages. It is based on well known concepts like classes, inheritance, attributes, and references.

The concrete representation of the visual language will be achieved by the specification of visual representations. To do so, the language designer defines a set of views that are based on a part of the abstract structure. A view consists of several buttons for language constructs and a three-dimensional canvas, which displays the graphical objects. To specify the visual representation, it is mostly sufficient to select visual patterns from a library to assign them to constructs of the abstract structure. These visual patterns can be parameterized in a declarative way. They automatically contribute layout and interaction properties, as described in the previous section. For example, all atomic constructs of the abstract structure are associated to the visual pattern *3D set*, which supports the interpenetration-free arrangement in the 3D space.

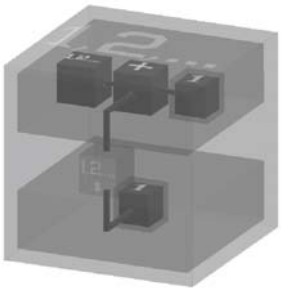
The concrete visual appearance of a language construct is realized by a visual pattern named *VP3DForm*, which refers to a textual specification of a so called *generic drawing*. Such a generic drawing basically consists of several *graphical primitives*, special areas (called *containers*) which determine placeholders, where sub elements can be nested, and methods to transform the drawing, if the required size of the sub elements exceeds the size of the container. The language designer can choose from a library of predefined primitives like spheres, boxes, cylinders, cones or tori, which are provided by the underlying *jMonkeyEngine* [12]. We use this 3D engine—which were originally developed to program 3D games—for all low-level drawing tasks. More complex primitives can be designed with modeling tools like Blender [13].

Analog to the visual representation, the language designer can define a set of code generators, which transform the 3D program into a textual representation. To achieve this, we use well known concepts from the area of compiler construction that are encapsulated in the *Eli system* [14].

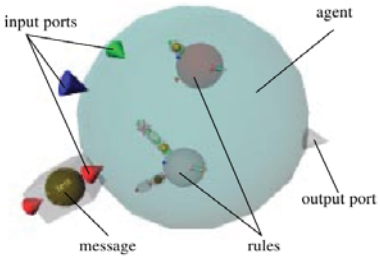
A structure editor, generated from these specifications, provides a multi-document interface, which shows, besides the tree view of the language model, the three-dimensional views, specified by the language designer. Another important issue is the possibility to define a coupling between different views. For instance, it is quite conceivable that a molecular editor has, besides a view that shows a ball-and-stick model, a view that shows the molecules as a *space filling model* (also known as *calotte model*). With the aid of synchronization functions, such coupled views can be kept consistent. As mentioned in the previous section, the purpose of the structure editor is to keep programs syntactically correct. In general this is achieved by the program composition with the special tailored insertion contexts that allow the insertion of objects only at valid positions. But for further language dependent checks, the language designer can implement special functions, which are called after each modification and report potential violations in an error view.

IV. Related Work

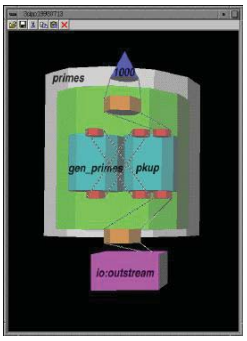
Two-dimensional visual languages have been established since circa three decades. Thus, the idea to automate the implementation of such languages was very obvious. Hence, many generator frameworks, based on different design principles, were established. Besides the DEViL system—which



(a) Cube: natural number generator (from [15]).



(b) SAM: representation of an agent (from [16]).



(c) 3D-PP: prime number computation (from [17]).

Fig. 6. Some historical three-dimensional languages.

was recently extended by a simulation concept [18]—there are many other generator frameworks for visual languages, e.g., *DiaGen* [19], *DiaMeta* [20], *VLDesk* [21], or *MetaEdit+* [22]. But to the best of our knowledge, DEViL3D is the first approach of a generator system for 3D languages. Hence, this section about related work focuses on historical three-dimensional visual languages and on a classification for such languages.

The most prominent three-dimensional language is *Cube* developed by Najork [15]. Cube is a three-dimensional, visual, logic programming language, which is semantically similar to Prolog and makes use of the data-flow paradigm. Each language construct, like predicates, data types, or values is represented by a cube, which may contain further cubes or may be connected by a pipe with others. The interior arrangement of a predicate definition cube consists of transparent boxes called planes, which represent semantic information—the vertical arrangement indicates disjunction, while the horizontal arrangement indicates conjunction. This visualizes an analogy to Prolog, since each plane corresponds to a clause of a Prolog program. The example of a Cube program in Fig. 6(a) shows a predicate definition cube that defines a natural number generator, which bases on the following recursive definition: a) 1 is a natural number, and b) if n is a natural number, then so is $n + 1$. Inside the cube, there are two planes which are stacked vertically to visualize the disjunction. The lower plane forms the base case and the upper plane forms the recursive case of the definition.

The *SAM* (Solid Agents in Motion) language [16], which is parallel, synchronous and state-oriented bases on the well known 2D language *Pictorical Janus*. A SAM program describes a set of agents that synchronously communicate by exchanging messages. SAM programs consist of an abstract representation for specifying the agent's behavior and a concrete representation showing the application domain. The abstract as well as the concrete representation is three-dimensional. 3D objects in SAM are 3D messages, agents with ports, and rules with a precondition and a sequence of actions (see Fig. 6(b)).

The 3D-PP language [17] bases on the concurrent logic programming language GHC and offers the possibility to construct three-dimensional programs—by combining pictorial program elements—that are equivalent to GHC programs. Program elements, e.g., atomic values, lists, or I/O data have a three-dimensional representation. The designer of 3D-PP developed an editor, which supports direct manipulation techniques to construct and manipulate the 3D program representation. Program elements can be represented semi-transparent to make objects visible that are nested inside another object, or that are hidden by an object in the foreground. Elements that are nested into each other can be seen in the example in Fig. 6(c). This example shows a program that computes the first 1000 prime numbers.

Stasko and Wehrli classify 3D visualizations [23] in three categories: *augmented 2D*, *adapted 2D*, and *inherent 3D* visualizations. Their categories can be applied to 3D languages as well. Augmented 2D representations require less than three spatial dimensions. The third dimension is added merely for aesthetic purposes. Such representations take no advantages from the third dimension and hence, they are less suitable. The adapted 2D representations require at least two spatial dimensions; however the third dimension is used to encode extra features. For example, Cube programs use the third dimension to encode conjunction or disjunction through arrangement of planes. The meaning of inherent 3D representations should be obvious and includes representations with inherently three-dimensional entities, such as all real-life objects. We extend this category to include representations where the arrangement of constructs to one another is inherently 3D. One example for this are the molecular models, because the arrangement of atoms—resulting from the electron cloud repulsion—is inherently three-dimensional.

V. Future work

Our generator DEViL3D is able to generate full usable three-dimensional editors, e.g., for molecular models. But we have further ideas that we would like to assemble. Our goal is to complete the three-dimensional aspects of the generated editors. Again, these aspects can be subdivided in techniques to navigate and interact with language constructs, together

with new layout mechanisms. All these techniques have to be developed in a generic manner and have to be integrated in DEViL3D, such that they can be used in any editor. In the following, we report on investigations related to these aspects.

At present, the navigation inside the 3D world relies on the unrestricted movement of the first-person camera. Another approach is the constraint-based navigation, that is, to focus special parts of interest automatically, or zoom in.

The layout of three-dimensional programs is mainly determined by the application of visual patterns. To increase layout alternatives, we have ideas for more visual patterns. For example, so called *cone trees* [24] are an effective method from the area of information visualization to visualize hierarchical information structures. Furthermore, it is conceivable to arrange objects as part of a three-dimensional cube which stacks several planes, which in turn are internally organized as matrices.

Up to now, we have implemented all layout concepts by hand. In general, we have to explore if the use of a constraint solver is beneficial. In domains, constituted of real-world objects, objects are constrained to move on a horizontal floor or have to dock on other objects. For such purposes, a constraint solver can automatically satisfy these requirements by computing a solution for a set of equations.

Until now, the generic drawings, mentioned in Section III, are specified textual. It would be helpful to have a three-dimensional editor, which allows the specification of generic drawings in a WYSIWYG-style. We plan to implement such an editor using DEViL3D and make use of the code generation module that have to output the specification code for visual specified drawings.

VI. Conclusion

In this paper we have presented an approach to automatize the implementation of structure editors for three-dimensional languages with the DEViL3D generator framework. The focus of this paper lies on techniques that make the visual composition of three-dimensional programs possible. First of all, navigation techniques allow the inspection of the whole 3D scene with different input devices and can avoid the occultation of objects. To insert language constructs into the 3D world, special insertion contexts are developed and gizmos allow their manipulation. The layout of the three-dimensional programs is defined by visual patterns that provide appropriate layout strategies, e.g., the non-interpenetration of objects that are organized in a 3D set. All these techniques are provided by the generator system and are available in each generated editor. Techniques that are not specific for 3D are adopted from the DEViL system.

DEViL3D is indented to effectively support the development of new three-dimensional visual domain specific languages, and thus encourage the use of the third dimension in modeling domains.

References

[1] "National Instruments LabVIEW," http://www.ni.com/labview. [Online; accessed 23-April-2012].

[2] M. Gogolla, O. Radfelder, and M. Richters, "Towards Three-Dimensional Representation and Animation of UML Diagrams," in *Proceedings of the International Conference on The Unified Modeling Language: Beyond the Standard*, Sep. 1999, pp. 489–502.

[3] E. P. Glinert, "Out of Flatland: Towards 3-D Visual Programming," in *Proceedings of the Fall Joint Computer Conference on Exploring technology: today and tomorrow*, 1987, pp. 292–299.

[4] C. Schmidt, U. Kastens, and B. Cramer, "Using DEViL for Implementation of Domain-Specific Visual Languages," in *Proceedings of the Workshop on Domain-Specific Program Development*, Jul. 2006.

[5] B. Cramer, D. Klassen, and U. Kastens, "Entwicklung und Evaluierung einer Domänenspezifischen Sprache für SPS-Schrittketten," in *Proceedings of the Workshop on Domain-Specific Modeling Languages*, Mar. 2008, pp. 59–73.

[6] C. Schmidt, P. Pfahler, U. Kastens, and C. Fischer, "SIMtelligence Designer/J: A Visual Language to Specify SIM Toolkit Applications," in *Proceedings of Second Workshop on Domain Specific Visual Languages in conjunction with OOPSLA*, Nov. 2002.

[7] B. Shneiderman, "Direct Manipulation: A Step Beyond Programming Languages," *IEEE Computer*, vol. 16, no. 8, pp. 57–69, Aug. 1983.

[8] D. B. Conner, S. S. Snibbe, K. P. Herndon, D. C. Robbins, R. C. Zeleznik, and A. van Dam, "Three-Dimensional Widgets," in *Proceedings of the Symposium on Interactive 3D Graphics*, Mar. 1992, pp. 183–188.

[9] T. R. G. Green and M. Petre, "Usability Analysis of Visual Programming Environments: A 'Cognitive Dimensions' Framework," *Journal of Visual Languages and Computing*, vol. 7, no. 2, pp. 131–174, 1996.

[10] P. Eades, "A Heuristic for Graph Drawing," *Congressus Nemerantium*, vol. 42, no. 11, pp. 149–160, 1984.

[11] C. Schmidt and U. Kastens, "Implementation of visual languages using pattern-based specifications," *Software – Practice and Experience*, vol. 33, no. 15, pp. 1471–1505, 2003.

[12] "jMonkeyEngine 3 – Documentation," http://jmonkeyengine.org/wiki/doku.php/jme3. [Online; accessed 3-April-2012].

[13] "Blender," http://www.blender.org/. [Online; accessed 23-April-2012].

[14] U. Kastens, P. Pfahler, and M. Jung, "The Eli System," in *Compiler Construction*, ser. Lecture Notes in Computer Science. Springer Berlin / Heidelberg, 1998, vol. 1383, pp. 294–297.

[15] M. A. Najork, "Programming in Three Dimensions," *Journal of Visual Languages and Computing*, vol. 7, no. 2, pp. 219–242, 1996.

[16] C. Geiger, W. Müller, and W. Rosenbach, "SAM – An Animated 3D Programming Language," in *Proceedings of the IEEE Symposium on Visual Languages*, Sep. 1998, pp. 228–235.

[17] T. Oshiba and J. Tanaka, ""3D-PP": Visual Programming System with Three-Dimensional Representation," in *Proceedings of the International Symposium on Future Software Technology*, Oct. 1999, pp. 61–66.

[18] B. Cramer and U. Kastens, "Animation automatically generated from simulation specifications," in *Proceedings of the IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, Sep. 2009, pp. 157–164.

[19] M. Minas, "Concepts and realization of a diagram editor generator based on hypergraph transformation," *Science of Computer Programming*, vol. 44, no. 2, pp. 157–180, 2002.

[20] ——, "Generating Meta-Model-Based Freehand Editors," in *Proceedings of the 3rd International Workshop on Graph Based Tools (GraBaTs'06), Natal (Brazil), Satellite event of the 3rd International Conference on Graph Transformation*, ser. Electronic Communications of the EASST, A. Zündorf and D. Varró, Eds., vol. 1, Sep. 2006.

[21] G. Costagliola, V. Deufemia, and G. Polese, "A Framework for Modeling and Implementing Visual Notations With Applications to Software Engineering," *ACM Transactions on Software Engineering and Methodology*, vol. 13, no. 4, pp. 431–487, Oct. 2004.

[22] *MetaEdit+ Domain-Specific Modeling (DSM) environment*, MetaCase Consulting, 2012, http://www.metacase.com/products.html. [Online; accessed 4-April-2012].

[23] J. T. Stasko and J. F. Wehrli, "Three-Dimensional Computation Visualization," in *Proceedings of the IEEE Workshop on Visual Languages*, Aug. 1993, pp. 100–107.

[24] G. G. Robertson, J. D. Mackinlay, and S. K. Card, "Cone Trees: Animated 3D Visualizations of Hierarchical Information," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Reaching Through Technology*, Apr. 1991, pp. 189–194.

Semi-Automated Generation of Domain-Specific, Natural Language-Based, Visual Programming Languages

Michael Wisely, Jennifer L. Leopold
Department of Computer Science
Missouri University of Science and Technology
Rolla, MO, USA
mwwcp2@mst.edu, leopoldj@mst.edu

Abstract—With the continued proliferation of digitally available data, it is likely that in the future more programming will need to be done by professionals who are untrained as computer scientists. To this end, domain-specific programming languages like MATLAB, Excel VBA, and LabVIEW G will continue to increase in popularity. Herein we briefly examine three domain-specific languages, considering positive and negative attributes of each. We then propose the creation of a tool that will facilitate semi-automated generation of new domain-specific programming languages. To best address the needs and technical background of the intended user population, the generated language will be graphical, with a syntax and semantics based on natural language.

1. INTRODUCTION

With the advent of domain-specific programming languages such as MATLAB and Excel VBA, engineers, accountants, and other non-programmers are learning to construct programs to solve problems specific to their fields. Though these languages are useful to those without degrees in computer science, to new users they are no easier to understand than programs written in traditional, general-purpose programming languages like C++ or Python. One still must learn to read and write programs that are heavily dependent on a precise syntax, expressed using non-intuitive combinations of letters and symbols. Despite the conciseness of such a textual representation, programmers understand and describe the semantics of their code by effectively translating it to natural language.

Some domain-specific languages, like LabVIEW's G language, take a different approach to programming, one that is not as focused on the use of text to express the computation. Using a graphical interface, users can drag, drop, and connect elements to specify the actions that are to be performed. The use of a visual interface in languages such as G simplifies the programming experience for end-users.

This paper briefly discusses advantages and disadvantages of some popular domain-specific programming languages. It then proposes a natural language-based tool to semi-automate the development and implementation of new visual, domain-specific programming languages. Intuitively, utilizing a representation that a user already understands (i.e., natural language) should allow him/her to more effectively develop algorithms to solve problems.

2. DOMAIN-SPECIFIC LANGUAGES

2.1. Importance of Domain-Specific Languages

Programming is a skill that can be used to efficiently solve many different types of problems. Computer scientists commonly use languages like Python and C++ because they are robust, augmented by an extensive set of libraries to accomplish any computing task. Experienced programmers are familiar with the imperative programming paradigm, and after an hour or so of reading documentation for a new language, they likely can begin working on a program to solve a problem. Those who are not well versed in languages such as C++ or Python have a much larger hurdle to overcome: learning to communicate to the computer a sequence of actions that must be performed to solve a problem.

Many programming languages are difficult or intimidating to learn due to their size [1]. These languages were designed so that they could be applied to a variety of problems in different domains. However, in many respects it is the flexibility of these languages that reduces their expressiveness. To this end, domain-specific programming languages were developed. Typically, these languages have a small set of features (as compared to general-purpose programming languages), but their syntaxes are much more expressive.

End-user programmers easily can become frustrated if they must write code to manage computer resources (e.g., memory, threads, etc.), or to define data structures. As such, domain-specific languages are typically high-level programming languages; issues like memory management, concurrency, and data structure implementation are handled by the language internals, and do not require explicit specification by the user [3].

2.2. Review of Existing Domain-Specific Languages

2.2.1. MATLAB: Imperative Textual Programming

Although domain-specific languages are not necessarily designed to be used by non-professional programmers, they

often are. For example, MathWorks' MATLAB programming language was designed so that users could easily solve problems that require matrices, vectors, and functions related to linear algebra [4]. The semantics of MATLAB follow the rules of linear algebra: the dimensions of multiplied matrices must agree, rows and columns are indexed by one, etc. [5]. Programs written in MATLAB leverage powerful linear algebra primitives, thereby facilitating the development of numerical computing applications.

However, it easily could be argued that MATLAB's syntax is no friendlier than a general-purpose programming language. In MATLAB, users can assign values to variables, perform arithmetic operations, and call functions using an imperative, C-like syntax. For experienced programmers, the syntax is familiar and reasonably intuitive. But for others, this syntax can be confusing.

2.2.2. LabVIEW: Declarative Graphical Programming

The G programming language, designed by National Instruments for LabVIEW, takes a much different approach to writing programs. Instead of writing text to specify the actions to be performed, users construct block diagrams. This is done by connecting input and output blocks to function blocks using a graphical user interface. LabVIEW then compiles the diagrams to machine code, and executes them [6]. G is often used in the development of measurement and control systems. Its user interface is friendly to those who are not experienced programmers; users easily can connect different inputs and functions together to produce output. As such, block diagrams provide a convenient way to write declarative programs [10]. However, not all problems can best be thought of in terms of the flow of data; hence such languages actually might be an obstacle to articulating a solution.

2.2.3. Alice: Imperative Graphical Programming

To help beginning programmers learn to write programs, researchers at Carnegie Mellon University developed Alice: an imperative visual programming language. Using Alice, users can solve computational problems by creating 3D scenes, and specifying various actions to manipulate the objects in the scenes [7]. Like MATLAB, Alice lets users define the step-by-step execution of their programs, and like LabVIEW, it has a graphical user interface, allowing users to drag-and-drop statements to create programs.

Unlike MATLAB and LabVIEW, each statement in Alice is represented in pseudo-English. Consequently, reading an Alice program is much like reading broken English. Similarly, writing Alice programs is as simple as dragging desired statements into the editor and filling in the blanks. As the user drops a statement into the editor, the editor prompts the user to fill in every necessary parameter. For example, in Figure 1 the user has dragged the "bunny turn" method into the editor. S/he then is prompted for a direction and amount for the turn before the statement is actually inserted into the program. While runtime errors still can occur, syntactic and semantic errors cannot.

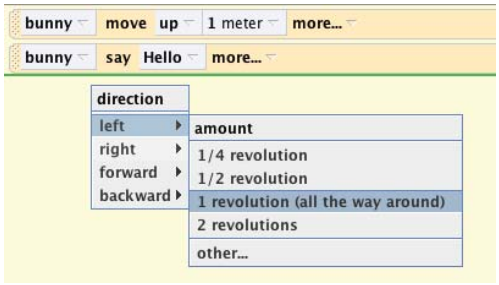


Figure 1: Alice sample code with editor parameter menu

III. MAKING PROGRAMMING MORE ACCESSIBLE

3.1. The Approach Used to Generate a Domain-Specific, Natural Language-Based, Visual Programming Language

By combining natural language and graphical programming, Alice helps its users to quickly learn to read and write programs. To this end, we propose the creation of a tool that is capable of *generating* a visual programming language using a natural language description of the problem domain. In brief, the tool will work as follows:

1. The language developer provides the tool with a description of what entities exist in the problem domain and what actions those entities can perform.
2. The tool parses the statements in the description, extracting information; for example:
   a. A knowledge base is searched for word association (e.g. "move" might imply "location").
   b. A dialog with the user is initiated when ambiguities in the description need to be clarified.
3. The tool generates a Python code framework that allows the language developer to write code to specify any additional details of the implementation for the new language. For example, if a user is designing a language to control a robot, the user will need to fill in the code framework to provide the logic necessary for communicating with the hardware.
4. The tool wraps the language with a drag-and-drop interface, making it visual.
5. The language developer can integrate the new language into other software projects so that the end-users effectively can program visually in a pseudo-natural language.

3.2. The Use of Natural Language in Programming

Since the mid-1960's, programming using natural language has been a hot topic of research. Intuitively, if users can articulate a solution using their natural language, the learning curve for programming should be greatly reduced, as there is no longer a need to learn a new language in which to "communicate." However, interpreting natural language can be incredibly challenging [8]. Natural language can be ambiguous, even to native speakers. For example, the phrase, "capital of the country" may refer to capital city of a particular country or to its wealth. At times, identifying the

parts of speech for a phrase can be difficult to determine depending on the context in which it is used.

Despite these challenges, researchers at MIT have developed a tool named Metafor that is capable of translating natural language descriptions of problems into code [9]. Users provide a description of a program to Metafor written in English, which subsequently is parsed to create code for a program that will accomplish the necessary tasks. This helps users to quickly convert their ideas from natural language into code. However, Metafor does not generate all of the logic for a program. It only creates the "scaffolding" for a program. Users must fill in the rest of the code to have a properly functioning program. Although it does not generate programs in their entirety, Metafor would make an excellent tool for software engineering purposes.

Extracting knowledge from natural language descriptions is the most challenging aspect of implementing our language generator tool. Metafor generates skeleton code from a natural language description, but our tool must do more. Since our goal is to generate a new programming language based on a natural language description of the objects and actions relevant to a particular problem domain, we also must generate a grammar for the new language. Because Metafor is designed to generate programs based on natural language stories told by users, it is capable of interpreting these stories as program descriptions and mapping them to statements in Python. As such, it is not necessary for Metafor to record *how* it mapped natural language statements to code. Once the scaffolding is generated, the mapping is no longer needed. Our tool needs to take one step beyond the accomplishments of Metafor, generating parser definition files in addition to skeleton code. These definition files will define a BNF grammar for the pseudo-English language, and make calls to the generated skeleton code.

The language processing techniques used by Metaphor could be leveraged to develop a grammar for a new programming language. During the disambiguation process, verbs and nouns identified by Metafor are mapped to functions and objects, respectively. By maintaining the user's natural language description of actions and objects, a collection of pseudo-natural language commands could be developed and formed into a formal grammar. Though this new language would be a usable domain-specific language, it would still require that programmers have a firm understanding of the language's syntax.

3.3. Visual Programming

As previously discussed for LabVIEW and Alice, visual programming has several advantages over textual programming. Foremost is the reduction of syntax and semantic errors in programs. Visual languages are much more "interactive" during the actual development of a program than are compiled languages and even REPL interpreters. Instead of letting users make syntactic errors and then displaying error messages, users are prevented from making errors in the first place [1]. Alice forces the programmer to choose parameters for each statement before inserting the statement into its program editor. By doing this, users cannot construct incomplete statements. Also as previously described, Alice forces users to select values of the correct type as parameters to statements, thereby preventing potential type errors.

Our tool certainly could be used to generate an Alice-like, stand-alone programming environment, but we envision its potential use to be much more extensive. Not only could new applications be more quickly developed, but the tool could be used to add visual programming features to existing products. By extending an existing command line programming interface with a visual programming interface, software developers can expand their target audience to include programmers and non-programmers alike.

IV. APPLICATION TO PYVRCC-3D

4.1. Objectives of the Application

VRCC-3D+ [10] is the implementation of a mathematical model that supports quantitative and qualitative spatial reasoning in 3D. To support the use of this tool as an automated spatial reasoner, an end-user programming language could be designed that would allow an untrained programmer to create visual "programs" representing a collection of 3D objects and the (generalized) spatial relations that hold between each pair of objects at different abstract time periods (i.e., "states").

Developing such a programming language could be automated by using our language generator tool. What follows are some of the steps that would be taken to create a visual language for the Python implementation of VRCC-3D+, which we shall call PyVRCC3D. For simplicity, this example assumes that objects in the 3D world only can be solid spheres.

A. The Domain Description

Objects exist in a 3-D Euclidian world. They are spherical and solid. They can obscure other objects from view. They can move through the world, but not through each other.

B. User Interaction with the Language Generation Tool

At this stage, the tool begins a dialog with the user (language developer) to clarify any ambiguities in the initial description. This stage also allows the user to provide extra details to the tool, which s/he may not have considered when writing the initial domain description. To clarify a potential ambiguity in the description, the tool may ask the user, "*You specified that 'Objects are spherical and solid.' Are all objects spherical and solid, or are these traits of only some objects?"* The user would then have an opportunity to confirm this trait or provide more details for the tool to process.

C. Generated Code Skeleton

At this point, the tool would have generated a Python code framework according to the discussion with the user. It should be noted that the language designer would have to add code to the generated functions to leverage existing VRCC-3D+ code that determines obscuration. However,

once the tool knows that the new language is to be associated with the VRCC-3D+ API, it may suggest available VRCC-3D+ functions that contain keywords such as "obscure" in the name.

D. Generated Language Grammar

The tool then generates skeleton code based on the information gathered during its dialog with the developer. Both the conversation and code generation capabilities are similar to the functions of Metafor. However, the unique feature of our tool is its ability to generate a new language as well.

Based on the interaction with the user, the tool creates a list of different object data types. In addition, it can reason when certain primitive types should be used. For example, floating points may be chosen to represent 3D coordinates. This could be determined either from user interaction or from data acquired from the knowledge base.

E. Visual Interface for the Generated Language

Using the code framework and associated domain specific language, an Alice-like visual interface can be generated. This generated interface allows users to write programs in the new language without concern for syntax. As previously mentioned, a visual interface guides users when writing programs, circumventing potential syntax errors. Dragging and dropping objects into function calls gives programmers and non-programmers alike a friendly interface to express algorithms in the new programming language.

IV. FUTURE WORK

For many problem domains, users may want to add concurrency to their programs. PyVRCC-3D users, for example, may want to move a group of objects at the same time. In the context of PyVRCC-3D, a "simultaneously" block would make sense.

Another priority for future work is providing the ability to create a container construct; for example, C and C++ have arrays, Python has lists, and MATLAB has vectors. To allow users to manipulate groups of objects, the language generation tool should allow designers to define abstract data types such as lists, and allow the programmers to apply actions to entire collections rather than just atomic entities.

Certainly a high priority is to conduct user testing to evaluate the potential usability and usefulness of both the programming language generator tool, as well as the generated languages. We intend to conduct such studies using untrained programmers from several different fields, including biology and engineering.

VI. CONCLUSION

The advantages gained by integrating natural language statements into visual programming can be obtained in problem domains beyond 3D movie making and education. Domain-specific languages, by nature, have a minimal set of commands. Consequently, this reduces the number of corresponding natural language statements, making the language more reasonable to implement graphically than a general-purpose language.

To aid in the creation of Alice-like languages, we have proposed the development of a language generation tool. This tool will generate a visual programming language that is based on a set of natural language statements provided by the language designer. The resulting visual programming language then can be used to develop applications from scratch for a particular problem domain, or be used to add functionality to existing software applications. To that end, it will facilitate programming for both the beginner and the experienced programmer.

VIII. REFERENCES

[1] Brusilovsky, P., Calabrese, E., Hvorecky, J., Kouchnirenko, A., & Miller, P. (1997). Mini-languages: a way to learn programming principles. *Education and Information Technologies*, 65-83.

[2] Spinellis, D. (1999). Notable design patterns for domain-specific languages. *The Journal of Systems and Software*, *56*, 91-99.

[3] Mernik, M. (2005). When and how to develop domain-specific languages. *ACM Computing Surveys*, 316-344.

[4] The MathWorks, Inc. (2011). *Matlab datasheet*. Retrieved from http://www.mathworks.com/tagteam/70533_91199v01_MATLAB-DataSheet_v9.pdf?s_cid=ML2012_bb_datasheet

[5] The MathWorks, Inc. (2011). *Matlab documentation*. Retrieved from http://www.mathworks.com/help/techdoc/?s_cid=ML2012_bb_doc

[6] National Instruments Coorporation. (2012). *The Benefits of Programming Graphically in NI LabVIEW*. Retrieved from http://www.ni.com/labview/whatis/graphical-programming/

[7] Conway, M., Audia, S., Burnette, T., Cosgrove, Dennis., and Christiansen, K. (2000). Alice: lessons learned from building a 3D system for novices. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (CHI '00). ACM, New York, NY, USA, 486-493.

[8] Bruckman, A., Edwards, E. (1999). Should we leverage natural-language knowledge? An analysis of user errors in a natural-language-style programming language. In *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit* (CHI '99). ACM, New York, NY, USA, 207-214.

[9] Liu, H., Lieberman, H. (2005). Metafor: visualizing stories as code. In *Proceedings of the 10th international conference on Intelligent user interfaces* (IUI '05). ACM, New York, NY, USA, 305-307.

[10] Johnston, W. M., Hanna, J. R. P., & Millar, R. J. (2004). Advances in dataflow programming languages.*ACM Computing Surveys*, *36*(1), 1-34.

Formalizing the Syntax of Codecharts

Jon Nicholson Aidan Delaney Gem Stapleton
Visual Modelling Group, University of Brighton
{j.nicholson,a.j.delaney,g.e.stapleton}@brighton.ac.uk

***Abstract*—Codecharts are a formal diagrammatic language for specifying the structure of object-oriented programs, frameworks, and design patterns. One of the principle concerns of codecharts is effective tool support in both prescriptive (e.g. design verification) and descriptive (e.g. design visualization) tasks. To ensure unambiguous interpretation of the language by both end users and machines it is essential that the syntax and semantics are formalized. To this end, we formally define the syntax of codecharts using the now standard approach of defining both a concrete and abstract syntax. The concrete syntax allows us to reason over the labelled geometric shapes in a diagram in order to decide the validity of a codechart. The abstract syntax allows us to reason about content without irrelevant geometric properties, providing the foundation on which to build a semantic account. To ensure an appropriate link between the concrete and abstract syntax we define a structure preserving mapping between the two syntax levels.**

I. Introduction

Large software development projects are often marred by substantial problems, a syndrome generally referred to as the 'software crisis'. Fitting with [3] and [6], we attribute difficulties that software development projects face to four inherent properties of software. *Complexity*: Software systems are some of the most complex artifacts ever manufactured by humankind; *Invisibility*: Software is intangible and invisible, hence its design flaws are particularly difficult to detect and resolve; *Conformance*: Designing complex programs and enforcing conformance to design decisions are distinctly difficult problems where manual verification, if at all possible, demands intensive effort and is largely impractical; and *Changeability*: Software and its documentation must be continuously evolved to meet ever changing requirements.

Approaches to software engineering address these properties by providing a framework for reasoning about the software system under development, for example [1], [2], [7], [8], [9]. Each of these approaches, like codecharts (LePUS3) [4], are visual languages for specifying different aspects of software. We believe that formal visual languages with strong attention to reasoning stand to make the biggest impact because of their, typically, user friendly and accessible nature.

However, most of the aforementioned notations are not formalized or are too expressive to enable efficient (or even decidable) reasoning to be undertaken. The notations in [1], [8], [9] have no formal syntax and their semantics have been presented intuitively by example, which are major limitations. By contrast, [2] and [7] have both a formally defined syntax and semantics, which are key to unambiguous interpretation, understanding and tool support. However, the expressive richness of these notations means that efficient reasoning about software models developed using them is not possible.

By contrast, codecharts have been designed to be a lightweight (relatively inexpressive) notation for effective tool support in both prescriptive (e.g. design verification) and descriptive (e.g. design visualization) tasks. Due to their relatively inexpressive nature, reasoning about the link between model and software becomes not only possible but tractable. In particular it is, at least in principle, possible to automatically produce a codechart from an implementation (i.e. design visualization), verify it against that implementation (i.e. design verification) and check that it is semantically consistent with an existing software model. The Two-Tier Programming Toolkit is a proof of concept that allows a user to create, reverse engineer, edit, and verify codecharts against Java implementations. However, in order to fully realize the benefits of codecharts, it is necessary to formally define their syntax and semantics. This will enable the development of an inference system that can be used to reason about codecharts.

This paper formalizes the visual syntax of codecharts using the approach first set out in [5] and seen in [7], in which the concrete and abstract syntax for a diagrammatic notation are considered as distinct concepts. In §II we introduce codecharts through a series of examples. We formalize our understanding of codecharts by defining a concrete (§III) and abstract (§IV) syntax, and the formal mapping between them (§V).

II. Informal Overview of Codecharts

Codecharts visually specify/model the structure of object-oriented programs, application frameworks and design patterns. Its vocabulary represents common concepts in class-based object-oriented program design as found in languages such as Java and C#: classes, methods and method signatures. Particular attention is given to sets of classes related by inheritance (inheritance class hierarchies), and sets of dynamically bound methods. Also represented are properties of, and relationships between, (sets of) classes and methods.

Table I presents the visual tokens of the codechart language[1]: LePUS3. These visual tokens are the building blocks of codecharts and we must provide constraints over how they are allowed to interact. Through this section we give examples of diagrams that are codecharts (i.e. satisfy the constraints that we will provide) as well as diagrams that use these visual tokens but are not codecharts. We note that [4], where codecharts are introduced, does not specify such constraints but instead illustrates use of each visual token by example.

[1]In this paper we focus on 'constants', and leave 'variables' for future work.

TABLE I
PRIMITIVE VISUAL TOKENS USED IN CODECHARTS

Class	A **rectangle** represents a class, labelled by a class name. An **offset rectangle**, such as in Fig. 1b, represents a set of classes, labelled by a set of class names.
Signature	An **ellipse** represents a method signature, labelled by a signature name. An **offset ellipse**, such as in Fig. 2b, represents a set of method signatures, labelled by a set of signature names.
(root, a set of classes)	A **triangle** represents an inheritance class hierarchy, labelled by a class hierarchy name. An **offset triangle**, which does not appear in this paper, represents a set of inheritance class hierarchies, labelled by a set of class name hierarchies.
R	An **inverted triangle** represents an unary relation, labelled by a unary relation name.
R	A **single-headed arrow** represents a relationship between the source and target, labelled by a binary relation name.
R	A **double-headed arrow** represents a pairwise relationship between the source and target, labelled by a binary relation name.

The examples in this paper are based on a simple case study for stock management of films and music in department stores. We demonstrate codecharts by specifying the system's structure and programming interface. Consider Fig. 1a, which contains three labelled rectangles (representing classes), two labelled single headed arrows (representing relationships) and a labelled inverted triangle (representing a property). This codechart states that the classes `Music` and `Film` both $Inherit$[2] from the $Abstract$ class `Media`.

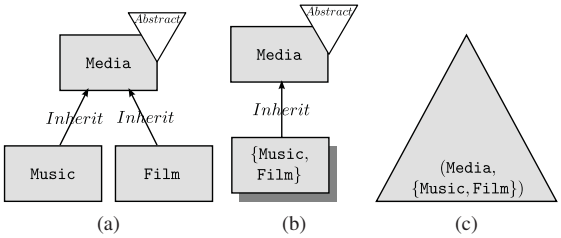


Fig. 1. Three codecharts showing class inheritance

Fig. 1b introduces a labelled offset rectangle that represents the set of classes $\{\texttt{Music}, \texttt{Film}\}$. The arrow sourced on the offset rectangle states that all those classes $Inherit$ from `Media`, thereby representing the same information as in Fig. 1a in a more succinct fashion.

Fig. 1c introduces a labelled triangle that represents an inheritance class hierarchy: a special set of classes that contains a single root class from which all other classes within the set $Inherit$. We label hierarchies using tuples that indicate the root class, in this case `Media`, and the remainder of the set, in this case $\{\texttt{Music}, \texttt{Film}\}$. This codechart succinctly represents the same class structure as in Fig. 1a and 1b, but no longer specifies that `Media` be $Abstract$. To overlap this triangle with $Abstract$ would inappropriately require that all classes in the hierarchy be $Abstract$.

[2]The semantics of such relationships are outside the scope of this paper.

While Fig. 1 introduced classes, properties and relations, Fig. 2 introduces methods and their signatures. Consider Fig. 2a which contains the classes `Film` and `Media` connected by $Inherit$, and three labelled ellipses that represent method signatures. The combination of a method signature and a class represents a specific method.

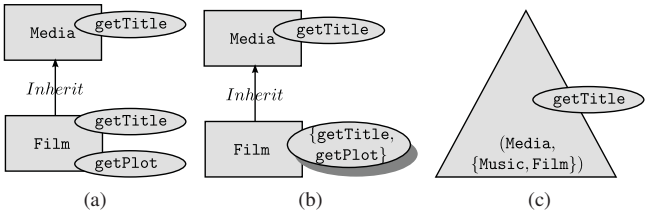


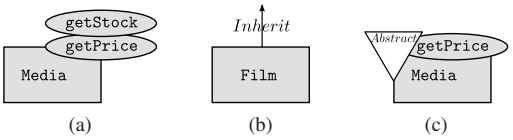
Fig. 2. Three codecharts showing methods

Fig. 2b introduces a labelled offset ellipse that represents the set of method signatures $\{\texttt{getTitle}, \texttt{getPlot}\}$. This offset ellipse overlapping `Film` specifies a set containing the two methods specified individually in Fig. 2a. Fig. 2b is therefore equivalent to Fig. 2a, yet uses fewer symbols.

Finally, Fig. 2c shows a method signature overlapping a hierarchy, which represents a set of methods which we will call M. M contains all methods with the signature `getTitle` implemented in `Media`, `Music` and `Film`. M must therefore contain the appropriate method in `Media` and any overriding methods in the subclasses (i.e. it may be a singleton set). This allows the structure of software to be designed without needing to specify implementation detail.

Our examples illustrate the syntax of codecharts with the exception of double headed arrows. Due to their more complex semantics, and space constraints, we do not discuss thier meaning in this paper. Instead we refer the reader to [4] for more details. However, we include them in our formalization of the syntax for completeness sake; syntactically they are subject to the same constraints as single headed arrows.

Through Fig. 1 and Fig. 2 we informally introduced syntactically valid codecharts. We now turn our attention to syntactically invalid 'codecharts', such as those in Fig. 3; an ellipse must overlap a (unique) rectangle (3a), an arrow must have a target (3b), and an inverted triangle must overlap a unique geometric shape (3c). Fig. 3a would require the method `getStock` to have a method, i.e. `getPrice`, as its domain which is clearly not meaningful. Fig. 3b does not specify from what the class `Film` inherits, and thus does not make a statement. Fig. 3c is ambiguous as it is unclear whether `getPrice` or `Media` is abstract. From the details given in [4] concerning the syntax of codecharts, these diagrams are not clearly prevented from being codecharts. By formally defining the concrete syntax of codecharts we become explicit as to what is, and is not, syntactically valid.



(a) (b) (c)

Fig. 3. Three syntactically invalid 'codecharts'

III. Concrete Syntax

To define our concrete syntax we observe the visual tokens of codecharts that are summarized in Table I. A codechart consists of a labelling function (λ) and nine finite sets, one for each permissible geometric shape: rectangles (R), ellipses (E), triangles ($T_{\triangle}$), their offset counterparts (OR, OE, $OT_{\triangle}$, respectively), inverted triangles ($T_{\triangledown}$), and single and double headed arrows ($A_{\blacktriangleright}$ and $A_{\blacktriangleright\blacktriangleright}$ respectively). Each geometric shape is labelled with a name drawn from specified, pre-defined, sets that give context to the codechart. Our approach of having pre-defined sets of labels (not seen previously in the existing codecharts literature) will help to define the semantics of codecharts via the abstract syntax.

A rectangle in R is labelled by a **class name** in the finite set $\mathcal{C}$. An offset rectangle in OR is labelled by a set of class names in the power set of $\mathcal{C}$ ($\mathbb{P}\mathcal{C}$). The labelling of ellipses and offset ellipses mirrors this structure, where we use **signature names** chosen from the finite set $\mathcal{S}$. Triangles in $T_{\triangle}$ are labelled by **class name hierarchies** that are elements of the following derived finite set:

$$\mathcal{CH} = \{(root, Set) : root \in \mathcal{C} \wedge Set \subseteq \mathcal{C} - \{root\}\}$$

Offset triangles in $OT_{\triangle}$ are labelled by sets of class name hierarchies in the power set of $\mathcal{CH}$ ($\mathbb{P}\mathcal{CH}$). Finally, inverted triangles in $T_{\triangledown}$ are labelled by **unary relation names** in the finite set $\mathcal{UR}$, and arrows in $A_{\blacktriangleright} \cup A_{\blacktriangleright\blacktriangleright}$ are labelled by **binary relation names** in the finite set $\mathcal{BR}$.

We formally associate labels with geometric shapes using a labelling function, λ. For example, consider Fig. 1a which contains three rectangles (r_1, r_2, r_3) an inverted triangle (t_1) and two single headed arrows (a_1, a_2). The rectangle r_1 is labelled by the class name `Media`, written $\lambda(r_1)$ = `Media`. Similarly r_2 and r_3 are labelled by the class names `Music` and `Film` respectively. The arrow a_1 that sources on r_2 and arrow a_2 that sources on r_3 are both labelled by the binary relation name $Inherit$ and both target r_1. Finally, the inverted triangle t_1 overlapping r_1 is labelled by the unary relation name $Abstract$. We formalize the notion of 'overlapping' geometric shapes as when the intersection of those sets of points constituting their interiors is not the empty set. We denote the interior of a geometric shape g by $int(g)$.

It is the geometric shapes and their labels that constitute a *concrete codechart*, subject to certain 'overlapping' constraints. This is defined as follows where, for convenience, we write PG (polygons) as shorthand for the set $R \cup T_{\triangle} \cup OR \cup OT_{\triangle}$ and PGC (polygons and curves) for $PG \cup E \cup OE$:

Definition 1. *A* **concrete codechart** *is a tuple*

$$(R, T_{\triangle}, E, OR, OT_{\triangle}, OE, T_{\triangledown}, A_{\blacktriangleright}, A_{\blacktriangleright\blacktriangleright}, \lambda)$$

where:

1) R *is a finite set of rectangles*
2) $T_{\triangle}$ *is a finite set of triangles*
3) E *is a finite set of ellipses*
4) OR *is a finite set of offset rectangles*
5) $OT_{\triangle}$ *is a finite set of offset triangles*
6) OE *is a finite set of offset ellipses*
7) $T_{\triangledown}$ *is a finite set of inverted triangles*
8) $A_{\blacktriangleright}$ *is a finite set of single headed arrows*
9) $A_{\blacktriangleright\blacktriangleright}$ *is a finite set of double headed arrows*
10) λ *is a labelling function that ensures*
    a) *for each* $x \in R$, $\lambda(x) \in \mathcal{C}$
    b) *for each* $x \in T_{\triangle}$, $\lambda(x) \in \mathcal{CH}$
    c) *for each* $x \in E$, $\lambda(x) \in \mathcal{S}$
    d) *for each* $x \in OR$, $\lambda(x) \in \mathbb{P}(\mathcal{C})$
    e) *for each* $x \in OT_{\triangle}$, $\lambda(x) \in \mathbb{P}(\mathcal{CH})$
    f) *for each* $x \in OE$, $\lambda(x) \in \mathbb{P}(\mathcal{S})$
    g) *for each* $x \in T_{\triangledown}$, $\lambda(x) \in \mathcal{UR}$
    h) *for each* $x \in A_{\blacktriangleright} \cup A_{\blacktriangleright\blacktriangleright}$, $\lambda(x) \in \mathcal{BR}$

such that the following constraints hold:

1) *each* $x \in E$ *overlaps a unique* $s \in PG$, $int(x) \cap int(s) \neq \varnothing$
2) *each* $x \in OE$ *overlaps a unique* $s \in PG$, $int(x) \cap int(s) \neq \varnothing$
3) *each* $x \in T_{\triangledown}$ *overlaps a unique* $s \in PGC$, $int(x) \cap int(s) \neq \varnothing$
4) *no two distinct* $a, b \in PG$ *overlap,* $int(a) \cap int(b) = \varnothing$
5) *each* $a \in A_{\blacktriangleright} \cup A_{\blacktriangleright\blacktriangleright}$ *is sourced on an element of* PGC*, written* $s(a) \in PGC$*, and targets an element in the same set, written* $t(a) \in PGC$

By this definition the following is the concrete codechart for Fig. 1a, where all sets not mentioned are empty:

- $R = \{r_1, r_2, r_3\}$, $T_{\triangledown} = \{t_1\}$, $A_{\blacktriangleright} = \{a_1, a_2\}$
- $\lambda(r_1)$ = `Media`, $\lambda(r_2)$ = `Music`, $\lambda(r_3)$ = `Film`, $\lambda(t_1) = Abstract$, $\lambda(a_1) = Inherit$, $\lambda(a_2) = Inherit$.

Having now defined the concrete syntax, we can discriminate between diagrams that use the prescribed visual tokens that are codecharts and those that are not codecharts.

IV. Abstract Syntax

The concrete syntax affords a logical structure over which we can reason about how a codechart is drawn, how many rectangles it contains and where they are positioned for example. However, in the semantics of codecharts the geometry is unimportant (other than overlaping, or the source/target of an arrow), whereas the topology of the diagram is important[3]. To avoid reasoning about irrelevent geometric properties and layout of codecharts we use an abstract syntax.

Our abstract syntax does not contain geometric information. It consists of the names used to label a concrete codechart, allowing us to reason about content without considering geometric properties that may be irrelevant, e.g. for determining the equivalence of codecharts. This approach, described in [5], is now standard in visual language development.

For example, Fig. 1a contains three rectangles in its R, labelled `Media`, `Music` and `Film` respectively. The appropriate *abstract codechart* does not contain any notion of

[3]We acknowledge that layout is highly important from a usability perspective, but that is not the focus of this paper.

rectangle, but instead collects these class names into the finite bag (multiset) $\mathcal{R}$. Similar bags exist for offset rectangles ($\mathcal{OR}$), (offset) ellipses ($\mathcal{E}$, $\mathcal{OE}$), (offset) triangles ($\mathcal{T}_{\triangle}$, $\mathcal{OT}_{\triangle}$) and inverted triangles ($\mathcal{T}_{\triangledown}$).

While the source and target of arrows are identified for free by virtue of their geometry in the concrete syntax, the abstract syntax must include this information explicitly. To this end, arrows are represented in the abstract syntax as tuples in the form $(label, source, target)$ in appropriate finite bags. For example, the arrow a_1 is represented as $(Inherit, \texttt{Music}, \texttt{Media})$ in $\mathcal{A}_{\blacktriangleright}$.

Now consider the three ellipses labelled `getTitle` and `getPlot` in Fig. 2a. The respective abstract codechart contains two occurances of `getTitle` and one occurance of `getPlot` in its finite bag of signature names, $\mathcal{E}$. Similar to arrows, we explicitly identify what each ellipse (and inverted triangle) overlaps using the function ω. To formally define this function we need access to it under domain restrictions, and therefore define the notation $\omega|_A$ to mean the function ω with its domain restricted to A. Below is the formal definition for the abstract syntax of codecharts where, for convenience, we write $\mathcal{PG}$ as shorthand for the set $\mathcal{R} \cup \mathcal{T}_{\triangle} \cup \mathcal{OR} \cup \mathcal{OT}_{\triangle}$ and $\mathcal{PGC}$ for $\mathcal{PG} \cup \mathcal{E} \cup \mathcal{OE}$:

Definition 2. *An* **abstract codechart** *is a tuple:*

$$(\mathcal{R}, \mathcal{T}_{\triangle}, \mathcal{E}, \mathcal{OR}, \mathcal{OT}_{\triangle}, \mathcal{OE}, \mathcal{T}_{\triangledown}, \mathcal{A}_{\blacktriangleright}, \mathcal{A}_{\blacktriangleright\blacktriangleright}, \omega)$$

where all bags are finite and:

1) $\mathcal{R}$ *is a finite bag whose elements are chosen from* $\mathcal{C}$
2) $\mathcal{T}_{\triangle}$ *is a finite bag whose elements are chosen from* $\mathcal{CH}$
3) $\mathcal{E}$ *is a finite bag whose elements are chosen from* $\mathcal{S}$
4) $\mathcal{OR}$ *is a finite bag whose elements are chosen from* $\mathbb{P}(\mathcal{C})$
5) $\mathcal{OT}_{\triangle}$ *is a finite bag whose elements are chosen from* $\mathbb{P}(\mathcal{CH})$
6) $\mathcal{OE}$ *is a finite bag whose elements are chosen from* $\mathbb{P}(\mathcal{S})$
7) $\mathcal{T}_{\triangledown}$ *is a finite bag whose elements are chosen from* $\mathcal{UR}$
8) $\mathcal{A}_{\blacktriangleright}$ *and* $\mathcal{A}_{\blacktriangleright\blacktriangleright}$ *are finite bags whose elements are chosen from*

$$\{(l, s, t) : l \in \mathcal{BR} \wedge s, t \in \mathcal{PGC}\}$$

9) $\omega : \mathcal{E} \cup \mathcal{OE} \cup \mathcal{T}_{\triangledown} \to \mathcal{PGC}$ *specifies overlaps such that:*
   a) $\omega|_{\mathcal{E}}$ *has codomain* $\mathcal{PG}$
   b) $\omega|_{\mathcal{OE}}$ *has codomain* $\mathcal{PG}$

By this definition the following is the abstract codechart for Fig. 2a, where the missing elements are empty bags and we view a bag as a set of pairs:

- $\mathcal{R} = \{(\texttt{Media}, 1), (\texttt{Film}, 1)\}$
- $\mathcal{E} = \{(\texttt{getTitle}, 1), (\texttt{getTitle}, 2), (\texttt{getPlot}, 1)\}$,
- $\mathcal{A}_{\blacktriangleright} = \{((Inherit, (\texttt{Film}, 1), (\texttt{Media}, 1)), 1)\}$,
- $\omega((\texttt{getTitle}, 1)) = (\texttt{Media}, 1)$,
  $\omega((\texttt{getTitle}, 2)) = (\texttt{Film}, 1)$,
  $\omega((\texttt{getPlot}, 1)) = (\texttt{Film}, 1)$.

V. Mappings between the Syntax Levels

Through §IV we indicated how concrete and abstract codecharts are related. For example, a rectangle in R is labelled by a class name that is a member of $\mathcal{R}$. That is, $\mathcal{R}$ contains as many occurrences of each class name as are used to label rectangles in R. To formally define this we require access to components of the concrete and abstract codecharts. Given a concrete codechart, c, (respectively abstract codechart C) the i-th component of c is denoted $\pi(c, i)$ ($\pi(C, i)$ respectively). For example, $\pi(c, 2) = T_{\triangle}$ and $\pi(C, 2) = \mathcal{T}_{\triangle}$. Given this, the mapping between syntax levels is formally defined as follows:

Definition 3. *Let c be a concrete codechart and let C be an abstract codechart. We say that c is a* **drawing** *of C provided there exists a function $f : PGC \cup T_{\triangledown} \to \mathcal{PGC} \cup \mathcal{T}_{\triangledown}$ such that the following hold:*

1) *for each i where $1 \le i \le 7$, $f|_{\pi(c,i)}$ is bijective with codomain $\pi(C, i)$ such that for each $x \in \pi(c, i)$ and some j: $f(x) = (\lambda(x), j)$*
2) *for each i where $8 \le i \le 9$, $f|_{\pi(c,i)}$ is bijective with codomain $\pi(C, i)$ such that for each $x \in \pi(c, i)$ and some j: $f(x) = ((\lambda(x), f(s(x)), f(t(x))), j)$*
3) *for all $x \in E$, x overlaps $s \in PG$ iff $\omega(f(x)) = f(s)$*
4) *for all $x \in OE$, x overlaps $s \in PG$ iff $\omega(f(x)) = f(s)$*
5) *for all $x \in T_{\triangledown}$, x overlaps $s \in PGC$ iff $\omega(f(x)) = f(s)$*

If c is a drawing of C then C is an **abstraction** *of c.*

VI. Conclusion

In this paper, we formally defined a concrete and abstract syntax of codecharts, and provided a mapping between them. The advantages of using both abstract and concrete syntax are well-documented and are now standard approaches in visual language development. Thus, our work has put codecharts on an equal footing with other formalized visual notations. The benefits of this approach in the case of codecharts will become more explicit when inference rules are developed and soundness and completeness properties are investigated.

In order to prove the soundness of inference rules it is necessary that the semantics are formally defined, which our work can be seen as a step towards. A key part in future work will be to describe how the core syntax of programming languages links to the names used in codecharts, such as linking **abstract** in Java to $Abstract$ and **extends** to $Inherit$. Given this we hope to investigate the applicability of codecharts to program metrics to ascertain an indication of program complexity or the level of dependencies and coupling present.

References

[1] C. Beeri, A. Eyal, S. Kamenkovich, and T. Milo. Querying business processes with BP-QL. *Information Systems*, 33(6):477–507, 2008.

[2] P. Bottoni, M. Koch, F. Parisi-Presicce, and G. Taentzer. A visualization of OCL using collaborations. In *Proceedings of the 4th International Conference on Unified Modeling Language (UML '01)*, volume 2185 of *LNCS*, pages 257–271. Springer, 2001.

[3] F. P. Brooks. No silver bullet: Essence and accidents of software engineering. *IEEE Computer*, 20(4):10–19, April 1987.

[4] A. H. Eden and J. Nicholson. *Codecharts: Roadmaps and Blueprints for Object-Oriented Programs*. Wiley-Blackwell, April 2011.

[5] J. Howse, F. Molina, J. Taylor, and S.-J. Shin. Type-syntax and token-syntax in diagrammatic systems. In *Proceedings of the international conference on Formal Ontology in Information Systems - Volume 2001*, FOIS '01, pages 174–185, New York, NY, USA, 2001. ACM.

[6] M. M. Lehman. Laws of software evolution revisited. In *Proceedings of the 5th European Workshop on Software Process Technology*, pages 108–124. Springer, October 1996.

[7] I. Oliver, J. Howse, G. Stapleton, E. Nuutila, and S. Törmä. Visualizing and specifying ontologies using diagrammatic logics. In *Proceedings of 5th Australasian Ontologies Workshop*, volume 112, pages 37–47, Melboune, Australia, December 2009. CRPIT.

[8] D. Stein, S. Hanenberg, and R. Unland. Query models. In *proceedings of the 7th International Conference on Unified Modeling Language (UML '04)*, volume 3273 of *LNCS*, pages 98–112. Springer, 2004.

[9] H. Störrle. VMQL: a visual language for ad-hoc model querying. *Journal of Visual Languages & Computing*, 22(1):3–29, 2011.

Does the Orientation of an Euler Diagram Affect User Comprehension?

Andrew Blake, Gem Stapleton
University of Brighton, UK
{a.l.blake,g.e.stapleton}@brighton.ac.uk

Peter Rodgers
University of Kent, UK
p.j.rodgers@kent.ac.uk

Liz Cheek, John Howse
University of Brighton, UK
{liz.cheek,john.howse}@brighton.ac.uk

***Abstract*—Euler diagrams, which form the basis of numerous visual languages, can be an effective representation of information when they are both *well-matched* and *well-formed*. However, being well-matched and well-formed alone does not imply effectiveness. Other diagrammatical properties need to be considered. Information visualization theorists have known for some time that orientation has the potential to affect our interpretation of diagrams. This paper begins by explaining why well-matched and well-formed drawing principles are insufficient and discusses why we should study the orientation of Euler diagrams. To this end an empirical study is presented, designed to observe the effect of orientation upon the comprehension of Euler diagrams. The paper concludes that the orientation of Euler diagrams does not significantly affect comprehension.**

I. Introduction

Euler diagrams represent set theoretic relationships using interconnected closed curves often drawn using circles or ovals. Curves are labelled, so affording context to the information or data therein. Figure 1 contains three Euler diagrams all representing the same information and illustrates that there are syntactic choices to be made when visualizing data. Each diagram tells us that Course Leaders are a subset of Lecturers, Lecturers are a subset of Academics and these staff could be Managers. Later in this section we will discuss the syntactic differences between the diagrams in figure 1.

Euler diagrams are regarded as a natural and effective way to depict sets and their relationships. They form the basis of numerous visual languages, including Swoboda and Allwein's Euler/Venn logic [1], Gil et al.'s spider diagrams [2], Kent's constraint diagrams [3], and Oliver et al.'s concept diagrams [4]. In the latter two cases, the visual languages are expressive enough to model complex properties of software. In addition, Euler diagrams are applied in a wide variety of other contexts including architecture [5], arts [6] and social media [7]. Wilkinson [8] presents a survey of natural science journals and online affiliated content from 2009 observing 72 occurrences of Euler diagrams. All of these uses of Euler diagrams demonstrate the importance of providing an account of how best to draw them in terms of user comprehension.

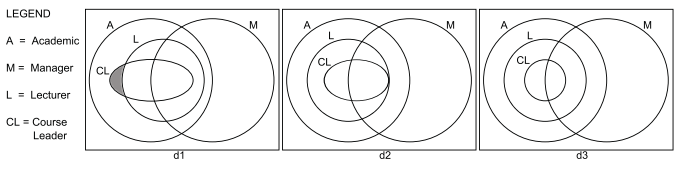


Fig. 1. Visualizations of a staff hierarchy.

We already have some insight into how best to draw Euler diagrams. In particular, we focus on two categories of so-called well-matched and well-formed drawing principles. These are designed to yield effective diagrams, where effective means reducing comprehension errors. Gurr, theorising well-matched diagrams, postulates that the most effective diagram is one with structure and property that matches, or closely matches, that which it strives to represent [9]. Well-formedness describes relationships between curves and regions in a diagram. There has been some work on empirically testing these well-formedness properties, observing the extent to which they impact comprehension [10], [11]. Gurr's theory tells us to select well-matched diagrams and the empirical work guides us to select well-formed diagrams in order to maximize effectiveness.

In figure 1, the diagram d1 is neither well-matched or well-formed. It is not well-matched as its shaded region denotes that there are no Course Leaders that are not Lecturers: the set of Course Leaders is contained by the set of Lecturers but CL is not contained by L. It is not wellformed as it has a disconnected zone: the region inside A and L but outside CL comprises two disconnected pieces. The diagram d2 is well-matched but not well-formed. It is not well-formed as it has a disconnected zone, as described earlier, and a brushing point where CL meets L. The diagram d3 is both well-matched and well-formed. It does not exhibit any extraneous properties and the relationship between its curves and regions are neither disconnected or brushing and, therefore, it is regarded the most effective at conveying this information pertaining to staff hierarchy. There are a number of other well-formedness conditions that a diagram can exhibit. One example, concurrency, exists when two or more curve segments follow the same path; see [11] for a full list.

To illustrate further differences in Euler diagram layout, the diagrams in figure 2 represent the same information as those in figure 1. Diagrams d3 to d6 are all well-matched and well-formed and are, by these properties, regarded as equally effective. However, there are clear visual differences between them. These differences can largely be attributed to the shape of their curves. Diagram d3 uses circles, diagrams d4 and d5 use ellipses and diagram d6 uses irregular shapes.

Diagrams d4 and d5 are basically identical except that diagram d4 has been rotated by 150 degrees to yield d5. Diagrams d3 to d6 visually illustrate that well-matched and well-formed drawing principles alone are too naive in yielding effective Euler diagrams. Given the current state of knowledge, we are unable to determine which of these diagrams is most effective.

Thus, in addition to well-matched and well-formed, there are other diagrammatical properties to consider when ascertaining the effectiveness of a visual representation. Perceptual theorists know that we are sensitive to the diagrammatical properties of orientation, shape and colour [12]. Aware of this phenomena, information visualisation theorists manipulate these properties, affecting our interpretation and, thus, comprehension of diagrams [13].

Conscious that other diagrammatical properties affect our interpretation of diagrams, this research aims to ascertain the whether orientation impacts user comprehension of Euler diagrams. This is a key question as studies of Euler diagram comprehension (such as [11]) have assumed that users' understanding of a diagram is not impacted by orientation. If this turns out to be a false assumption then such studies have additional confounding variances, not taken into account by the investigators. The remainder of this paper focuses on the question of orientation and, in doing so, presents an empirical study addressing the general question: does the orientation of an Euler diagram affect our comprehension?

The remainder of this paper is structured as follows. In section II we present the design of the experiment. Section III describes our research vehicle and section IV presents our experiment execution and results. Finally, section V discuses our conclusions and future work.

II. Experiment Design

We are aiming to establish whether the orientation of an Euler diagram affects user comprehension. In order to investigate this, we designed an empirical study which requires participants to answer questions concerning the information conveyed by Euler diagrams. In particular, the study uses a parallel group design with repeated measures within each of the two groups; we call these groups *participant group A* and *participant group B*. We chose a set of Euler diagrams which were displayed to the two groups of participants, with group A being shown the diagrams with one orientation and group B being shown the same diagrams in a different orientation. In each case, participants were asked a question concerning the information within the diagram.

Consistent with other researchers who have investigated user comprehension [14], [15], [16], we recorded the time taken to answer the questions as the primary dependent variable. Moreover, we also considered error rate. The independent variables were diagrams and rotation. If orientation impacts on comprehension then we would expect to see, for some diagram, a significant difference between the mean time taken to answer the posed question by participant group A to the mean time taken by participant group B.

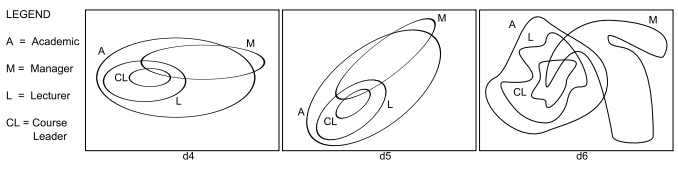


Fig. 2. Further illustrations of the staff hierarchy.

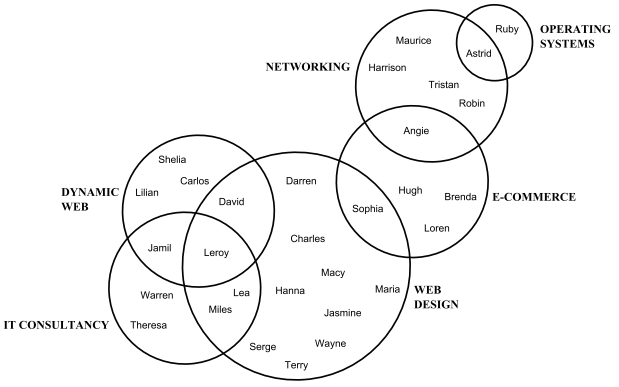


Fig. 3. An Euler diagram with 6 curves.

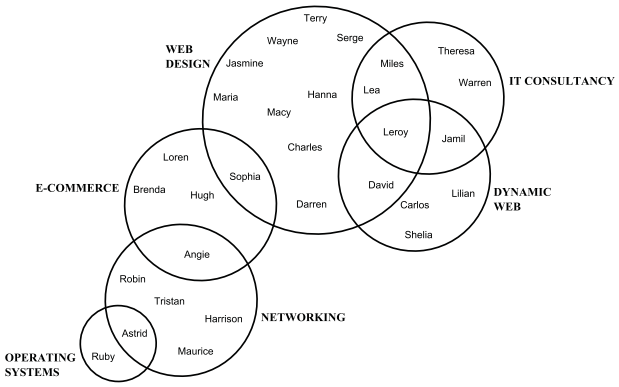


Fig. 4. Figure 3 rotated 180 °.

In designing the study, we have considered the following factors. First, we identified the types of information conveyed by Euler diagrams, to enable the construction of a range of questions for the study. Second, we carefully considered choices in diagram layout, in order to ensure that we minimize unwanted variation across diagrams. The subsections that now follow expand upon the considerations just described.

A. Euler Diagram Specification

To execute the study, we had to produce a range of Euler diagrams of which to ask questions. As in [11], we placed data items within the curves in order to enable meaningful questions to be asked. For instance, figure 3 tells us that the set of students studying the module OPERATING SYSTEMS is disjoint from the set of students studying E-COMMERCE. Figure 5 expresses that DATA STRUCTURES is being studied by the student Victor. Scaled versions of these diagrams were used in the study.

When drawing Euler diagrams, even those which are both well-matched and well-formed, there are numerous choices to be made, such as curve thickness or the relative positioning of labels. In order to minimize confounding variables, we adopted the following drawing conventions:

1) all diagrams were monochrome, drawn in an area of 765×765 pixels,
2) the curves used all had a 2 pixel stroke width and were circles,
3) the curve labels were written using upper case letters in Times New Roman, 14 point size, font in bold,
4) data items were written using lowercase letters, except that the first letter was capitalised, and with Ariel 12 point size font,
5) each curve label was positioned closest to its corresponding curve, and
6) data items were evenly distributed within the regions (called *zones*).

These conventions ensured that different diagrams had consistent layout features. Each diagram used in the study contained curves of three sizes, as seen in figure 5. Moreover, conforming to previous observations concerning user comprehension, all diagrams were well-matched and well-formed.

While striving to minimize confounding variables, it was deemed important there was some diversity in the diagrams, so that participants had to read and understand each diagram before being able to answer the posed question. The diagrams were chosen to have the following characteristics:

1) type 1: 4 curves, 9 zones and 20 data items,
2) type 2: 6 curves, 13 zones and 30 data items, and
3) type 3: 8 curves, 17 zones and 40 data items.

The premise for these choices is it allows diagrams to exhibit the range of basic set theoretic concepts, namely set inclusion, disjointness, and set intersection. Moreover, the diagrams needed to exhibit a reasonable level of complexity in order to demand cognitive effort on the part of the participant; having only a few curves, zones, or data items was deemed insufficient.

Our study used 6 diagrams for each of the three characteristic types, giving 18 diagrams in total for each set of participants. Each of the drawn diagrams was randomly rotated by an angle between $45\,^\circ$ and $315\,^\circ$ in order to remove possible bias arising from the manner in which the facilitator had drawn the diagram. These (rotated) diagrams were allocated to participant group A. Figures 3 and 5 are examples of two diagrams allocated to participant group A. These diagrams were copied and each randomly rotated a second time. These diagrams were allocated to participant group B. Figures 4 and 6 are examples of two diagrams allocated to participant group B which are the rotated copies of figures 3 and 5 respectively. The second random rotations were designed so that no diagram was within $\pm 45\,^\circ$ of either the original diagram or that obtained under the first rotation.

B. Data and Questions

We had to choose a context for the information displayed in our Euler diagrams. Our aim was that participants should be familiar with the context of the information, so that they did not need to learn anything except for how to interpret the Euler diagrams. Moreover, it was also considered important that the participants did not have any pre-exposure to the actual information represented. Since we anticipated that our participants would be university students, we decided to visualize information about fictional university modules and the students studying those modules. The module names were based on those commonly found in British undergraduate computing courses. Student names were taken to be first names only, a mixture of both male and female names, and reflected a variety of ethnicities.

Three styles of question were specified: 'Which', 'Who' and 'How'. Example questions are:

1) Which module is being taken by 5 students?
2) Who is taking INTERACTION DESIGN, HCI and OBJECT-ORIENTED DESIGN but not UML?

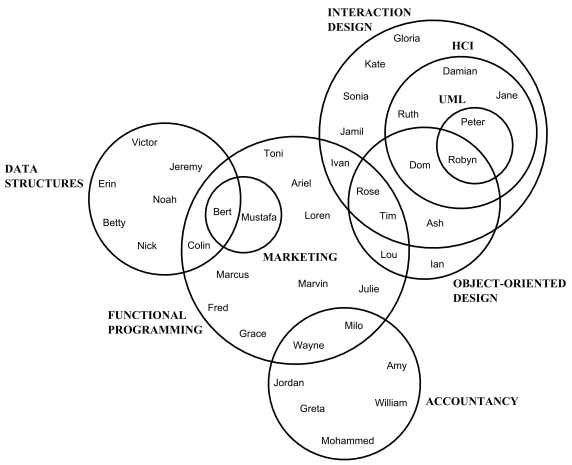


Fig. 5. An Euler diagram with 8 curves.

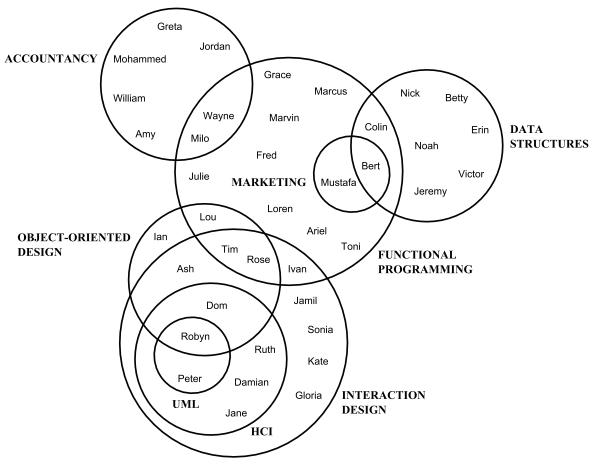


Fig. 6. Figure 5 rotated $148\,^\circ$.

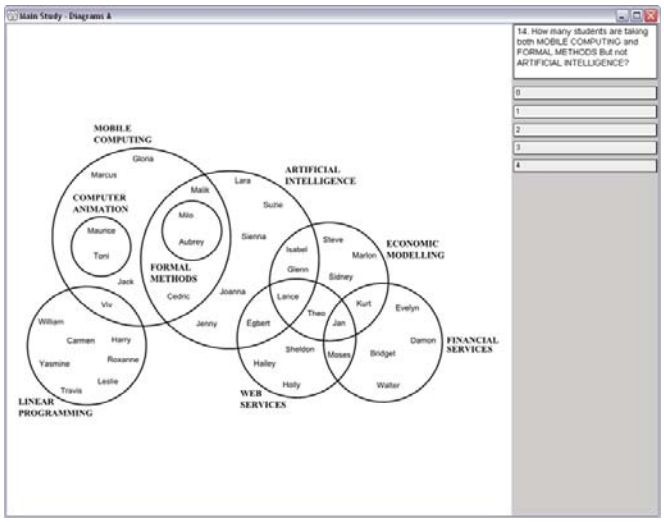


Fig. 7. Research Vehicle

3) How many students are taking both MOBILE COMPUTING and FORMAL METHODS but not ARTIFICIAL INTELLIGENCE?

Here, the first question was that asked of the diagram in figure 3 and of the rotation of it in figure 4. The second question was that asked of the diagram in figure 5 and of the rotation of it in figure 6.

There were 18 different questions in total, one for each of the 18 diagrams used in the study. The six diagrams of each characteristic type were allocated, between them, two of each style of question. All questions were multiple choice and had either 4 or 5 choices of answers; the correct answer was always unique.

III. Research Vehicle

To collect data during the study, we used a software tool (which we call the research vehicle) to present the questions to participants, gather the answers given to the questions and the time taken to reach each answer. Each time the participant answered a question, the research vehicle would ask the participant to indicate when they were ready to proceed to the next question, thus allowing them to pause between questions. Further, there was a maximum time limit of two minutes for each question. This was to ensure that each experiment did not continue indefinitely. Figure 7 is a screen shot of the research vehicle. It presents the third style of question, 'How', as specificed in section II-B. The research vehicle was used for two phases of the experiment, a *training phase* and a *data collection phase*.

The training phase was designed to give participants the opportunity to practise interpreting Euler diagrams and using the software, to avoid any learning effect during the actual data collection stage. In the training phase, each participant was presented with 6 Euler diagrams and their questions, one after the other, in a fixed order; these 6 diagrams were distinct from those used in the actual study. They were exposed to two examples of each question type and examples of 4, 6 and 8 curve diagrams.

The data collection phase presented the 18 chosen diagrams to each participant in a random order. The randomizing of the order of questions was an attempt to negate potential learning effects had the diagrams been presented in the same order throughout the study. It is the data from this phase that we analyze in order to test our hypothesis.

IV. Experiment Execution and Results

We are aiming to establish whether the orientation of Euler diagrams affects user comprehension. Specifically we want to test the null hypothesis that there does not exist an Euler diagram where the mean time taken to interpret the diagram is different when the diagram is oriented differently versus the alternative hypothesis that there exists at least one Euler diagram where the mean time taken to interpret the diagram is different when the diagram is oriented differently.

Our study recruited 32 participants, including six during a pilot phase. The participants were randomly allocated to either group A or group B; these groups were equal sizes. They were all undergraduate students from the University of Brighton's School of Computing, Engineering and Mathematics and they spanned all undergraduate years. The participants performed the experiment on campus within a usability laboratory which affords a quiet environment free from noise and interruption. Each participant was alone during the experiment, in order to avoid distractions, with the exception of an experimental facilitator who was present throughout. The same computer and monitor was used by each participant. The experiment took approximately 1 hour and participants were paid £6 to take part.

There were three phases to the experiment. Before participants entered the aforementioned training phase (which introduces the participants to the research vehicle), participants were introduced to the notion of Euler diagrams and the types of questions to be asked. This was achieved using hard copy printouts of three diagrams, with four, six and eight curves respectively, and with one question of each style. Participants were given a few minutes to study the diagrams and questions after which the experimental facilitator explained how to answer the questions. When the facilitator was happy that the participant clearly understood how each answer was derived, the participant was asked whether they were happy to proceed with the experiment.

The participants then entered the training phase, where they had the opportunity to use the research vehicle to answer questions. When all six questions were answered participants were shown data indicating questions answered correctly and how long each question took to answer. If a question was answered incorrectly the facilitator went through the question with the participant. The participants then entered the data collection phase of the study, where we collected the quantitative data.

Initially, a pilot study was undertaken involving six participants. The experimental design, method and research vehicle proved robust, with no changes required. Subsequently, the

main study was instigated involving a further 26 participants. All questions were attempted and completed comfortably within the two minutes allowed. There were no differences between the execution of the pilot study and the main study so their data sets were combined, consistent with [17]. Consequently, the following results are based on 32 participants and $18 \times 32 = 576$ observations.

A. Results and Analysis

In order to explore whether orientation impacts user comprehension, it is insightful to examine the box and whisker plot in figure 8. This illustrates that, for each diagram, the times taken to answer the question by participants in group A are very similar to the times taken to answer the question by participants in group B. Considering diagram 1, for example, we see that the interquartile ranges are almost identical across participant groups. In fact, the interquartile ranges for each diagram, by participant group, overlap substantially except perhaps for diagram 16. Despite these substantial overlaps, we can see that there is variation between the different diagrams, indicating that the study design is robust and fit-for-purpose. In summary, this plot indicates that orientation is unlikely to impact comprehension.

Source	DF	SS	MS	F	P
group	1	1.1012	1.1012	0.83	0.369
diagram	17	65.4768	3.8516	46.36	0.000
group*diagram	17	1.4599	0.0859	1.03	0.419
subj(group)	30	39.7617	1.3254	15.95	0.000
Error	510	42.3739	0.0831		
Total	575	150.1735			

TABLE I
ANOVA FOR LOG TIME.

To verify this insight, we conducted an analysis of variance (ANOVA) test, taking into account the diagram and participant group. In order to conduct this test, we require the data to be normal. However, a normal probability plot (not included here) revealed that the data are not normally distributed, but applying a transformation (in this case taking the log of the time taken) resulted in a normal data set. Using this transformed data set, the statistical calculations are included in table I.

First, we consider the row for group, which concerns difference in time taken between the two groups. Here, a p-value of 0.369 indicates that there was no significant difference in the mean time taken to answer the questions by the participants group A with the mean time taken by participants in group B. By contrast, there were significant differences between the mean times taken to answer questions about each diagram (ignoring the breakdown by participant group), with a p-value of 0.000 seen in the row for diagram. This indicates that there was a significant amount of diversity in our selected diagrams. Thus, these two p-values mean that we can safely and rigorously use the data to compare the affect of orientation.

The pertinent row, with regard to our hypothesis, is that for the interaction of group and diagram there is a differential effect of rotation among diagrams. A p-value of 0.419 means that there is insufficient evidence to reject the null hypothesis and we conclude that orientation does not affect user comprehension.

B. Error Results

Of the 576 observations there were a total of 19 errors giving an error rate of 0.03 or 3%. Table II lists each diagram which incurred errors. For each diagram, errors are distributed between participant groups.

	Group A	Group B	Total Errors
diagram 2	1	2	3
diagram 4	1	0	1
diagram 6	2	2	4
diagram 8	2	1	3
diagram 12	1	0	1
diagram 13	2	1	3
diagram 14	2	2	4
	11	**8**	**19**

TABLE II
ERRORS FOR EACH DIAGRAM BY PARTICIPANT GROUP

As can be seen from table II, the maximum error for a diagram under a participant group is 2. Therefore, there is little useful information that can be derived from this error data regarding orientation of Euler diagrams affecting user comprehension. It was observed in the introduction of this paper that well-matched [9] and well-formed [11] drawing principles are designed to reduce comprehension errors of Euler diagrams. The very low error rate found here reinforces the premise of these principles.

With the exception of diagram 13, the remaining 6 diagrams listed in table II conveyed information about a subset relationship which accounted for 16 of the 19 total errors. Of these 16 errors, 12 errors were for questions that referenced a curve that was completely contained by another curve, thus conveying a subset relationship. These 12 errors occurred from 4 questions phrased either:

1) 'Who is taking module A and module B but not module C?' or,
2) 'How many students are taking module A and module B but not module C?'

Of these 16 errors 4 other errors occurred from diagrams exhibiting subsets about which their questions did not reference. These questions required participants to count the number of students in a module and were phrased 'Which module is being taken by n students?' By contrast, of the 11 diagrams with no errors only three exhibited set inclusion. While well-matched and well-formed drawing principles appear to contribute to a very low error rate there is a notable bias in the nature of errors that do occur, specifically with diagrams exhibiting set inclusion.

V. CONCLUSIONS AND FURTHER WORK

In this paper we set out to establish whether the orientation of Euler diagrams affects user comprehension. To establish this we designed a parallel group study with repeated measures.

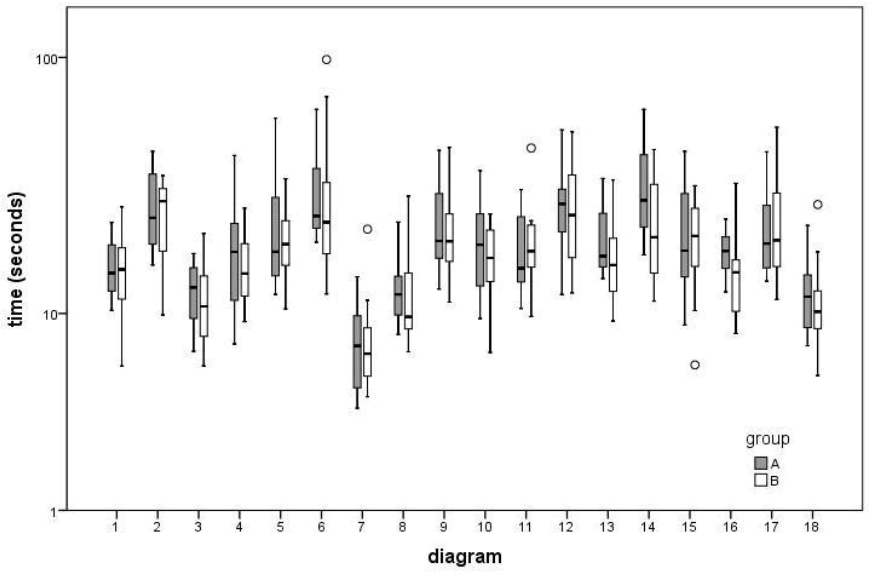


Fig. 8. Times taken (log scale) broken down by diagram and participant group.

We paid particular attention to the layout of the diagrams used within the study as well as their complexity. To ensure a degree of difficulty to the questions in our study the diagrams had up to 8 curves present, 17 zones and 40 data items. To ensure the participants, who were undergraduate students, were familiar with the question domain the diagrams visualised information about modules and students enrolled for them. This was to avoid any bias in the data due to learning a new context; their learning was limited to the diagrams. Our analysis of the collected data demonstrated that orientation does not affect user comprehension. The next phase of this research will be to explore the effect of curve shape and colour upon the comprehension of users.

Our result has implications for Euler diagram layout as well as future usability studies. In particular, people who draw Euler diagrams need not worry about the orientation from an effectiveness perspective and can now focus on other properties. In addition, our work supports current techniques for automated Euler diagram layout methods, such as [8], [18], [19], [20], [21], which do not pay any regard to orientation.

In terms of usability studies, our work underpins that in [11], which assumed that Euler diagram orientation does not impact user comprehension. Furthermore, this gives flexibility to the design of future studies, whereby empiricists no longer need to concern themselves with this aspect of diagram layout.

Acknowledgements Rodgers holds a Royal Society Industrial Fellowship. Stapleton is supported by an Autodesk Education Grant.

REFERENCES

[1] N. Swoboda and G. Allwein, "Heterogeneous reasoning with Euler/Venn diagrams containing named constants and FOL," *Euler Diagrams 2004*, ser. ENTCS, vol. 134. Elsevier Science, 2005.

[2] J. Gil, J. Howse, and S. Kent, "Formalising spider diagrams," in *IEEE Symposium on Visual Languages*. IEEE, 1999, pp. 130–137.

[3] S. Kent, "Constraint diagrams: Visualizing invariants in object oriented modelling," in *OOPSLA97*. ACM Press, 1997, pp. 327–341.

[4] I. Oliver, J. Howse, G. Stapleton, E. Nuutila, and S. Törma, "Visualising and specifying ontologies using diagrammatic logics," *Australasian Ontologies Workshop*, CRPIT vol. 112, 2009, pp 37-47.

[5] S. F. Lam, http://www.aadip9.net/shenfei//, accessed Sept 2011.

[6] Lukelab, http://www.lukelab.com/graphs/portrait-of-the-artists-interior-as-venn-diagram/, accessed Sept 2011.

[7] D. Skau, http://blog.visual.ly/euler-and-venn-diagrams/, accessed Sept 2011.

[8] L. Wilkinson, "Exact and approximate area-proportional circular Venn and Euler diagrams," *IEEE Trans. on Visualisation and Computer Graphics*, vol. 18, pp. 321–330, 2012.

[9] C. Gurr, "Effective diagrammatic communication: Syntactic, semantic and pragmatic issues," *Journal of Visual Languages and Computing*, vol. 10, no. 4, pp. 317–342, 1999.

[10] A. Fish, B. Khazaei, and C. Roast, "User-comprehension of Euler diagrams," *Journal of Visual Languages and Computing*, vol. 22, pp. 340–354, 2011.

[11] P. Rodgers, L. Zhang, and H. Purchase, "Wellformedness properties in Euler diagrams: Which should be used?" *IEEE Trans. on Visualisation and Computer Graphics, to appear*.

[12] A. Treisman, "Preattentive processing in vision," *Computer Vision, Graphics, and Image Processing*, vol. 31, pp. 156–177, 1985.

[13] R. Mazza, *Introduction to Information Visualisation*. Springer, 2009.

[14] P. Isenberg, A. Bezerianos, P. Dragicevic, and J. Fekete, "A study on dual-scale data charts," *IEEE Trans. on Visualisation and Computer Graphics*. IEEE, 2011, pp. 2469 – 2478.

[15] N. Riche and T. Dwyer, "Untangling Euler diagrams," *IEEE Trans on Visualisation and Computer Graphics*, vol. 16, pp. 1090–1097, 2010.

[16] H. Purchase, "Which aesthetic has the greatest effect on human understanding?" *International Symposium on Graph Drawing*. Springer, 1997, pp. 248–261.

[17] D. Altman, N. Burton, I. Cuthill, M. Festing, J. Hutton and L. Playle, http://www.nc3rs.org.uk/downloaddoc.asp?id=400, accessed April 2012.

[18] J. Flower and J. Howse, "Generating Euler diagrams," *Diagrams*. Springer, 2002, pp. 61–75.

[19] P. Rodgers, L. Zhang, and A. Fish, "General Euler diagram generation," *Diagrams*. Springer, 2008, pp. 52–68.

[20] G. Stapleton, L. Zhang, J. Howse, and P. Rodgers, "Drawing Euler diagrams with circles," *Diagrams*. Springer, 2010, pp. 23–38.

[21] P. Simonetto, D. Auber, and D. Archambault, "Fully automatic visualisation of overlapping sets," *Computer Graphics Forum*, vol. 28, no. 3, 2009, pp. 967–964.

A Video Game to Learn KeyScretch

Gennaro Costagliola, Mattia De Rosa, Vittorio Fuccella, Fabrizio Torre
University of Salerno
Via Ponte Don Melillo
84084 Fisciano (SA), Italy
{gencos, matderosa, vfuccella, ftorre}@unisa.it

Abstract

KeyScretch is a text entry method for mobile devices equipped with touch-screens, based on a menu-augmented soft keyboard. It improves the previously studied menu-based methods by enabling the interpretation of compound strokes, corresponding to the input of particularly frequent character sequences. Although good performances can be reached in reasonably short times (2-3 hours of practice), the method does have a learning curve and the initial low performance could discourage the users. In this paper we present a video game to learn KeyScretch. The game helps the users to learn the method while having fun, letting them practice on the most frequent gestures associated to text and reach an acceptable expertise level.

1 Introduction

Text entry is one of the main tasks for mobile device users. Mostly, text is entered to compose SMSs, even though the growing computing capacity of the devices suggests that they will be increasingly used for richer text-based applications. Nevertheless, text entry on mobile devices is inefficient mainly due to their small size.

KeyScretch is a text entry method for mobile devices equipped with touch-screens, based on a menu-augmented soft keyboard. A menu-augmented soft keyboard is a keyboard whose keys have a radial menu surrounding them. Each menu item contains a (particularly frequent) character. The menu is shown as soon as the key is pressed and the selection of a menu item through a radial swipe is a shortcut to enter the character associated to it. *KeyScretch* improves the previously studied menu-based methods by enabling the interpretation of compound strokes, corresponding to the input of particularly frequent character sequences.

The *KeyScretch* text entry method has been presented in a previous paper [1]. The results of an analysis of the method's performance showed that *KeyScretch* significantly outperforms the baseline tapping-based method, with a crossover occurring after about two hours and a half. Although good performances can be reached in such reasonably short times, the method does have a learning curve and the initial low performance may discourage the users. Furthermore, the text chunks entered with the method have a variable length and the users may initially have some difficulties in correctly segmenting the text.

For the above reasons, we designed a playful solution that allows the user to have fun while learning the method correctly and quickly. Ideas of this sort are not new in literature: many typing tutors have been developed for physical keyboards; in fact typing-games are useful to improve typing speed [5]. One of the most famous is *Typing of the Dead* of Sega Entertainment (based on FPS *The House of the Dead*) in which a player must write a sentence that appears next to a zombie in order to kill him; the player loses a life if s/he fails to write the sentence before being reached by the zombie. A variant of this game was proposed for learning how to write Japanese characters [7]. Two other very popular games, focusing on proper typing the words on the keyboard and reserved for children of primary school, are *Tux Type* [8] and *GCompris* [3]. They both are distributed with an Open Source license. Furthermore, although using different skins, the two games share the same goal: typing letters on the keyboard before, falling from the top of the screen, they would disintegrate on the ground.

In the context of touch keyboards, a game [4] was proposed to facilitate the learning of the *Shape Writer* [6] text entry method. In this game the player must explode displayed balloons by typing (with *Shape Writer*) the words contained inside the balloons; the balloon is lost (and it becomes impossible to explode it) if it reaches the top of the screen before the user has typed the word.

Our typing-game is called *TypeJump* and is custom designed for *KeyScretch*. The player must type a presented word (or a short phrase) before the expiring of a timer. This objective is embedded in a game skin, where a stick man has to reach a destination before the timer expires. The game tells the user how to correctly segment the text, thus helping the user in one of the most difficult tasks occurring at the beginning of method's learning.

This research is a work in progress. In the present paper we introduce the game and the ideas behind it. Only preliminary tests with pilot users have been carried out so far.

Figure 1: The QWERTY keyboard layout augmented with a menu.

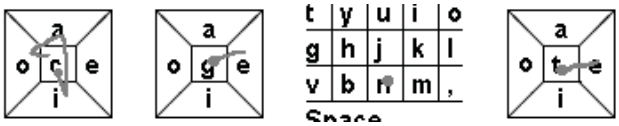


Figure 2: The 4 strokes needed to enter the text 'ciao gente'.

Deeper studies are left for future work.

The rest of the paper is organized as follows: the next section briefly describes the *KeyScretch* method; Section 3 presents the design of the *TypeJump* video game; finally, Section 4 offers our conclusions and outlines the future work.

2 The *KeyScretch* Text Entry Method

The *KeyScretch* method has been recently introduced by Costagliola et al [2, 1]. This section is a summary of the description of the method and its performances, partially presented in the above cited papers and partially still unpublished. The method allows the user to enter text on a soft keyboard through both *taps* and more articulated strokes. In this method each soft key is augmented with a menu containing frequent letters which appears whenever the key is pressed. Each gesture is initiated on a key and is driven by the key surrounding menu. A space character is inserted by ending the gesture inside the character key area or through a direct use of the space bar.

KeyScretch can be instantiated differently for different languages. In particular, the number of menu items, the characters to associate to the menu items and their arrangement must be chosen. In a recently tested instance, optimized for the Italian language, the vowels 'a', 'e', 'i' and 'o' are associated to each of the sides of the character keys, as shown in Figure 1. With this menu layout, the interaction sequence necessary to enter the Italian text *ciao gente* (*hello folks*, in English) is shown in Figure 2. The string is ten characters long but it can be entered with a sequence of four strokes (*taps* or *gestures*). The strokes correspond to the input of the following sequence of text chunks $\{ciao\ \}\{ge\}\{n\}\{te\}$. Three text chunks out of four are entered through a stroke and only one through a *tap*. The first stroke allows us to enter up to five characters.

In general, an instance of *KeyScretch* induces a segmentation of the text into a sequence of *text chunks*, each produced through a stroke. The way an instance segments the text can be expressed through a rule put in the form of a regular expression. In *KeyScretch*, using a menu containing n characters $x_1, \ldots, x_n$ a stroke produces a *text chunk* described by the following regular expression:

$$.[x_1 x_2 \ldots x_n] + [\ ]? \tag{1}$$

The above pattern matches a text chunk starting with any character (specified in (1) by the starting '.') chained to a sequence of one or more characters from the $\{x_1, x_2, \ldots, x_n\}$ set, (specified by $[x_1 x_2 \ldots x_n]$+) possibly ending with a space character, ([ ]?).

The *KeyScretch* method has been extensively evaluated. In particular, the performance of the above introduced Italian instance has been assessed in a study with eight users; the user satisfaction has been evaluated through questionnaires; the expert performances of instances optimized for different Western languages have been estimated through a predictive model. The results of the evaluation can be summarized as follows:

- In the user study, performed by typing on a soft keyboard with a Qwerty layout, *KeyScretch* significantly outperformed the traditional method based on tapping on a soft keyboard. In particular, a speed crossover (the point where the performance with *KeyScretch* exceeds that obtained with the traditional method) was obtained after about two hours and a half of practice. At the end of the study (after about six hours of practice), the users reached a typing speed close to 40 WPM with *KeyScretch*, outperforming the traditional method of 17%. Furthermore, the users expressed a positive judgement on the method, appreciating its speed and comfort above all.
- The capacity of the users of correctly segmenting the text was evaluated by comparing the written text to the original text segmented through a text scanner. As expected, we observed that the users improved their capacity over the experiment, with the largest improvement obtained in the first 15 minutes. In the remaining time, their way of segmenting the text did not differ much from the ideal segmentation.
- The estimation of the expert performance revealed that specific instances of the method enable improvements within the range of 30%-49% against the traditional method on the Qwerty layout, according to the language used. The method always outperforms the simple menu-based method and improvements are also obtained against *Shape Writer* [6] with most language/layout configurations.

3 The *TypeJump* Video Game

The objective of the video game is to allow the user to quickly learn *KeyScretch*. In particular, the user can:

- Learn the correct way to segment the text. This is critical in the earliest learning phases: although the experiments have shown that a reasonably short time (about a quarter of hour) is necessary to learn an acceptable way to perform segmentation, if the user does

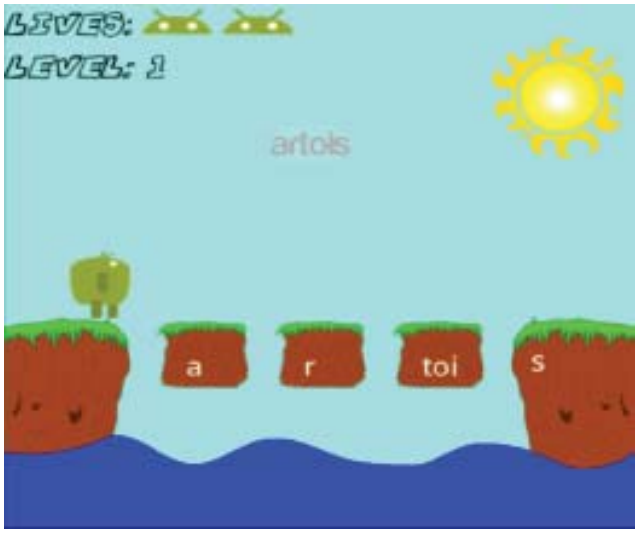


Figure 3: The *TypeJump* interface.

not understand one of the most important concepts underlying the method, s/he may abandon the use of that method, without investing additional time in learning. Learning by example can accelerate this process and prevents users from disaffecting to the method before fully understanding its functioning.

- Acquire an adequate writing speed. The experiments presented in the previous section show that more than two hours are necessary to obtain a writing speed equal to that of the traditional method. The fact that these hours are spent playing, makes the learning time enjoyable and prevents users suffering from the frustration of having to type slowly in other text-based applications.

3.1 The Interface

A screenshot of the video game interface is shown in figure 3. The game stage is entirely contained in the screen area. The game requires the user to personify a character, which has the appearance of the Android robot. The robot is initially located on the mainland at the left of the screen and his objective is to reach the mainland on the right of the game stage. In order to do this, the character has to jump through a sequence of rocks, located between the two mainlands.

The game is controlled through the *KeyScretch* soft keyboard. In order to reach the target mainland, the user has to enter a short text sentence. Each rock and the target mainland contain a text chunk of the sentence. A stroke (a simple tap or a gesture) must be performed to let the robot jump to the next rock. The whole text contained in the next rock must be entered through a single stroke, otherwise the android performs an ineffective jump (not long enough to reach the rock) and falls down into the sea. An ineffective jump means the loss of a life. The game is timed, in fact

(a) The initial screen

(b) The final screen when the game has been won

(c) The game over screen

Figure 4: Screenshots of the *TypeJump* game interface.

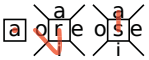
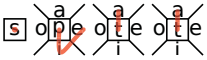
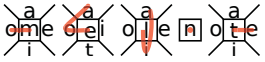
Chunks	Level	Sentence	Typing
2	1-5	rea-to	
3	1-5	a-rio-sa	
4	3-5	s-pie-ta-ta	
5	3-5	mo-toa-lia-n-te	

Table 1: Sentence examples.

the sea level rises as time elapses. As soon as the sea level reaches the surface of the rocks, the android falls down into the sea all the same.

The game ends when:

- all lives are terminated and the game is over (see the screenshot in Figure 4c);
- the user has been able to write the sentences loosing at most two lives. In this case, the user wins the game (see the screenshot in Figure 4b).

The game is divided into different levels. The difficulty increases from level to level: at the earliest levels, the sentences to enter are shorter and the sea level rises slowly, while at the most advanced levels, sentences of increasing complexity must be entered and the sea level quickly reaches the rocks.

3.2 The Sentences

The sentences are picked at random from a repository. Each sentence is divided in text chunks respecting the correct *KeyScretch* segmentation. In the easiest 2 levels, only the sentences whose number of segments is smaller than 3 are selected. In subsequent levels, also sentences with 4 or 5 chunks can be presented. Sentences with more than 5 chunks have been avoided since the inclusion of a high number of rocks in the game stage has an unpleasant visual effect.

The sentence repository contains a balanced number of the most frequent gestures. Sample sentences from the repository, composed of a different number of text chunks,

are shown in Table 1. The last column of the table reports the sequence of strokes needed to enter the sentence.

3.3 Implementation

Both the game and the *KeyScretch* method have been developed for the Android system. While the game has been implemented as an application, the keyboard has been implemented as an Android service. The difference is that a service can also be used with other applications, such as text editing applications and so on. A picture of a smartphone showing both the video game and the keyboard is in figure 5.

The application implementing the game is composed of a single Android Activity, which is instantiated for every single sentence. The application is designed to receive input from the *KeyScretch* keyboard. A *TextWatcher* module checks the changes in the input text and performs the animation associated to a jump upon the reception of the text.

The keyboard prototype was implemented as a customization of the Android *SoftKeyboard* sample project. The original view was slightly modified to enable the visualization of the menu around the peripheral keys: the width of the keys of the upper and central key lines have been narrowed to leave enough space on the sides, while the vertical size of the view has been enlarged to leave a space in the upper part of the keyboard. Although the keys have a rectangular shape, the menu is squared. The proportion between width and height is different between keys placed in different key lines, but is constant between keys in the same line. As a form of feedback, after a gesture has been performed by the user, the recognized template is printed in superposition to the stroke for a fraction of a second.

4 Conclusions and Future Research

In this paper we have presented *TypeJump*, a video game to learn the *KeyScretch* text entry method. The game helps the users to learn the method while having fun, letting them practice on the most frequent gestures associated to text and teaching them the text segmentation induced by the method. This way, the effort required to reach an acceptable expertise level is alleviated by the pleasure of playing and a lower disaffection to the method is expected.

At present, only an initial test with a few pilot users has been carried out. Users have expressed their appreciation for the game, which was considered enjoyable and exciting. Furthermore, they responded affirmatively when asked whether the game encouraged them in keeping using the *KeyScretch* method.

Future studies will be aimed at completing the experiments. In particular, we would like to gather experimental results on the learning of the method through the video game. In the presence of a positive satisfaction level with the game, just matching the learning curve already obtained in previous experiments would be considered a good result, since the same time is spent in a pleasant activity.

Figure 5: The interface showing both the *TypeJump* application and the *KeyScretch* service.

5 Acknowledgement

The authors thank Dott. Giuseppe Faggiano for his help in the development of the video game.

References

[1] G. Costagliola, V. Fuccella, and M. Di Capua. Text entry with keyscretch. In *Proceedings of the 16th international conference on Intelligent user interfaces*, IUI '11, pages 277–286, New York, NY, USA, 2011. ACM.

[2] G. Costagliola, V. Fuccella, M. Di Capua, and G. Guardi. Performances of multiple-selection enabled menus in soft keyboards. In *DMS'09 -The 15th International Conference on Distributed Multimedia Systems*, pages 359–364, Redwood City, CA, USA, Sept. 2009.

[3] GCompris. http://gcompris.net/, 2012.

[4] P. O. Kristensson and S. Zhai. Learning shape writing by game playing. In *CHI '07 extended abstracts on Human factors in computing systems*, CHI EA '07, pages 1971–1976, New York, NY, USA, 2007. ACM.

[5] C.-H. Lin and E. Z.-F. Liu. Transactions on edutainment iii. chapter A comparison between drill-based and game-based typing software, pages 48–58. Springer-Verlag, Berlin, Heidelberg, 2009.

[6] Shapewriter. http://www.shapewriter.com/, 2010.

[7] K. Stubbs. Kana no senshi (kana warrior): a new interface for learning japanese characters. In *CHI '03 extended abstracts on Human factors in computing systems*, CHI EA '03, pages 894–895, New York, NY, USA, 2003. ACM.

[8] TuxType. http://tux4kids.alioth.debian.org/tuxtype/, 2012.

DMS/DET Demo

Semi-automatic requirement tracing in modified code: An Eclipse Plugin

Daniel Hanspeter, Andrea Janes, Alberto Sillitti, Giancarlo Succi
Free University of Bozen/Bolzano
Center for Applied Software Engineering
Bolzano, Italy
daniel.hanspeter@stud-inf.unibz.it, {andrea.janes, alberto.sillitti, giancarlo.succi}@unibz.it

***Abstract*—**Software developers are interested in requirement traceability to e.g., verify if all requirements are covered by a system design specification. Based on the assumption that related artifacts contain related terms, researchers have developed, used, and extended algorithms that identify related terms and subsequently infer which artifacts are related (i.e., there is a traceability link between them). Source code is not as verbose as a natural language description, which reduces the applicability of algorithms that precisely rely on such a commonality. This paper extends the Vector Space Model using tf*idf term weights to improve the identification of traceability links between source code and requirements. To this extent, we modify the way how requirements are identified and to include user feedback. We show that the inclusion of user feedback significantly improved the number of correctly identified requirements and present an application as an Eclipse[1] plugin.

Keywords-requirement traceability; information-retrieval; recommender systems

I. INTRODUCTION

Algorithms that identify traceability links between natural language artifacts are less accurate when applied between formal language (like source code) and natural language artifacts [1][2]. This is mainly because source code is not as verbose as a natural language description. For example, redundancies are avoided to improve maintainability, while in natural language redundancy can be used to ensure that something is understood correctly. Similarly, ambiguities in natural language is a form to communicate e.g., irony. Formal or controlled natural languages have explicit constraints on grammar, lexicon, and style to reduce ambiguity and complexity [3]. These differences reduce the applicability of algorithms that precisely rely on such a commonality [1].

This paper proposes to extend an IR algorithm to retrieve requirements based on source code and to include user feedback to improve the identification of traceability links between source code and requirements. The modifications are inspired from recommender systems in which similarity is used to improve recommendations and user feedback is used to personalize the recommendation [4].

II. APPROACH

The tool consists in an Eclipse view that displays the predicted requirement for the currently modified / observed source code. This code is extracted by the tool from the user's cursor position in the editor. Furthermore, the user can give feedback by proposing the right requirement for the current prediction. In this way, the tool adapts and changes its decisions for future predictions.

We divided the task of combining an IR algorithm with user feedback to improve the identification of traceability links between source code and requirements in four steps: 1) identify the most suited IR algorithm that identifies traceability links between source code and requirements in our context, 2) implement a basic version of the identified algorithm, 3) extend the algorithm to improve the precision for source code similarity, and 4) consider user feedback.

A. Selection of the IR algorithm

We performed a literature review from which we found out that the VSM in combination with TF*IDF or LSI is the most popular algorithm (e.g. [1][5][6][7][8][9]). In a first experiment, we determined which of the two term weighting methods are more suitable for us to identify requirement to source code traceability links.

We used 20 source code fragments, to evaluate if the average number of correctly recognized requirements differ between VSM/TF*IDF and VSM/LSI. We performed a one-tailed Mann–Whitney–Wilcoxon signed-rank test [10] with H_0: VSM/TF*IDF and VSM/LSI identify the correct requirement equally well and H_1: VSM/TF*IDF identifies more identify traceability links than LSI. With a p-value of 0.03764 and Wilcoxon's test statistic $W = 62$ we rejected H_0 and confirmed that – in our context – VSM/TF*IDF performs significantly better (i.e., identifies more requirements correctly) than VSM/LSI.

B. Basic implementation of VSM/TF*IDF

We implemented the VSM/TF*IDF algorithm within an Eclipse plugin. The plugin downloads the requirements from an issue management system and the source code from a configuration management system. Terms are obtained tokenizing the text representing the requirements and the source code after removing stop-words, boundaries, abbreviations, and phrases [11] and after reducing words to their stem form.

The plugin populates the RTM on startup, marking which requirement document r and which source code document d contains a specific term t. When the programmer modifies source code, the terms contained in the method surrounding the current cursor position are extracted and become the query

[1] http://www.eclipse.org

vector used to identify similar requirements.

As said in the introduction, source code is not as verbose as a natural language description. Consequently, methods – that are usually short units of source code – contain only few terms that can be used to search for related requirements, which renders this approach unfeasible in practice.

C. Extending the query vector to consider related source code for the similarity calculation

To reduce the problem of having too few terms to identify the related requirements, we extend the query vector to include more terms than only those of the modified method. This allows to predict requirements not only by their key identifying terms in the current modified method, but also by using those terms to identify similar requirements linked to methods having similar content.

Assuming that the modified source code fragment f is part of a document d_m and contains the terms t_q, for every term t in the RTM that is not yet part of t_q, we verify if f has a similarity of $sim(d_x, f)$ to another source code document d_x, $x \neq m$, higher than a defined threshold sim_{min}[2]. If this is the case, we add the missing term t to f with a term weight of $sim(d_x, f) \times tf*idf(t, d_x)$ if and only if t is part of a requirement. $sim(d_x, f)$ is the VSM/TF*IDF similarity between d_x and f, $tf*idf(t, d_x)$ is the tf*idf weight of the term t in the document d_x.

This approach adds terms to the query vector that are contained in similar source code artifacts but with a weight that models also the similarity of the source code fragment f to d_x.

D. Adding user feedback

We use user feedback to improve the similarity calculation between source code and the requirements. The user can use the Eclipse plugin to a) see the predicted requirement and b) to change it.

When the user chooses a requirement r for source code fragment f, the plugin adds the terms contained in f to the terms contained in requirement r. When the similarity between the source code fragment and requirement r is calculated now, it is higher, since more similar words exist in both texts.

III. Evaluation

We tested our prototype with five different sets of source code and requirements. The first set (20 methods and 13 requirements) is from an Android[3] application developed during a bachelor thesis, the second set (30 methods and 18 requirements) is from a Java project developed during a research project, the third set (30 methods and 13 requirements) is from Freemind[4], the fourth set (30 methods and 12 requirements) is from JGraphT[5], and the fifth set (30 methods and 13 requirements) is from JavaNCSS[6]. The last three projects are all Open Source Java software.

We were interested to understand how the inclusion of similar source code documents to modify the search query and correct user feedback influences the tracing of source code artifacts back to their originating requirement. We formulated the null and alternative hypotheses as follows:

- $H_{0,1}$: The adaptation of the similarity calculation does not influence automated traceability of requirements in source code artifacts.
- $H_{0,2}$: The use of user feedback does not influence automated traceability of requirements in source code artifacts.
- $H_{a,1}$: The adaptation of the similarity calculation influences automated traceability of requirements in source code artifacts.
- $H_{a,2}$: The use of user feedback influences automated traceability of requirements in source code artifacts.

The evaluation consisted in the comparison of results achieved using a) the use of pure VSM/TF*IDF and b) the modified VSM/TF*IDF including the correct user feedback for each artifact. The null hypothesis has been tested using a paired, nonparametric test, the Mann– Whitney–Wilcoxon paired test [10]. With a significance level of 95%, the null Hypothesis $H_{0,1}$ was accepted, the null Hypothesis $H_{0,2}$ rejected. $H_{0,2}$ was rejected in the first project with a Wilcoxon's test statistic of W = 42, in the second project with W=63,5, in the third with W=76, in the forth with W=34 and in the fifth with W=62.5. The results gathered for both projects show that user feedback has a significantly higher average correct prediction.

IV. Discussion

The experiment described in this paper showed that in our two projects the adapted algorithm could not significantly improve the identification of traceability links. The consideration of user feedback on the other hand significantly increased the correctly identified the trace to the requirements.

References

[1] G. Antoniol, G. Canfora, G. Casazza, A. De Lucia, and E. Merlo, "Recovering Traceability Links between Code and Documentation," Transactions on Software Engineering, 28(10):970-983, 2002.

[2] A. Marcus, J. I Maletic, and A. Sergeyev, "Recovery of traceability links between software documentation and source code," International Journal of Software Eng. and Knowledge Engineering, 15(5):811-836, 2005.

[3] R. I. Kittredge, "Sublanguages and controlled languages," Oxford University Press, 2003.

[4] F. Ricci, L. Rokach, B. Shapira, and P.B. Kantor. Reccomender Systems Handbook. 1st Edition, Springer-Verlag, 2011.

[5] J.H. Hayes, A. Dekhtyar, and S.K. Sundaram, "Improving after-the-fact tracing and mapping: supporting software quality predictions," IEEE Software, 22(6):30-37, 2005.

[6] A. De Lucia, F. Fasano, R. Oliveto, and G. Tortora, "Recovering traceability links in software artifact management systems using information retrieval methods," ACM Transactions on Software Engineering and Methodology, 16(4):13-63, 2007.

[7] H.G. Gross and M. Lormans, "Towards Software Component Procurement Automation with Latent Semantic Analysis," Proceedings of the Third International Workshop on Coordination and Adaption Techniques for Software Entities, pp. 51-68, 2007.

[8] J. Cleland-Huang, R. Settimi, D. Chuan, and Z. Xuchang, "Utilizing supporting evidence to improve dynamic requirements traceability" 13th IEEE International Conference Requirements Engineering, pp. 135-144, La Sorbonne, France, 2005.

[9] Hsin-Yi 16Jiang, T.N. Nguyen, Ing-Xiang Chen, H. Jaygarl, and C.K. Chang, "Incremental Latent Semantic Indexing for Automatic Traceability Link Evolution Management," Proceedings of the 2008 23rd IEEE/ACM International Conference on Automated Software Engineering, pp. 59-68, L'Aquila, Italy, 2008.

[10] W.J. Conover, Practical Nonparam. Statistics, 3rd Edition, Wiley, 1998.

[11] R. Grishman, "Natural language processing," Journal of the American Society for Information Science, 35(5):291–296, 1984.

[2] We chose a threshold of 0.30 in our experiments. The optimal value of this threshold has to be evaluated in future investigations.

[3] http://www.android.com

[4] http://freemind.sourceforge.net

[5] http://www.jgrapht.org

[6] http://www.kclee.de/clemens/java/javancss

An Interactive Learning and Assessment System for Simulation-based Science Education Using Cloud Computing Technology

Mengxia (Michelle) Zhu
Associate Professor
Department of Computer Science
Southern Illinois University
Carbondale, IL 62901, USA
Email: mzhu@cs.siu.edu

Frackson Mumba
Associate Professor
Department of Curriculum and Instruction
Southern Illinois University
Carbondale, IL 62901, USA
Email: frackson@siu.edu

***Abstract*—There are many concepts in physics, chemistry, biology and earth science that involve complicated molecular dynamics in the K-12 curriculum. It is hard for teachers to instill the minds of young children with the knowledge using traditional instructional materials since the process is invisible and hard to explain. Research shows that multimedia interactive computer simulations can offer a highly attractive inquiry-based learning environment and provide active problem solving experience to students. In addition, such simulation-based virtual lab provides an alternative to real lab practices, which might be expensive, impossible or too long or dangerous to conduct. Consequently, computation simulation and visualization techniques are becoming indispensable components in many E-learning systems to promote comprehensive and thorough understanding of science and math concepts. Currently, there are many existing and ongoing federal and state funded projects dedicated to developing open-sources or non-open source simulation and animation software packages for K-12 students. Several such systems include Molecular Workbench (http://workbench.concord.org), PhET (http://phet.colorado.edu), Chemistry Experiment Simulations (http://cse.edc.org/products/simulations/) and NanoHUB (http://nanoHUB.org]. However, these systems are not widely used as many teachers are either not aware of their existence or cannot f nd the lesson and assessment questions to their particular interests.**

In order to address these above-mentioned issues, we design and develop a cloud computing driven E-learning and assessment system, which integrates popular free open-source simulations developed for K-12 education with assessment capacity. Users can search and download simulations at this one-stop warehouse without visiting different web sites. Graphical user interface enables a user-friendly environment for teachers to create courses, accept or reject enrollment request, link simulation, create and post assignments, and grade assignment. Many question types such as multiple choices, true or false, word bank, match question and free answer questions are supported to meet various teaching needs. Some statistic analysis functions are included to help teacher evaluate the performance of the students.

This cloud computing driven E-learning system exploits many advanced computer technologies including object-oriented databases, PHP, Javascript and Ajax. In order to avoid the IT maintenance cost, we choose Google App Engine (GAE) to build and deploy our system at remote Google site. GAE allows fast development and deployment with simple system administration. We do not need to maintain any hardware, power facility or make backups and the system availability is high. More importantly, the cloud system can automatically scale up with more hardware and network resources to meet the increasing user demands in an effortless manner.

In comparison with the commercial blackboard system, our system conveys the learning and assessment processes through rich-content multimedia with strict user management and access control to protect sensitive materials in addition to grading and communication functionalities. On the other hand, the on-demand resource allocation strategy at the remote Google sites can dynamically schedule resources to meet the current system needs which is very cost effective.

ACKNOWLEDGMENTS

The authors would like to thank the Illinois State Board of Education for sponsoring this project under the title of Partnership for Improved Science Achievement through Computational Science funded in 2010 (http://www.piascs.siuc.edu). We also need to thank our undergraduate and graduate students for their hard work on this project.

Reviewer's Index

A
Carolina Lagos Aguirre
Junia Coutinho Anacleto
Danilo Avola

B
Arvind K. Bansal
Dorothea Blostein
Jeff Browne

C
Pasquina Campanella
Tolga Capin
Alfonso F. Cardenas
Augusto Celentano
Maiga Chang
Ing-Ray Chen
Shu-Ching Chen
William Cheng-Chung Chu
Yuan-Sun Chu
Gennaro Costagliola
Philip Cox
Alfredo Cuzzocrea

D
Sergiu Dascalu
Andrea De Lucia
Aidan Delaney
Vincenzo Deufemia
Pierpaolo Di Bitonto
Sergio Di Martino
Tiansi Dong
David H. C. Du

E
Larbi Esmahi

F
Filomena Ferrucci
Daniela Fogli
Manuel J. Fonseca
Farshad Fotouhi
Rita Francese
Vittorio Fuccella
Kaori Fujinami

G
Ombretta Gaggi
Nikolaos Gkalelis
Angela Guercio

H
Niklas Hallberg
Robert Heller

I
Pedro Isaias

J
Joaquim A Jorge
Erland Jungert

K
Levent Burak Kara
Yau-Hwang Kuo

L
Maria Laterza
Robert Laurini
Lian Li
Fuhua Lin
Hong Lin
Alan Liu
Jonathan Liu

M
Paolo Maresca
Andrea Molinari

N
Paolo Nesi
Kia Ng

O
Ian Oliver

P

Sethuraman Panchanathan
Ignazio Passero
Joseph J. Pfeiffer
Antonio Piccinno
Fabio Pittarello
Beryl Plimmer
Giuseppe Polese

R

Michele Risi
Veronica Rossano

S

Giuseppe Scanniello
Tevf k Metin Sezgin
Buntarou Shizuki
Peter Stanchev
Gem Stapleton
Mahbub Syed
Mahbubur Syed

T

Genny Tortora

V

Stefano Valtolina
Giuliana Vitiello

W

Ellen Walker

Y

Kazuo Yana
Chengjiu Yin
Atsuo Yoshitaka

Z

Ing Tomas Zeman
Kang Zhang
Roberto Zicari

Author's Index

A
Moataz Ahmed, 124
José C. Amante, 48
Bruno Apolloni, 155

B
Annalisa Banzi, 75
C. Becchi, 60
Andrew Blake, 185
Paolo Buono, 54

C
Daniele Cenni, 113
Shi-Kuo Chang, 9, 17, 25, 29, 60
Liz Cheek, 185
Wen-Hui Chen, 29
Dana Codreanu, 107
Francesco Colace, 17, 60
Luigi Colazzo, 150
Gennaro Costagliola, 191
Maria Francesca Costabile, 54, 79

D
Mattia De Rosa, 191
Massimo De Santo, 17, 60
Aidan Delaney, 181
Thomas M. Deserno, 25
Pierpaolo Di Bitonto, 139
Tiansi Dong, 3

E
Nathan Eloe, 67, 71
Francesco Epifania, 161

F
Felipe Furtado, 130
Salvatore Ferrandino, 17
Diego R. Fiori, 89
Raffaella Folgieri, 75, 155
Manuel J. Fonseca, 48
Elda Frezza, 139
Vittorio Fuccella, 191
Borko Furht, xv

G
Itana M. S. Gimenes, 83, 89
Ingo Glöckner, 3
Angela Guercio, 144

H
Daniel Hanspeter, 95
Robert Heller, 101
John Howse, 185

I

J
Andrea Janes, 95
Edson Jordão, 130

K
Amr Kamel, 119
Ian Kottman, 71
Liqun Kuang, 9

L
Rosa Lanzilotti, 79
Young Lee, 39
Jennifer Leopold, 67, 71, 177
Kuo-Jen Liao, 39
Fuhua Lin, 101
Robyn Littleton, 71

M
José C. Maldonado, 83, 89
Ana-Maria Manzat , 107
Paolo, Maresca, 144
Adaleigh Martin, 71
Grant McClure, 101
Andrea Molinari, 150
Juan Mora, 39
Yuichi Murata, 42

N
Paolo Napoletano, 60
Paolo Nesi, 113
Jon Nicholson, 181

O
Edson A. Oliveira Junior, 83, 89

P
Michela Paolucci, 113
Kang-Lin Peng, 25
Elvira Piccinno, 139

R
Peter Rodgers, 185
Teresa Roselli, 139
Veronica Rossano, 139

S
Chaman Sabharwal, 67, 71
Florence Sedes, 107
Khaled Mohamed Shams, 119
Buntarou Shizuki, 42
Alberto Sillitti, 95
Lidia Stanganelli, 144
Gem Stapleton, 181, 185
Giancarlo Succi, 95
Yaser Sulaiman, 124
Yao Sun, 9

T
Jiro Tanaka, 42
Fabrizio Torre, 191
Yu-Chee Tseng, 25

V
Giorgio Valle, 155
Sandeep Virwaney, 101

W
Yingze Wang, 9
William Webb, 39
Michael Wisely, 177
Jan Wolter, 171

Y
Nuri Yilmazer, 39
Zhaozheng Yin, 67

Poster/Demo Presenter's Index

H
Daniel Hanspeter, A-3

J
Andrea Janes, A-3

M
Frackson Mumba, A-5

S
Alberto Sillitti, A-3
Giancarlo Succi, A-3

Z
Mengxia (Michelle) Zhu, A-5

DMS
2012

Proceedings of the Eighteenth International Conference on Distributed Multimedia Systems

Copyright © 2012
Printed by
Knowledge Systems Institute
3420 Main Street
Skokie, Illinois 60076
(847) 679-3135
office@ksi.edu | www.ksi.edu
Printed in USA, 2012
ISBN 1-891706-32-2 (Paper)