International Journal of Interactive Multimedia and Artificial Intelligence

March 2013, Vol 2, Number 1, ISSN: 1989-1660

Instead of triying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the childs?. If this were the subjected to an appropriate course of education one would obtain the adult brain

Alan Touring

SPECIAL ISSUE ON **ARTIFICIAL INTELLIGENCE AND SOCIAL APPLICATION**

http://www.ijimai.org

www.elasticbox.com
info@elasticbox.com
blog.elasticbox.com
Twitter: elasticbox

elasticbox

Deploy applications, not servers.

Challenges In Cloud Computing

Automation

- How long to deploy an application?
- What version do I use?
- How do I upgrade applications?

Portability

- How do I change providers?
- What is being used?
- How much does it cost?

Auto-Scaling

- Can my application auto-scale?
- How do I configure auto-scaling?

Disaster Recovery Planning

- Can my application tolerare faults?
- How do I recover my system?

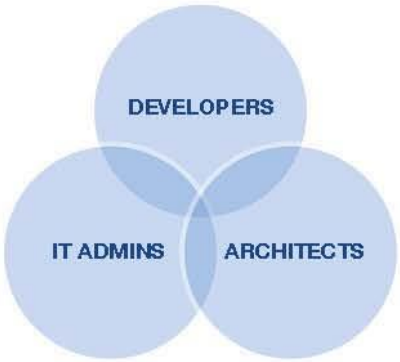
ElasticBox Solution

Automation

- ✔ Automatic Deployments
- ✔ Automatic Configuration
- ✔ Disaster Recovery

Runtime Environment

- ✔ Application Scaling
- ✔ Fault Tolerance
- ✔ Resource Clean-up
- ✔ Replication



Framework Design

- ✔ Architecture Policies
- ✔ Versioning
- ✔ Platform Management

Infrastructure Control

- ✔ Cost Analysis
- ✔ Policy Management
- ✔ Traceability

amazon web services™ · Windows Azure · hp · rackspace · openstack · EUCALYPTUS

INTERNATIONAL JOURNAL OF ARTIFICIAL INTELLIGENCE AND INTERACTIVE MULTIMEDIA

ISSN: 1989-1660 - VOL. II, NUMBER 1

Editor's Note

The International Journal of Interactive Multimedia and Artificial Intelligence provides an interdisciplinary forum in which scientists and professionals can share their research results and report new advances on Artificial Intelligence and Interactive Multimedia techniques.

This special issue "Artificial Intelligence and Social Application" includes extended versions of selected papers from Artificial Intelligence and Education area of the 13th edition of the Ibero-American Conference on Artificial Intelligence, held in Cartagena de Indias - Colombia, November, 2012. The issue includes, thus, five selected papers, describing innovative research work, on Artificial Intelligence in Education area including, among others: Recommender Systems, Learning Objects, Intelligent Tutoring Systems, Multi-Agent Systems, Virtual Learning Environments, Case-based reasoning and Classifiers Algorithms. This issue also includes six papers in the Interactive Multimedia and Artificial Intelligence areas, dealing with subjects such as User Experience, E-Learning, Communication Tools, Multi-Agent Systems, Grid Computing. IBERAMIA 2012 was the 13th edition of the Ibero-American Conference on Artificial Intelligence, a leading symposium where the Ibero-American AI community comes together to share research results and experiences with researchers in Artificial Intelligence from all over the world. The papers were organized in topical sections on knowledge representation and reasoning, information and knowledge processing, knowledge discovery and data mining, machine learning, bio-inspired computing, fuzzy systems, modelling and simulation, ambient intelligence, multi-agent systems, human-computer interaction, natural language processing, computer vision and robotics, planning and scheduling, AI in education, and knowledge engineering and applications.

The Ibero-American Society of Artificial Intelligence (IBERAMIA) is a legally-constituted non-profit association, with the primary objective of promoting scientific and technological activities related to Artificial Intelligence in Ibero-American countries. IBERAMIA joins the Ibero-American associations of Artificial Intelligence, strengthening common bonds, promoting activities and projects related to teaching, research, technology transfer and innovation related with Artificial Intelligence.

The Artificial Intelligence is present in our everyday life. Its application in distributed environments, such as the Internet, electronic commerce, mobile communications, wireless devices, distributed computing, and so on is increasing and is becoming an element of high added value and economic potential, both industrial and research. These technologies are changing constantly as a result of the large research and technical effort being undertaken in both universities and businesses. Interactive Multimedia Applications are also benefiting from advances in distributed systems research. Combining AI, applications become more customized, optimized and promote a better user experience. The exchange of ideas between scientists and technicians from both academic and business areas is essential to facilitate the development of systems that meet the demands of today's society.

We would like to thank all the contributing authors, as well as the members of the Program Committee and the Organizing Committee for their hard and highly valuable work. Their work has helped to contribute to the success of IBERAMIA conference. Finally, the Guest Editors wish to thank Editors-in-Chief of International Journal of Interactive Multimedia and Artificial Intelligence for the publication of this special issue that notably contributes to improve the quality of the conference. We hope the reader will share our joy and find this special issue very useful.

Dra. Elisa Boff, Dr. Juan Pavón

WELCOME TO NEW MEMBERS

Francisco Chiclana Ph. D. is Professor of Computational Intelligence and Decision Making, De Montfort University, Leicester, UK (May 2012). Co-Director of DIGITS - De Montfort University Interdisciplinary Group in Intelligent Transport Systems, Faculty of Technology, De Montfort University, Leicester, UK (2011-present). Coordinator of REF UOA 11: Computer Science and Informatics. Member of the Faculty of Technology Scholarship Panel, De Montfort University, Leicester, UK (2008-present). Member of the School of Computer Science and Informatics Research Committee. Member of the School of Computer Science and Informatics Industrial Liaison Committee

Arrabales is an ICT expert and PhD in Artificial Intelligence. He has published over 40 papers on Machine Consciousness and Cognitive Robotics and he is co-researcher in several AI research projects. He has a strong background in management, R&D and innovation projects and digital entertainment market. His main research interest is the application of artificial cognitive systems in video games and serious games. Arrabales has a working experience of over 12 years both in public university and privately held companies. He has worked as project manager and IT consultant in IPTV and VoD industry for several technology companies. Arrabales holds a PhD and B.Sc. in Computer Science and Technology from Carlos III University of Madrid and a B.Eng. in Computer Engineering from Polytechnic University of Madrid. He also holds an MBA from the EOI Business School. Currently, Arrabales is the Digital Content Engineering Area Manager at U-Tad.

DOI: 10.9781/ijimai.2013.210

TABLE OF CONTENTS

OPEN ACCESS JOURNAL

ISSN: 1989-1660

COPYRIGHT NOTICE

BROA: An agent-based model to recommend relevant Learning Objects from Repository Federations adapted to learner profile

[1,2]Paula A. Rodríguez, [1,2]Valentina Tabares, [2]Néstor D. Duque, [1]Demetrio A. Ovalle, [3]Rosa M. Vicari
[1]*Departamento de Ciencias de la Computación y la Decisión*, Universidad Nacional de Colombia-Medellín, Colombia
[2]*Departamento de Informática y Computación*, Universidad Nacional de Colombia-Manizales, Colombia
[3]*Instituto de Informática*, Universidade Federal do Rio Grande do Sul, Brasil

***Abstract* — Learning Objects (LOs) are distinguished from traditional educational resources for their easy and quickly availability through Web-based repositories, from which they are accessed through their metadata. In addition, having a user profile allows an educational recommender system to help the learner to find the most relevant LOs based on their needs and preferences. The aim of this paper is to propose an agent-based model so-called BROA to recommend relevant LOs recovered from Repository Federations as well as LOs adapted to learner profile. The model proposed uses both role and service models of GAIA methodology, and the analysis models of the MAS-CommonKADS methodology. A prototype was built based on this model and validated to obtain some assessing results that are finally presented.**

***Keywords* — Artificial Intelligent in Education, GAIA, Learning objects repository federations, MAS–CommonKADS, Multi-agent Systems, Student-centered recommender systems**

I. INTRODUCTION

THE growth of digital information, high-speed computing, and ubiquitous networks has allowed for accessing to more information and thousands of educational re-sources. This fact has led to the design of new teaching-learning proposals, to share educational materials, and also to navigate through them [1]. Learning Objects (LOs) are distinguished from traditional educational resources for their easy and quickly availability through Web-based repository, from which they are accessed through their metadata. In order to maximize the number of LOs to which a student could have access, to support his/her teaching-learning process, digital repositories have been linked through centralized repository federations sharing in this way educational resources and accessing resources from others [2]. LOs must be tagged with metadata so that they can be located and used for educational purposes in Web-based environments [3]. Recommender systems are widely used online in order to assist users to find relevant information [4]. Having a user profile allows a recommender system to help the student to find the most relevant LOs based on the student's needs and preferences. Intelligent agents are entities that have sufficient autonomy and intelligence to be able to handle specific tasks with little or no human supervision [5]. These agents are currently being used almost as much as traditional systems, making it a good choice to solve problems where autonomous systems are required and thus they work not only individually but also cooperate with other systems to achieve a common goal. The aim of this paper is to propose a model for LO searching, retrieving, recommendation, and evaluator modeled through the paradigm of multi-agent systems from repository federations. For doing so, the searching process needs a query string that is entered by the user and a similar relevance user profile according to the student's learning style (LS). The LO searching process is performed using local and remote repositories, or repository federations, that are accessible via web along with LO descriptive metadata. Since LO Repositories (LORs) are distributed, are different in design and structure, and not handle the same metadata standards. There is also a coordinator to be responsible for directing the search to different repositories according to their characteristics. The recommendation is made through collaborative filtering, searching for a similar profile to the user who is doing the quest to deliver a user pair LOs evaluated positively.

The rest of the paper is organized as follows: Section 2 outlines main concepts involved in this research. Section 3 describes some related works to the model pro-posed. Section 4 introduces the multi-agent model proposal based on both role and service models of GAIA methodology, and the analysis models of the MAS-CommonKADS methodology. A validation of the system's operation can be visualized is shown in Section 5. Finally, conclusions and future work are presented in Section 6.

II. BASIC CONCEPTS

A. Learning objects, repositories and federations

According to the IEEE, a LO can be defined as a digital

DOI: 10.9781/ijimai.2013.211

entity involving educational design characteristics. Each LO can be used, reused or referenced during computer-supported learning processes, aiming at generating knowledge and competences based on student's needs. LOs have functional requirements such as accessibility, reuse, and interoperability [6][7]. The concept of LO requires understanding of how people learn, since this issue directly affects the LO design in each of its three dimensions: pedagogical, didactic, and technological [7]. In addition, LOs have metadata that describe and identify the educational resources involved and facilitate their searching and retrieval. LORs, composed of thousands of LOs, can be defined as specialized digital libraries storing several types of resources heterogeneous, are currently being used in various e-learning environments and belong mainly to education-al institutions [8]. Federation of LORs serve to provide educational applications of uniform administration in order to search, retrieve, and access specific LO contents available in whatever of LOR groups [9].

B. Recommender Systems

Recommender systems are aimed to provide users with search results close to their needs, making predictions of their preferences and delivering those items that could be closer than expected [10],[11]. In the context of LOs these systems seeks to make recommendations according to the student's characteristics and its learning needs. In order to improve recommendations, recommender systems must perform feedback processes and implement mechanisms that enable them to obtain a large amount of information about users and how they use the LOs [2],[12].

C. Multi-Agent Systems

Agents are entities that have autonomy in order to perform tasks by achieving their objectives without human supervision. The desirable characteristics of the agents are as follows [13]: ***Reactivity***: they respond promptly to perceived changes in their environment; ***Proactivity***: agents can take initiative; ***Cooperation and Coordination:*** they perform tasks communicating with other agents through a common language; ***Autonomy***: agents do not require direct intervention of humans to operate; ***Deliberation***: they perform reasoning processes to make decisions, ***Distribution of Tasks***: each agent has definite limits and identified the problems to be solved; ***Mobility***: they can move from one machine to another over a network; ***Adaptation:*** depending on changes in their environment they can improve their performance, and ***Parallelism***: agents can improve performance depending on changes in their environment.

Multi-agent Systems (MAS) are composed of a set of agents that operate and interact in an environment to solve a specific and complex problem. This paradigm provides a new way of analysis, design, and implementation of complex software systems and has been used for the development of recommender systems [14].

D. Student Profile

The student profile stores information about the learner, its characteristics and preferences, which can be used to obtain search results according to its specificity. To handle a user profile can be used to support a student or a teacher in the LO selection according to its personal characteristics and preferences [14]. Gonzalez et al. [15] include in the student profile contextual characteristics that can be seen as transient values that are associated with environmental changes during one student's learning system session along with different physical and technological variables. Duque (2009) presents a combination of VARK y FSLSM models with good results to characterize the students profile and thus, provide students with learning materials tailored to their specific learning styles [16].

III. Related Works

Morales et al. (2007) present an architecture based on the multi-agent paradigm to identify and retrieve relevant LOs using the information request supplied by the user. In addition, this proposal includes aspects related to quality of LOs which are specified within their semantic description to improve the LO selection [17]. Authors propose a multi-agent architecture to retrieve LOs, however they do not use student cognitive characteristics such as learning styles in order to make recommendations.

Gerling (2009) proposes an intelligent system to assist a user in finding appropriate LO, according to the search's subject by using the user profile which takes into account its characteristics, preferences, and LO *relative importance*. The recommender system incorporates an intelligent agent in order to retrieve educational resources on the Web, considering student's learning style [18]. However, the system's design only considers the utilization of one intelligent agent.

Duque (2009), in his doctoral thesis, proposes a multi-agent system for adaptive course generation. The system is composed of several intelligent agents: an agent for the student profile information; a domain agent, having the structure of the virtual course and the teaching material (TM); a HTN planner agent, and finally, a TM recovery agent which makes the process of TM search and retrieval [16]. This work focuses on creating customized virtual courses taking into account student's learning styles; however, it does not focus on LOs.

Prieta (2010) proposes a multi-agent architecture for the process of search and retrieval of LO in distributed repositories. An additional functionality is provided to the system is making the LORs statistics on the number of results found and response time, then make future consultations in the LORs rated [19]. This system offers neither recommendations to the user nor customized searches based on the LO metadata. Casali (2011) presents an architecture and implementation of a recommender system prototype based on intelligent agents, whose goal is to return an ordered list of the most appropriate LOs according to the parameters that characterizes the user profile, language preferences and the interaction degree that the user wants to have with the LO. The search is performed in repositories having descriptive LO metadata which involves

educational characteristics [14]. The main limitation of this research is that although some student characteristics are considered into the user profile the user learning styles were not taken into account.

IV. Model Proposed

BROA (Spanish acronym for Learning Object Search, Retrieval & Recommender System) is a multi-agent system for searching, retrieving, recommendation and evaluator of LO, according to a search string entered by the user. The LOs resulting from the search are recommended based on the student's style of learning and other users' assessments. The Web-based LO search is performed over local and remote repositories, or by using LO repository federations through metadata descriptive LOs. Considering that LORs are distributed, they are different in design and structure, and hence they do not handle the same metadata standards. BROA was built under the MAS approach in order to exploit their advantages as follows: the *Parallelism of Tasks* for simultaneously searching in both local and remote LOR; the *Deliberation* ability for making decisions on which of LORs must perform the search and for performing user recommendations; *Cooperation*, *Coordination* and *Distribution* of tasks among agents by clearly identifying the problems to be solved by each agent and to define its limits. In our model each agent knows how LOs are stored and how each LO can be searched, accessed, and retrieved.

A. *Development Methodology*

There are different kinds of methodologies for modeling MAS, such as GAIA characterized for analyzing and designing agent-oriented systems. The main key concepts of GAIA are the following: roles, which are associated with responsibilities, permissions, activities, and protocols [5]. Another well-known MAS design methodology is MAS-CommonKADS proposed by Iglesias in his doctoral thesis [20] which integrates knowledge and software engineering along with object-oriented protocols. An integration of both methodologies was used in order to model the BROA system. The following is a brief description of each of these models: *Role Model* (GAIA): Allows the system designer to identify the expected functions of each of the entities that composes the system (goals, responsibilities, capabilities, and permissions). *Service Model* (GAIA): This model identifies all the services associated with each of the roles, its inputs, outputs, pre-conditions, and post-conditions. *Agent Model*: Describes the characteristics of each agent, specifying name, kind of agent, role, description, skills, services, activities, and goals. According to the GAIA methodology an agent can play several roles as shown in Figure 1, thus, a changing role diagram must be used in this case. *Task Model:* This model describes all the tasks that agents can perform along with the objectives of each task, its decomposition, and troubleshooting methods to solve each objective. Figure 2 shows the BROA's task diagram. *Expertise Model:* Describes the ontologies (knowledge and its relationships) that agents need to achieve their objectives. *Communication Model:* Describes main interactions among humans and software agents along with human factors involved for the development of these interfaces. *Organization Model*: This model aims to describe the human organization in which the multi-agent system is involved along with the software agent organization structure. *Coordination Model*: Dynamic relationships among software agents are expressesed through this model. For doing so, all the conversations among agents must be described: interactions, protocols, and capabilities required. Figure 3 shows the BROA's sequence diagram that specifies main interactions among agents.

B. *BROA's Architecture*

The design phase of MAS-CommonKADS methodology takes as input all the models got from the analysis phase and transforms their specifications for implementation. In addition, the architecture of each agent and the global network architecture must be provided [20].

Figure 4 shows the multi-agent architecture of the model proposed. This architecture was used to develop the BROA system, implemented using JADE (Java Agent Development Framework) agents [21]. The next Section describes each of the agents of the BROA system along with the main interactions that exist among them.

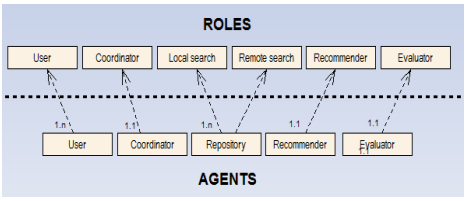


Fig. 1. Transformation of roles in agent's diagram

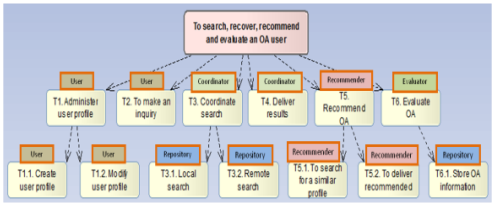


Fig. 2. Task Diagram

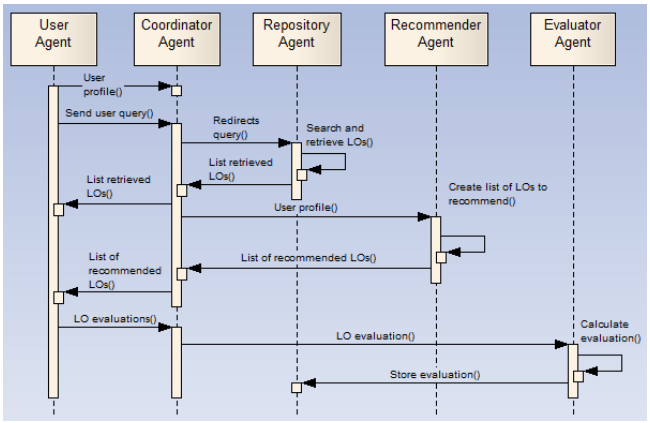


Fig. 3. Sequence diagram

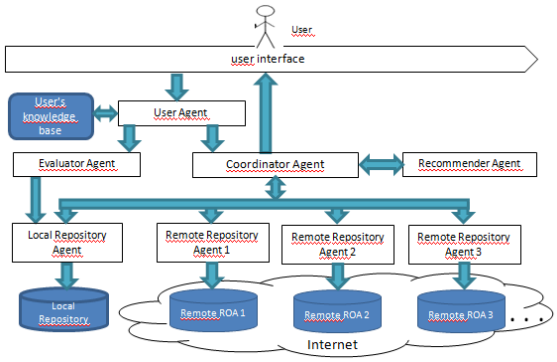


Fig. 4. BROA/SRLO's Architecture

Fig. 5. BROA/SRLO's Web-based Interface

Agent Description

User Agent: This agent communicates directly with the human user and whose role is representing him within the system along with communications with other agents (Coordinator and Evaluator). Also, the user agent manages the user's profile, enabling the creation and modification of profile's characteristics and preferences. Finally, this agent sends the query string to the coordinator agent in order to perform the LO search and the evaluation of the recommended LOs. *Coordinator Agent:* This agent is of deliberative nature since it takes care of redirectioning queries that the user makes to both the local and the remote repositories. This agent knows the repositories associated with the system and the information that each of them manages. In addition, it can access the user agent's profile to know what characteristics are useful for making a recommendation (learning style, educational level, language preference, among others). *Local and Remote Repository Agent:* Repository agents are responsible for making accurate LO searches in both local and remote repositories. This agent recognizes how LOs are stored inside the repositories, under what standard and type of metadata that manages. Also knows the type of search that can be performed within the repository and how to recover a particular LO stored. The local repository agent is also responsible for storing the LO evaluation given from an evaluator agent. Similarly, in the proposed architecture there is a repository for each LOR agent federation is local or remote. *Recommender Agent:* This agent makes two recommendations; the first stage is to find users registered in the system with similar profile to the user so having the same learning style and education level. The LOs selected by those users having a score greater or equal than 4 are shown. The second stage of recommendation is based on LOs recovered in all different repositories, based on user's LO query. This recommendation is based on the student's learning style. It is important to highlight that in the model proposed, the recommendation is based on the metadata that describes the LO and the information of learning style, educational level, and language preference of the registered student. In order to represent the agent's knowledge production rules were used, such as the following rule:

$$\begin{aligned} &\textit{LearningStyle(Visual-Global)} \wedge \\ &\textit{LearningResourceType (figure)} \vee \\ &\textit{LearningResourceType(graph)} \vee \\ &\textit{LearningResourceType(slide)} \vee \\ &\textit{LearningResourceType(table)} \wedge \\ &\textit{InteractivityLevel(medium)} \vee \\ &\textit{InteractivityLevel(high).} \end{aligned}$$

When there is a failure of similar users, the system shows only the results of the second recommendation and then stored the user profile information and evaluated LOs, within the knowledge base. *Evaluator agent:* This agent manages the evaluation performed by a user to some of the LOs that have been explored. The evaluation is made through explicit qualification that is given by the selected student who rates the specific LO from 1 to 5 according to his/her own satisfaction level.

Platform Design

The BROA's agent architecture was developed in JAVA, using JDOM for handling XML user's profiles. The local repository manager is stored in the PostgreSQL database that is characterized to be stable, with high performance, and great flexibility. The agent creation and management is made by using JADE platform using FIPA-ACL performatives [21]. The ontology creation was performed by using Protégé and finally, the Web integration was made based on the ZK framework. Figure 5 shows BROA's Web interface with the recommended and retrieved LOs. For the LO search process there was a student who had a Visual-sequential learning style and the search string used was: "computer science". Thus, a total of 196 LOs were recovered and only recommended, after a "learning style" filtering just 45 of them.

V. Experiments and Results

BROA system provides to the human user LO lists by using its interface. The first list is the result of the search made by the user according to his criteria. The second list presents list of recommended items to the user, which correspond to those LOs that are the most adapted to his own learning style.

To validate the BROA system a test was performed based on the keyword "computer science". In addition, a comparison was made with the results given by the system concerning the LOs recommended for users with different learning styles proposed by Duque [16]: Visual-Global, Visual-sequential, Auditory-Global, Auditory-sequential, Kinesthetic-Global, Kinesthetic-sequential, Reader-Global, Reader-sequential.

Thus, virtual users with different profiles were generated and LOs from real repositories were recovered. Figure 6 shows the results got for the tests being performed. A total of 196 LOs were retrieved after the search process for students with different learning styles. Figure 5 shows the quantity of LOs for each of the different learning styles. The BROA system makes a good recommendation process since the LOs provided are well adapted to the student's learning profile.

In order to evaluate the results of recommendations given by the system the Precision measure [22] was used which purpose is to analyze the quality of the retrieval.

$$Precision = \frac{Relevant\ LOs}{Relevant\ LOs + Retrieved\ LOs} \quad (1)$$

Figure 6 shows the results obtained by applying the Precision measure formula (1) to each learning style and additionally comparing values obtained with and without recommendation when accessing LO Merlot repository.

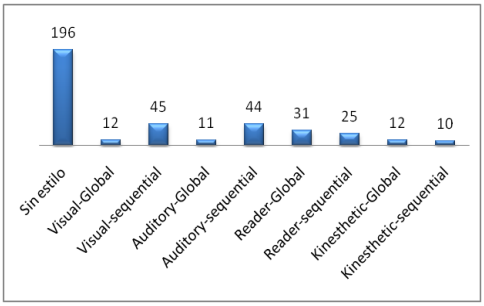


Fig. 5. LOs retrieved after search

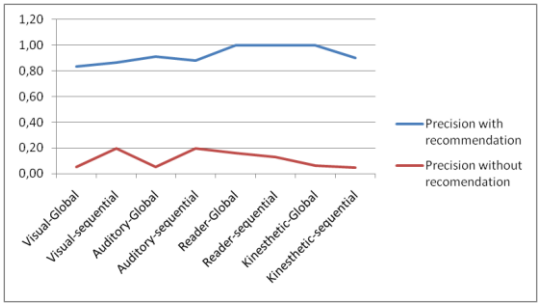


Fig. 6. Precision

VI. Conclusion and Future Work

This paper proposes a model for learning object searching, retrieving, recommendation, and evaluation modeled through the paradigm of MAS from repository federations. The BROA (Spanish acronym for Learning Object Search, Retrieval & Recommender System) system was built using this model. There is an agent in the BROA system dedicated to each repository accessed by the system. In order to facilitate LO searches, the agent knows how the LOs are stored, how is the way of accessing and recovering them, under what standard and type of metadata the LOs are stored and handled by the specific repository. Those searches are performed in a local LOR, where the already evaluated LOs are stored, and also performed in remote LORs associated to the system. The BROA system offers two types of recommendation; the first one is based on finding similar profiles. This first recommendation approach has not already been implemented in this prototype. The second type of recommendation is by searching the metadata of the LO, taking into account the query performed by the user, the results are presented at the right side on Figure 5. The model proposed in this paper addressed issues such as working on LOs and learning styles and making recommendations by the system to the user based on customized searches using the LO metadata. In addition, the problem modeling using a MAS technique was an excellent option, which allowed the disintegration into functional blocks, without losing the systemic point of view, which leads to distributing the solution in diverse entities that require specific knowledge, processing and communication between each other. The MAS allowed a neutral vision in the model proposed.

It is envisaged as future work to add an interface agent to make context-aware adaptations, along with the list of LOs delivered by the system considering other issues such as type of device from where the query is made, bandwidth, among others. For the evaluation process, it is intended to make templates for the user to rate its opinion about recommended LOs (explicit evaluation). The agent should analyze the results of the explicit evaluation and use logs, to assign a rating to each LO. Also it is envisaged to improve the theoretical and practical basis of the first stage of recommendations made by the system through collaborative filtering techniques. The learning style for this prototype should be selected by the user, an additional future work aims to propose a learning style test that will define which kind of learning style the user who is logged in the system has.

Acknowledgment

The research reported in this paper was funded in part by the COLCIENCIAS project entitled "ROAC Creación de un modelo para la Federación de OA en Colombia que permita su integración a confederaciones internacionales" Universidad Nacional de Colombia, with code 1119-521-29361.

References

[1] C. I. Peña, J.-L. Marzo, J. L. De la Rosa, and R. Fabregat, "Un Sistema de Tutoría Inteligente Adaptativo Considerando Estilos de Aprendizaje," VI Congreso Iberoamericano de Informática Educativa, pp. 1 – 12, 2002.

[2] J. Z. Li, "Quality, Evaluation and Recommendation for Learning Object," International Conference on Educational and Information Technology (ICEIT 2010), pp. 533–537, 2010.

[3] A. B. Gil and F. García, "Un Sistema Multiagente de Recuperación de Objetos de Aprendizaje con Atributos de Contexto," ZOCO'07/CAEPIA, 2007.

[4] K. Niemann, M. Scheffel, M. Friedrich, U. Kirschenmann, H.-C. Schmitz, and M. Wolpers, "Usage-based Object Similarity," Journal of Universal Computer Science, vol. 16, no. 16, pp. 2272–2290, 2010.

[5] M. Wooldridge, N. R. Jennings, and D. Kinny, "The Gaia Methodology for Agent-Oriented Analysis and Design," Autonomous Agents and Multi-Agent Systems, vol. 3, pp. 285 – 312, 2000.

[6] Y. Ouyang and M. Zhu, "eLORM: Learning Object Relationship Mining-based Repository," Online Information Review, vol. 32, no. 2, pp. 254–265, 2008.
[7] D. Betancur, J. Moreno, and D. Ovalle, "Modelo para la Recomendación y Recuperación de Objetos de Aprendizaje en Entornos Virtuales de Enseñanza/Aprendizaje," Avances en Sistemas e Informática, vol. 6, no. 1, pp. 45–56, 2009.
[8] F. De Prieta and A. B. Gil, "A Multi-agent System that Searches for Learning Objects in Heterogeneous Repositories," Advances in Intelligent and Soft Computing, vol. 71, pp. 355–362, 2010.
[9] H. Van de Sompel, R. Chute, and P. Hochstenbach, "The aDORe Federation Architecture: Digital Repositories at Scale," International Journal on Digital Libraries, vol. 9, no. 2, pp. 83–100, 2008.
[10] F. Chesani, "Recommendation Systems," Corso di laurea in Ingegneria Informatica, pp. 1–32, 2007.
[11] K. Mizhquero and J. Barrera, "Análisis, Diseño e Implementación de un Sistema Adaptivo de Recomendación de Información Basado en Mashups," Revista Tecnológica ESPOL-RTE, 2009.
[12] O. Sanjuán, E. Torres, H. Castán, R. Gonzalez, C. Pelayo, and L. Rodriguez, "Viabilidad de la Aplicación de Sistemas de Recomendación a entornos de e-learning," 2009.
[13] N. R. Jennings, "On Agent-based Software Engineering," Artificial Intelligence, vol. 117, no. 2, pp. 277–296, Mar. 2000.
[14] A. Casali, V. Gerling, C. Deco, and C. Bender, "Sistema Inteligente para la Recomendación de Objetos de Aprendizaje," Revista Generación Digital, vol. 9, no. 1, pp. 88–95, 2011.
[15] H. González, N. Duque Méndez, and D. Ovalle C., "Técnicas Inteligentes para la Actualización Dinámica del Perfil del Usuario en un Sistema de Educación Virtual," in Tendencias en Ingeniería de Software e Inteligencia Artificial, 2009.
[16] N. Duque, "Modelo Adaptativo Multi-Agente para la Planificación y Ejecución de Cursos Virtuales Personalizados - Tesis Doctoral," Universidad Nacional de Colombia, 2009.
[17] E. Morales and A. Gil, "Arquitectura para la Recuperación de Objetos de Aprendizaje de Calidad en Repositorios Distribuidos," SCHA: Sistemas Hipermedia Colaborativos y Adaptativos. II Congreso Español de Informática CEDI 2007, vol. 1, no. 1, pp. 31 – 38, 2007.
[18] V. B. Gerling, "Un Sistema Inteligente para Asistir la Búsqueda Personalizada de Objetos de Aprendizaje," Universidad Nacional de Rosario, 2009.
[19] F. de la Prieta, A. G. González, J. M. Corchado, and E. Sanz, "Sistema Multiagente Orientado a la Búsqueda, Recuperación y Filtrado de Objetos Digitales Educativos," VIII Jornadas de Aplicaciones y Transferencia Tecnológica de la Inteligencia Artificial, pp. 65 – 74, 2010.
[20] C. Á. Iglesias Fernández, "Definición de una Metodología para el Desarrollo de Sistemas Multiagentes," Universidad Politécnica de Madrid, 1998.
[21] F. Bellifemine, A. Poggi, and G. Rimassa, "JADE – A FIPA-compliant Agent Framework," Proceedings of PAAM, 1999.
[22] G. Shani and A. Gunawardana, "Evaluating Recommendation Systems," Recommender Systems Handbook, pp. 257 – 297, 2011.

Paula A. Rodriguez is Manager of Information Systems from Universidad Nacional de Colombia, Manizales and currently she advances Masters Studies in Systems Engineering program from Universidad Nacional de Colombia. She participates in the Research Group in Adaptive Intelligent Environments GAIA. Her research interests are in the areas of Artificial Intelligence in Education and Multi-Agent Systems (MAS), Learning Object Repository Federations, among others.

Valentina Tabares is Manager of Information Systems from Universidad Nacional de Colombia, Manizales and currently she advances Masters Studies in Systems Engineering program from Universidad Nacional de Colombia. She participates in the Research Group in Adaptive Intelligent Environments GAIA. Her research interests are in the areas of Artificial Intelligence in Education and Multi-Agent Systems (MAS), Learning Object Repository Federations, among others.

Néstor D. Duque is Associate Professor from Universidad Nacional de Colombia, campus Manizales and head of the Research Group in Adaptive Intelligent Environments GAIA. He received his master's degree in Systems Engineering from Universidad Nacional de Colombia, and his PhD in Engineering from Universidad Nacional de Colombia, campus Medellin. His PhD thesis was entitled "*Adaptive Multi-Agent Model for Planning and Execution of Adaptive Virtual Courses*" with Cum Laude honors. Author of a number of articles in scientific journals and book chapters including topics on their research and academic work, speaker at major national and international events; head in the development process of national and international research projects, member of academic committees of a dozen national and international journals, academic review in post-graduate academic programs and special events. He has also received some meritorious distinction for researching and teaching from School of Administration at Universidad Nacional de Colombia, campus Manizales.

Demetrio A. Ovalle is Full Professor at Universidad Nacional de Colombia, campus Medellin. He was head of the Department of Computer Science at School of Mines (2006-2008). He is presently head of GIDIA (Artificial Intelligence Research & Development Group), categorized as A1, maximum category established by Colombian National Science & Technology Bureau. He has published over 150 refereed publications in Journals, Books and International Conferences. He received his Master's and PhD degrees in Informatics respectively from Institut National Polythechnique in 1987, and Université Jospeh Fourier in 1991, in Grenoble, France. His research interests are in the areas of Distributed Artificial Intelligence and Multi-Agent Systems (MAS), Virtual Learning Environments, Intelligent Tutoring Systems, Mobile & Ubiquitous Computing, Ambient Intelligence, Autonomic Systems and Wireless Sensor Network. He has also received some meritorious distinctions in research from School of Mines at Universidad Nacional de Colombia, campus Medellin.

Rosa M Vicari is Full Professor at Federal University of Rio Grande do Sul (UFRGS), Brazil. She received her PhD in Computer Science at University of Coimbra, Portugal. Her research interests are in the areas of Artificial Intelligence in Education and Multi-Agent Systems (MAS), Virtual Learning Environments, Intelligent Tutoring Systems, Semantic Web and Ontologies, Learning Object Repository Federations, among others. She has also received some meritorious distinctions in research from UFRGS.

Improving the family orientation process in Cuban Special Schools trough Nearest Prototype classification

Villuendas-Rey, Y.[1], Rey-Benguría[2], C., Caballero-Mota, Y.[3], García-Lorenzo, M. M.[4]
[1]*Computer Science Department, University of Ciego de Ávila, Ciego de Ávila, Cuba*
[2]*Educational Research Center, University "Manuel Ascunce", Ciego de Ávila, Cuba*
[3]*Computer Science Department, University of Camagüey, Camagüey, Cuba*
[4]*Computer Science Department, University of Las Villas, Villa Clara, Cuba*

***Abstract* — Cuban Schools for children with Affective – Behavioral Maladies (SABM) have as goal to accomplish a major change in children behavior, to insert them effectively into society. One of the key elements in this objective is to give an adequate orientation to the children's families; due to the family is one of the most important educational contexts in which the children will develop their personality. The family orientation process in SABM involves clustering and classification of mixed type data with non-symmetric similarity functions. To improve this process, this paper includes some novel characteristics in clustering and prototype selection. The proposed approach uses a hierarchical clustering based on compact sets, making it suitable for dealing with non-symmetric similarity functions, as well as with mixed and incomplete data. The proposal obtains very good results on the SABM data, and over repository databases.**

***Keywords* — special schools, nearest prototype classifiers, mixed data, non-symmetric similarities**

I. INTRODUCTION

In Cuba, the Ministry of Education has special educational schools for dealing with children with singular educational needs. Among them, there are Schools for children with Affective-Behavioral Maladies (SABM).SABM had been designed with the goal of offering a special educational context. In them, the needs of the children that had show maladies in their affective development and/or in their behavior are resolved. Therefore, the children that had have delinquent or anti-social behaviors are bewared in a personalized way in SABMs. The family is the basic cell of society, and in it is the closest educational context for children. When children get out of SABMs, they return to their homes and to their neighborhoods, where they often do not have the correct models to follow. The adequate orientation to the children's family plays a key role to correct the deficiencies, and to insert effectively these children into society. That is why the personnel in charge of the family orientation process in the SABM of the province of Ciego de Ávila characterize the familiar dynamics of each family, and then proceed to design a personalized strategy for each group of families with similar dynamics.

To give an adequate orientation to the families, the headings of the SABM proceed on two stages: Clustering and Classification. On stage 1, they cluster the families according to their characteristics, and on stage 2, they assign a new arrived family to the group of its closest family, using Nearest Prototype Classification (see figure 1).

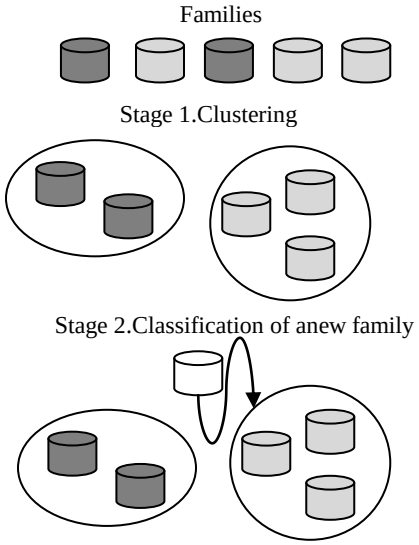


Fig.1. Stages of the Family orientation process at SABM.

Despite the challenges attached to clustering data, there is a need of structuralizing data in SABM School. In this domain, the description of each family has mixed and incomplete attributes. The sociologists associated to SABM selected these attributes to characterize the family dynamics of the SABM families. The data of the families of the SABM School of Ciego de Avila has fourteen attributes (Table I). These attributes measure the attitude of the family to the inclusion of a child in the SABM School, as well as the peculiarities of the family dynamic.

DOI: 10.9781/ijimai.2013.212

TABLE I
DESCRIPTION OF THE ATTRIBUTES OF SABM DATA

Att.	Name	Description
1.	impact	If exists impact or shock in the family
2.	attitude	The attitude adopted about the inclusion of a child in the SABM
3.	change	How the family reacts to the change, if they oppose (O), they resist (R), they have resignation (G) or they agree (A)
4.	guilty	If there are or there are not guilty feelings in the family
5.	clime	The kind of emotional clime, if it is positive or negative
6.	communication	The kind of communication that prevails in the family
7.	handling	The way the family handles the fact of including a child into the SABM
8.	relations	The way the interpersonal relations are developed into the family
9.	crisis	The kind of emotional crisis, by demoralization, disarranging, frustration, impotence or no crisis
10.	estimation	The way the self estimation of the family is
11.	consciousness	If there is or not consciousness of the reality
12.	linkage	If there is or not a favorable link with the SABM
13.	hopes	The hopes the family has to the future
14.	time	The time (in months) the child is at the SABM

To compare in effective way two families, and to decide whether the families have similar dynamics, it was needed to work together with the family orientation experts and the sociologists associated to SABM in Ciego de Ávila. After analyzing several similarity functions proposed in the literature for dealing with mixed and incomplete data, the experts decided that those similarities were not adequate for comparing SABM data.

It was decided then to design a personalized similarity function to deal with the peculiarities of SABM data.The sociologists and the family orientation experts of SABM decide that classical comparison criteria for nominal attributes were adequate to compare the nominal features of SABM data, but the 3rdattribute, "change".

To compare the values of the 3rd attribute, it was needed to establish a non-symmetric comparison matrix as feature comparison criterion, due to the semantics of the different values of this attribute (table 2). For the numerical attribute, "time", the selected comparison criterion was normalized difference.

From analysis with different expert and sociologist associated to SABM, a similarity function to compare the families is designed. It is a non-symmetric similarity, due to the non-symmetric comparison matrix for the 3rd attribute, change. Let be two families, f_i and f_j, and $f_i[k]$ the value of the k-th attribute (A_k) in the f_i family. The similarity for comparing SABM data is defined by:

$$S(f_i, f_j) = 1 - D(f_i, f_j) \quad (1)$$

where $D(f_i, f_j) = \sum_{k=1}^{14} D_k(f_i, f_j)$. For nominal attributes but 3rd attribute, the function $D_k(f_i, f_j)$ is as follows:

$$D_k(f_i, f_j) = \begin{cases} 0 & if \quad f_i[k] = f_j[k] \\ 1 & if \quad f_i[k] \neq f_j[k] \\ 0.5 & if \quad f_i[k] = "?" \vee f_j[k] = "?" \end{cases} \quad (2)$$

On the other hand, for the numerical attribute, the function $D_k(f_i, f_j)$ is as follows:

$$D_k(f_i, f_j) = \frac{|f_i[k] - f_j[k]|}{\max(A_k) - \min(A_k)} \quad (3)$$

In the case of the third attribute, "change", the different attribute values have a peculiar meaning. Due to, their similarity depends of each value combination. This attribute defines the attitude the family adopts to face the fact that one of the family members, a child, will be allocate into the SABM.

Table II shows the comparison matrix of values for the attribute "change". As shown, the dissimilarity between values "Resistance" (R) and "Resignation" (G) differ from "Resignation" to "Resistance".

TABLE II
COMPARISON MATRIX OF THE VALUES FOR THE ATTRIBUTE "CHANGE"

Value	*O*	*R*	*G*	*A*
O	0	0.2	0.8	1
R	0.2	0	**0.4**	0.8
G	0.8	**0.8**	0	0.4
A	1	0.8	0.4	0

Each cell shows the dissimilarity values of the pair (row vs. column). In bold the non-symmetric values

The rest of the paper is as follows: section II introduces the proposed hierarchical clustering, based on Compact Sets structuralizations, and the proposed Nearest Prototype selection algorithm. Section III addresses the selection of the adequate cluster number for the families in SABM, to improve the family orientation process. Sections IV and V review some previous works on clustering mixed data and nearest prototype selection for mixed data, respectively. Section VI offers the numerical experiments comparing the proposals with respect other clustering and prototype selection algorithms, over SABM data and repository data. The paper ends with the conclusions and future works.

II. CLUSTERING AND NEAREST PROTOTYPE SELECTION BASED ON COMPACT SETS

A. Hierarchical clustering based on Compact Sets

Taking into consideration the nature of the problem of clustering and classifying SAMB data, described by mixed and incomplete features, and with a non-symmetric similarity function used to compare the families; it is necessary to develop a novel clustering algorithm able to deal with all these restrictions simultaneously. This section introduces a

hierarchical clustering algorithm based on Compact Sets, to deal with SABM data.

Compact Sets structuralization was described in [1], and this structuralization is based on the concept of Maximum Similarity Graphs. Maximum Similarity Graphs (MSG), are directed graphs such that each instance $x \in X$ is connected to its most similar instance. A connected component of a MSG is a Compact Set (CS).

Formally, let be $G = (X, \theta)$ a MSG for a set of objects X, with arcs θ. In this graph, two objects $x_i, x_j \in X$ form an arc $(x_i, x_j) \in \theta$ if $\max_{x \in X}\{sim(x_i, x)\} = sim(x_i, x_j)$, where $sim(x_i, x_j)$ is a similarity function. Usually $sim(x_i, x_j) = 1 - \Delta(x_i, x_j)$ and $\Delta(x_i, x_j)$ is a dissimilarity function. In case of ties, the Maximum Similarity Graph establishes a connection between the object and each of its nearest neighbors. As mentioned before, Compact Sets are the connected components of such graph.

Formally, a subset $N \neq \emptyset$ of X is a Compact Set if and only if [1]:

$$a) \forall x_j \in X \left[x_i \in N \wedge \left(\max_{\substack{x_i \in X \\ x_i \neq o_j}} \{sim(x_i, x_j)\} = sim(x_i, x_j) \right) \vee \left(\max_{\substack{x_i \in X \\ x_i \neq x_j}} \{sim(x_j, x_i)\} = sim(x_j, x_i) \right) \right] \Rightarrow x_j \in N$$

$$x_i = x_{i_1} \vee j$$

$$b) \forall x_i, x_j \in N, \exists x_{i_1}, \cdots, x_{i_q} \in N \quad \max$$

$c)$ Every isolated object is a Compact Set, degenerated.

All the instances connected between them belong to the same CS, such that the nearest neighbor of each instance is also in the same CS (figure 2). The proposed method follows a hierarchical agglomerative approach to clustering, but merging CSs instead of objects.

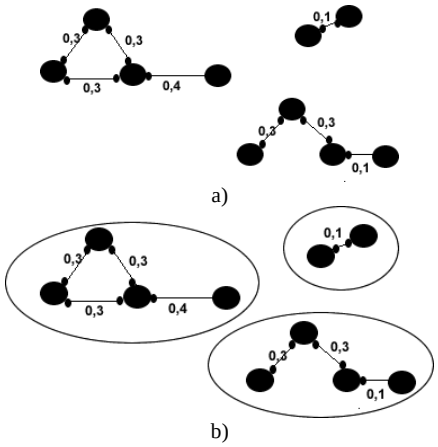


Fig. 2. a) Maximum Similarity Graph of instances and b) Compact Sets of instances.

As many other hierarchical agglomerative clustering algorithms [2], the proposed Compact Set Clustering (CSC) uses a multilayer clustering to produce the hierarchy. The algorithm (figure 3) starts by computing the Maximum Similarity Graph from dataset. Second, it is defined as initial groups each CS in the MSG. Then it merges the groups, until having the desired number of clusters. The merging is making with all possible groups that are more similar in a single step and it is avoided order dependence.

CSC algorithm uses the similarity between cluster representatives as inter group similarity function. Let be x and y the representatives of clusters Ci and Cj, respectively, and $S(x, y)$ is the similarity between those representatives. The similarity function between those clusters is:

$$sim(C_i, C_i) = S(x, y) \tag{5}$$

The instances that maximize the overall inter-group similarity correspond to the representatives of the clusters. Formally, the representative instance r of a group Cj will be:

$$r = arg \max_x \{S(x, y)\} \tag{6}$$

(4)

Compact Sets Clustering (CSC)	
Inputs:	k: number of groups S: inter objects similarity function *T*: training set
Output:	C: resulted clustering

1. C = ϕ
2. Create a Maximum Similarity graph of the objects in *T* using the similarity function S
3. Add to C each connected component of the graph created at step 1
   3.1. Select the cluster representative instance as in (6)
4. While |C| < k
   4.1. Merge all more similar groups, using (5)
   4.2. Recalculate cluster representative instance
5. Return C

Fig. 3. Compact Set Clustering (CSC) algorithm.

Thus, the CSC algorithm selects real objects to represent the clusters, avoiding the construction of artificial centroids. This approach obtains compact and separated clusters, and it is able to detect the true partitions of data.

B. Nearest Prototype selection based on Compact Sets

In the classification phase of the SABM data, each new family must be compared to every family already in SABM, using the Nearest Neighbor (NN) classifier [3]. Despite the NN classifier is one of the most popular supervised

classification methods in Pattern Recognition, it suffers from important drawbacks. NN has high storage and computational requirements, because it storages the entire training set, requiring large space. In addition, to determine the class of a new object, NN needs to compare it with every object in the training set. Another drawback of NN is its sensitivity to noisy and outlier objects.

To overcome these drawbacks, researchers have proposed the Nearest Prototype (NP) classification. NP classification use prototype selection methods to obtain a reduced set of representative objects (prototypes) as training data for classification.As NP classification has been extensively used for supervised classification with very good results [4] , [5], it was decided that the classification stage of SABM data was carried out using NP classification.

As stated before, the SAMB data is described by mixed and incomplete features, and it also uses a non-symmetric similarity function to compare the families; so, it is necessary to develop a novel prototype selection algorithm able to deal with all these restrictions simultaneously. This section introduces a prototype selection algorithm (figure 4) based on Compact Sets structuralization [6].

The proposed Prototype Selection (PS) algorithm allows deciding the desired amount of prototypes for the Nearest Prototype classification. It is also able to deal with arbitrarily similarity functions; due to the similarity to compare objects is a parameter of the algorithm.

Prototype Selection algorithm	
Inputs:	k: desired number of prototypes S: inter objects similarity function *T*: training set
Output:	P: prototype set

1. P = ϕ
2. C = CSC(k, S, T)
3. For each cluster $C_i \in C$
   5.1. Select the cluster representative as in (6)
   3.1. Add to P the cluster representative
4. Return P

Fig. 4. Prototype Selection (PS) algorithm.

The PS algorithm starts with an empty prototype set. Then, it structuralizes the training set *T* using the Compact Sets Clustering (CSC) method, finding as many clusters as desired prototypes. Then, the PS algorithm will select the representing object of each cluster, and will add it to the prototype set.

The PS algorithm proposed includes several novel characteristics, differentiating it from previous prototype selection algorithms. It structuralizes data using a hierarchical clustering algorithm based on Compact Set structuralization. It also uses a data-dependant similarity function, which makes it applicable to several domains with non-metric similarities, such as social sciences and medicine. It also selects representing objects of clusters as prototypes instead of constructing artificial objects for the Nearest Prototype classification stage.

III. Finding the Adequate Clustering for SABM Data

As mentioned before,the data of the families of the SABM School of Ciego de Avila is described by mixed attributes that measure the attitude of the family to the inclusion of a child in the SABM, as well as the peculiarities of the family dynamic. It is also used a non-symmetric similarity (1) to compare family descriptions.

The first stage of the family orientation process is to cluster the families of the SABM. As no predefined number of clusters exists, it is needed to obtain several candidates clustering, and then select the one that best fits data. Internal cluster validity indexes allow comparing several candidate clustering, and deciding which of them best fits data. To determine the adequate cluster number of SABM data, it was clustered with cluster number varying from two to nine clusters, and then it were used internal cluster validity indexesto select the partition that best fits data. Among unsupervised cluster validity indexes, the Dunn's index measure how compact and well separated the clusters are. Let be d(Ci,Cj) the dissimilarity between clusters, and Δ(Ci) the cluster size, the Dunn's validation index is the ratio between the minimum dissimilarity between two clusters and the size of the largest cluster.

$$D = \frac{\min_{i=1..nc, j=1..nc, i\neq j}(d(C_i, C_j))}{\max_{1..nc}(\Delta(C_i))} \quad (7)$$

Where $d(C_i, C_j)$ is the dissimilarity between clusters, and $\Delta(C_i)$ is the cluster size.

Dunn's index was used with complete – linkage as dissimilarity measure and with single – linkage as cluster size measure. In figure 5, there are shown the results the Dunn's index with cluster number varying from two to nine clusters. The best partition has seven clusters.

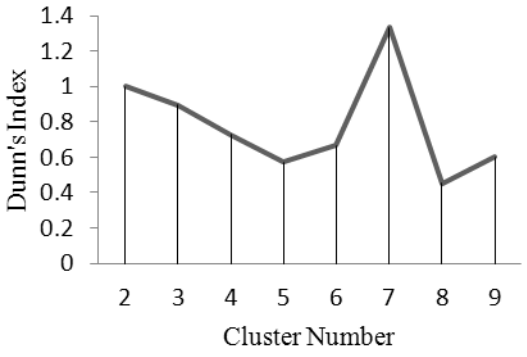


Fig.5. Values of the Dunn's index obtained by CSC using different cluster number.

In addition, it was also used the Silhouette index [7]. The Silhouette is the average, over all clusters, of the Silhouette width of their points.

If x is an object in the cluster c_i and n_i is the number of objects in c_i, then the Silhouette width of x is defined by the

ratio:

$$S(x) = \frac{b(x) - a(x)}{\max\{a(x), b(x)\}} \quad (8)$$

where a(x) is the average dissimilarity between x and all other objects in c_i, and b(x) is the minimum of the average dissimilarities between x and the objects in the other clusters.

$$a(x) = \frac{1}{n_i - 1}\sum_{y \in C_i, y \neq x} d(x,y) \quad (9)$$

$$b(x) = \min_{h=1..k, h \neq i}\left\{\frac{1}{n_h}\sum_{y \in C_h} d(x,y)\right\} \quad (10)$$

Finally, the global Silhouette is as follows:

$$S(C) = \frac{1}{k}\sum_{i=1}^{k}\frac{1}{n_i}\sum_{x \in C_i} S(x) \quad (11)$$

For a given object x, its Silhouette width ranges from −1 to 1. If the value is close to –1, then it means that the object is more similar, on average, to another cluster than the one to which it belongs. If the value is close to 1, then it means that its average dissimilarity to its own cluster is significantly smaller than to any other cluster. The higher the Silhouette, the more compact and separated are the clusters.

In figure 6 it is shown the results of the Silhouette index, with cluster number varying from two to nine clusters. The best partition also had seven clusters.

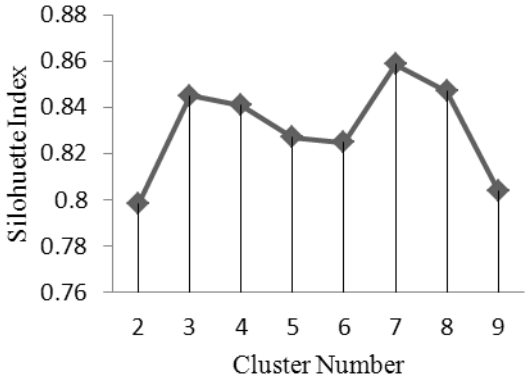


Fig.6. Values of the Silhouette index obtained by CSC using different cluster number.

According to both Dunn's and Silhouette indexes, the structuralization that best fits the SABM data is the one with seven clusters. This structuralization will be used later in the classification stage of the family orientation process.

For the classification stage, each instance had as class label the number of the cluster it belongs. By this, the resulted clustered families of stage one, will constitute the training matrix for the supervised classifier.

IV. Previous Works on Clustering Mixed Data

It is impossible to address clustering techniques without referring to the k-means algorithm. The k-means algorithm is one of the oldest clustering techniques, and it has a proved efficiency to find compact and well separated clusters. At the first step, k-means randomly select a set of cluster centers from data. Then, it assigns each object to its closest center, using the Euclidean distance. After that, the algorithm iterates until no change is made on cluster centers. In the iterative process, it computes the new cluster centers, as the mean of all objects in the cluster, and reassigns every object to its closest center.

Several authors have proposed modifications to this simple, yet powerful technique, to handle mixed and incomplete data. All of them include a redefinition of the distance function, as well as the cluster centers.

In 1997, Huang proposed the k-prototypes (KP) algorithm [8]. The KP algorithm redefines cluster center as the mean of the numerical attributes, and the mode of the nominal attributes. Also, it uses as dissimilarity function, with weights $\omega=\{\omega_1, \ldots, \omega_d\}$ of each attribute. Although the KP algorithm deals with mixed type attributes, it does not handle missing data.

In 2007, Ahmad and Dey proposed another modification of the classical k-means algorithm [9]. They redefined the dissimilarity function. The proposed dissimilarity includes attribute weights. For categorical attributes, the dissimilarity takes into account the co-occurrences of each value pair, and then set as more similar the low frequency values pairs.

Ahmad and Dey [9] also redefine the cluster center. In their definition, the center consists on a cluster description. The description includes the mean of each numerical attribute, and a set of pairs (value, count) for each categorical attribute. Each pair has the attribute value and the count of objects in a cluster that have this value.

In 2011, the same authors [10] proposed a modification of the algorithm proposed previously in 2007. They do not give in the paper any name for the new method, so this paper refers to it as AD2011 (Ahmad and Dey proposal of 2011). The new method discretizes numeric attributes before the clustering process, using the Equal Width Discretization procedure. It also includes in the dissimilarity function the contribution λ of each attribute to the cluster.

The AD2011 algorithm includes two user-defined parameters. The first is the γ parameter, included in the attribute contribution computation, having a suggested value of 20, and the second is the S parameter, included in the discretization procedure of numeric attributes, having a suggested value of 5.

Among the main drawbacks of k-meansbased clustering are that the algorithms depend on the definition of cluster centers. They are also unable to form arbitrary shapes clusters.

Another family of clustering algorithms is the family of hierarchical algorithms. Hierarchical clustering algorithms create a decomposition of the objects by forming a binary tree called dendogram. All objects are at the root, at the intermediate nodes are groups of objects, and at leafs are single objects. The tree is usually created top down (divisive algorithms) or bottom up (agglomerative algorithms). In the last, each object is considered as a group, and at each step the two more similar groups are joined. The stopping condition is usual that all objects are in the same group, or the desired number of groups is reached. These methods are referred as Hierarchical Agglomerative Clustering (HAC). A HAC method for dealing with mixed and incomplete data is the HIMIC algorithm [11].

Other kind of clustering algorithms are model – based clustering. In these methods, a model or metaheuristic is used to evolve clusters. Each candidate clustering is a solution, having certain optimization value (cluster quality). The model or heuristic iterates, until it finds the desired clustering. Among model based clustering are the Genetic Algorithm cluster based AGKA, proposed by Roy and Sharma in 2010 [12] and the Flocking based method proposed in [13].

The AGKA algorithm is based on Genetic algorithms. Genetic Algorithms (GA) are one of the most used techniques in Artificial Intelligence, Pattern Recognition, and Data Mining. They offer a feasible solution for a huge number of optimization and classification problems.

The AKGA method uses a genetic procedure, and includes the dissimilarity proposed by Ahmad and Dey in 2007 [9] in the fitness function.

AGKA codifies each candidate clustering as an individual, using an integer array of length equal to the object count. Each position (gene) of the array indicates the cluster assigned to the object in that position. It has a mutation strategy that changes an object to its most probable cluster, offering a quickly convergence. It also has an elitist survival strategy.

Model based clustering can be applied in a huge number of situations, and they have numerous variants according to the parameters, evolution strategies, solution generation, and others. However, the algorithms belonging to this approach are often stochastic, and the quality of the resulted clustering depends on the parameter setting and internal evolution strategies used by a particular model.

V. PREVIOUS WORKS ON MIXED DATA PROTOTYPE SELECTION

As one of the main drawbacks of NN classifiers is its sensitivity to noisy and mislabeled objects (section II), there is a research interest in the Artificial Intelligence and Pattern Recognition community to overcome this difficulty [14], [15].

The algorithms to obtain a prototype set for the NN classifier are divided into prototype selection methods and prototype generation methods. This work is focused on prototype selection methods; due to these methods obtain a subset of the training matrix.

Prototype selection methods are divided into condensing algorithms, editing algorithm and hybrid algorithms [15]. Condensing algorithms aim at reducing the NN computational cost by obtaining a small subset of the training matrix, maintaining the accuracy as high as possible, while editing algorithms aim at improve classifier accuracy by deleting noisy and mislabeled objects. Hybrid methods usually combine both condensing and editing strategies in the selection procedure.

The first editing algorithm is the Edited Nearest Neighbor (ENN), proposed by Wilson in 1972 [15]. The ENN algorithm deletes the objects misclassified by a k-NN classifier, where k is a user-defined parameter, usually $k = 3$.

Another classical editing method is MULTIEDIT, proposed by Devijver and Kittler in 1980 [17]. MULTIEDIT works as follows: first, it divides the training matrix in ns partitions, in each partition it applies the ENN method, using a 1-NN classifier trained with the next partition. The last partition is trained with the first one. After each iteration, it joins the remaining objects in each partition and repeats the process until no change is achieved in successive iterations.

In 2000, Hattori and Takahashi [18] proposed a new editing method, referred in this paper as NENN. The method computes the k neighbors of each object, including all objects that have the same dissimilarity value of the last k neighbor. If at least one of the neighbors it is not of the same class of the object, it deletes the object of the training matrix.

In 2002, Toussaint used proximity graphs to obtain a reduced prototype set [19].

Caballero *et al*. introduced other editing algorithms in 2007, the EditRS1 and EditRS2 methods [20]. They used elements of the Rough Set Theory to obtain lower and upper approximations of the training matrix, and to compute the limit regions of each class. Both methods use a reduct as base of the editing process.

Condensing methods were proposed first by Hart in 1968 with the Condensed Nearest Neighbor (CNN) algorithm [21]. In this work, he introduced the concept of consistent subset, a subset of the training matrix such as training a NN classifier with this subset, every instance in the original training matrix is correctly classified.

The Reduced Nearest Neighbor (RNN) consists on a post processing of the CNN algorithm. After computing CNN, RNN tries to delete every object, if the deletion does not introduce any inconsistency. Gates [22] demonstrated that if a minimum consistent subset is a subset of the CNN result, the RNN methods always find it.

Another modification to classic CNN is the Generalized Condensed Nearest Neighbor (GCNN) method. It was proposed by Chou *et al*. in 2006 [23]. The GCNN treats CNN as a particular case, and includes a set of rules to "absorb" prototypes.

Other condensation method is the PSR, introduced by Olvera-López at al. in [24], which selects the prototype set based on prototype relevance. More recently, García-Borroto *et al*. proposed the CSESupport method [25]. It deletes the less important objects, guaranteeing the consistency of the subset

by a mark strategy.

The mark strategy consists on the following: when deleting an object, it marks every object that supports it (in a Support Graph), and at least one of them must be included in the condensed subset. A support graph is a directed graph, such as it connects each object all objects of its same class closer than the NUN object [25].

The NUN (Nearest Unlike Neighbor) is the object of different class closest to *x* [26].In this strategy, when an object is the last with a mark, it is included in the result, same if an object does not have any outward edges in the graph.

The method initiates with all training matrix as a consistent subset, and at each iteration deletes the less important objects. It also updates the objects NUN, and builds the support graph with every object in the training matrix, to maintain the subset consistency [25].

CSESupport method handles missing and incomplete data, as well as asymmetric and non-symmetric dissimilarities. However, it does not allow defining the desired number of prototypes.

VI. Numerical Experiments

Numerical experiments were carried out using nine mixed and incomplete databases of the Machine Learning repository of the University of California at Irvine (UCI) [27].

TABLE III
DESCRIPTION OF DATABASES USED IN NUMERICAL EXPERIMENTS

Databases	Nominal Attributes	Numerical Attributes	Classes
autos	10	16	6
colic	15	7	2
dermatology	1	33	6
heart-c	7	6	5
hepatitis	13	6	2
labor	6	8	2
lymph	15	3	4
tae	2	3	3

The first experiment was to compare the performance of state of the art clustering algorithms with CSC, over the SABM data and over repository data, and the second experiment was to compare the performance of the proposed prototype selection procedure with respect to other prototype.

A. Numerical experiments on clustering mixed data

The family orientation process on SABM involves both clustering and Nearest Prototype classification. It was decided to consider both internal and external cluster validity indexes to compare the performance of the proposed CSC algorithm with respect to AD2011 [10] and AGKA [12] algorithms over the SABM data. Both AD2011 and AGKA had a predefined dissimilarity, and the CSC algorithm used the similarity function designed for SABM data (section I). It were selected both Dunn's index and the Silhouette index for internal clustering validation and Entropy and Cluster Error indexes for external clustering validation. The amount of clusters to obtain by each algorithm in SABM data was defined to be equal to seven. It was because seven clusters was the best partition of SABM data (section III). The results of the compared algorithms over the SABM data are shown in Table IV.

TABLE IV
RESULTS OF THE CLUSTERING ALGORITHMS OVER SABM DATA

Algorithms	Internal indexes		External indexes	
	Dunn's index	*Silhouette index*	*Cluster Error*	*Entropy*
AD2011	0.0077	-0.1427	0.6364	2.5331
AGKA	0.1968	-0.2850	0.5686	1.7730
CSC	**1.3333**	**0.8585**	**0**	**0**

External evaluation measures for clustering can be applied when class labels for each data object in some evaluation set can be determined *a priori*. The clustering task is then used to assign these data points to any number of clusters. In each cluster must be all and only those data objects that are members of the same class [28].To compare the clustering results produced by the different algorithms, it is used the Cluster Error andthe Entropy measure.

Cluster Error [9] consists on counting the amount of objects not belonging to the majority class of each cluster. Let be C the resulted clustering, C_ia cluster in C, and n_i the number of object belonging to the majority class in the i-th cluster. The Cluster Error of C with respect to class labels is given by:

$$CE(C) = \sum_i \frac{|C_i| - n_i}{|C_i|} \tag{12}$$

Lower values of Cluster Error indicate a high performance of the algorithms.

The Entropy index, as described in [29], measures the dispersion of the classes in the clusters. Low Entropy indicates high similarity of clusters and classes. Let be C the resulted clustering, ci the i-th cluster in C, n_i^j the number of object of the j-th class in the i-th cluster and N the amount of objects. The Entropy of C with respect to class labels is given by:

$$E(C) = -\sum_i \frac{|c_i|}{N} * \sum_j \frac{n_i^j}{|c_i|} \log\left(\frac{n_i^j}{|c_i|}\right) \tag{13}$$

To compare the results of the selected clustering algorithms with respect to the proposed CSC over repository data, both Cluster Error and Entropy external indexes were selected. It was used as cluster count for each algorithm the amount of class each database has.

The CSC algorithm was applied to repository data using the HOEM dissimilarity function proposed by Wilson and Martinez [30]. The results of Cluster Error and Entropy over repository data are shown in table V and figure 7, and in table VI and figure 8, respectively. Then, to establish if the differences in performance were significant or not, the Wilcoxon test was applied.

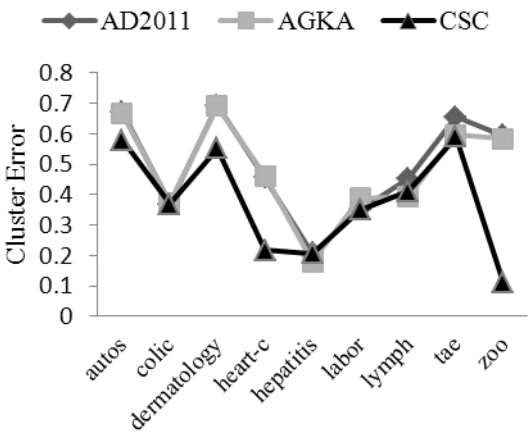


Fig.7. Results of the methods over UCI databases according to Cluster Error.

TABLE V
RESULTS OF THE METHODS ACCORDING TO CLUSTER ERROR

Databases	**AD2011**	**AGKA**	**CSC**
autos	0.6731	0.6650	0.**5804**
colic	**0.3695**	0.3724	**0.3695**
dermatology	0.6939	0.6910	**0.5519**
heart-c	0.4554	0.4615	**0.2178**
hepatitis	0.2064	**0.1803**	0.2064
labor	**0.3508**	0.3880	**0.3508**
lymph	0.4527	**0.3933**	0.4121
tae	0.6556	0.6158	**0.5894**
zoo	0.5940	0.5841	**0.1089**
Times Best	**2**	**2**	**7**

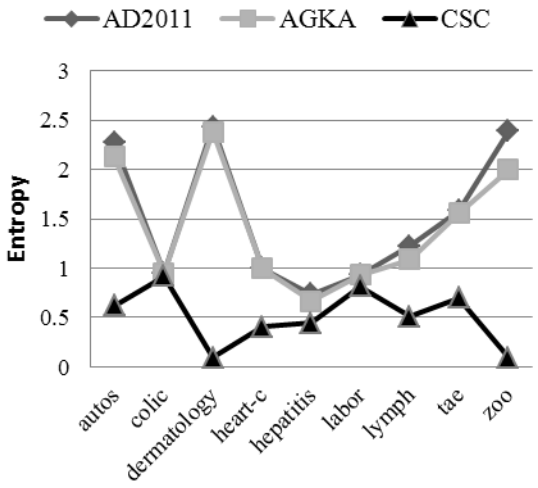


Fig.8. Results of the Entropy of the methods over the UCI databases.

TABLE VI
RESULTS OF THE METHODS ACCORDING TO ENTROPY

Databases	**AD2011**	**AGKA**	**CSC**
autos	2.2725	2.1314	**0.6198**
colic	0.9503	0.9525	**0.9205**
dermatology	2.4326	2.3793	**0.0947**
heart-c	0.9943	0.9956	**0.4063**
hepatitis	0.7346	0.6663	**0.4424**
labor	0.9348	0.9311	**0.8155**
lymph	1.2277	1.0914	**0.5120**
tae	1.5845	1.5593	**0.6985**
zoo	2.3906	1.9988	**0.0909**
Times Best	**0**	**0**	**9**

The Wilcoxon test (table VII) helps determining if the CSC significantly outperforms the other algorithms according to Cluster Error and Entropy. It is define the null hypothesis as no differences in performance, and the alternative hypothesis as the proposed method outperforms the other method. It is used an alpha value of 0.05, with a 95% confidence level.

TABLE VII
RESULTS OF THE WILCOXON TEST FOR PAIR WISE CLUSTERING COMPARISON

Our method	**Asymptotical Significance**	
	According to Cluster Error	*According to Entropy*
vs. AD2011	**0.028**	**0.008**
vs. AGKA	**0.036**	**0.008**

The proposed method has a significant better performance than the AD2011 and AGKA methods. This may be due to it uses a similarity function data dependant, which makes it applicable to several domains with non-metric similarities, such as social sciences and medicine. It also selects a cluster representative instead of constructing fictional cluster centers, guaranteeing a real object represents each cluster. Therefore, the proposed algorithm is able to detect the true partitions of data and to handle mixed and incomplete databases.

B. Numerical experiments on Nearest Prototype classification

This section offers the results of comparing the performance of the proposed Prototype Selection (PS) approach with some other prototype selection algorithms for mixed data [18], [23], [24], [25] and with the original classifier (ONN), using all objects.

The proposed PS method was applied to SABM data with cluster count equal to seven (selecting one prototype per class), and it was applied over repository data with cluster count equal to 50, so 50 prototypes were selected from each database, one for each cluster. PS used the HOEM proposed by Wilson and Martinez [30] as dissimilarity function for repository data.

The Classifier Error measure was used to compare the

performance of the algorithms. Classifier Error (CE) is calculated as the ratio between the amount of misclassified objects and the amount of instances in the original training set. Let be $\alpha(x)$ the true class off the object x, and $NN(x)$ the class assigned to x by the Nearest Neighbor classifier. The Classifier Error is given by:

$$CE = \frac{|x \in T : \alpha(x) \neq NN(x)|}{|\text{Training set}|} \quad (14)$$

Another quality measure of prototype selection methods is Retention Rate. Retention Rate (RR) is calculated as the ratio between the amount of selected prototypes and the amount of instances in the original training set.

$$RR = \frac{|\text{Prototype set}|}{|\text{Training set}|} \quad (15)$$

The 10 fold cross validation procedure facilitates testing the performance of the Prototype Selection stage. On SABM data (table VIII), the classifier trained with the whole data obtained zero testing error, despite the use of a non-symmetric similarity. In addition, several prototype selection methods were able to classify correctly every instance in the testing sets, having zero error too. The PRS method deletes the entire dataset, whereas the GCNN method does not achieve any data reduction.

TABLE VIII
CLASSIFIER ERROR AND RETENTION RATES OF PROTOTYPE SELECTION METHODS OVER SABM DATA

Algorithm	Classifier Error	Retention Rates
CSESupport (CSES)	**0**	**0.1812**
GCNN	**0**	1
NENN	0.395	0.8421
PRS	-	0*
PS	**0**	**0.1812**
ONN	**0**	1

* The PRS method deletes the entire database.

The Classifier Error and Retention Rate results of the methods over repository data are shown in tables IX and X, respectively. Figures 9 and 10 also show these results.

TABLE IX
CLASSIFIER ERROR OF PROTOTYPE SELECTION METHODS OVER REPOSITORY DATA

Databases	CSES	GCNN	NENN	PRS	PS	ONN
autos	0.3026	0.3023	0.6054	0.3311	0.3450	**0.2926**
colic	0.2310	*0.1956*	**0.1819**	0.2176	0.3012	0.2064
dermatology	0.1172	0.0681	**0.0572**	0.0873	**0.0572**	0.0599
heart-c	0.2576	0.2282	**0.1621**	0.2312	*0.2213*	0.2282
hepatitis	0.2325	0.1875	0.2079	0.2325	0.1937	**0.1741**
labor	**0.1000**	0.1566	0.1700	0.2066	*0.1233*	0.1400
lymph	0.2361	0.2033	0.2433	0.2461	0.2033	**0.1823**
tae	0.3801	0.3841	0.7554	0.5691	0.5366	**0.3641**
zoo	*0.0300*	*0.0300*	0.1081	0.0490	0.0600	0.0400
Times better than ONN	**2**	**2**	**3**	**0**	**3**	

Error lower than original classifier in italics and sub-rayed, and best results in bold.

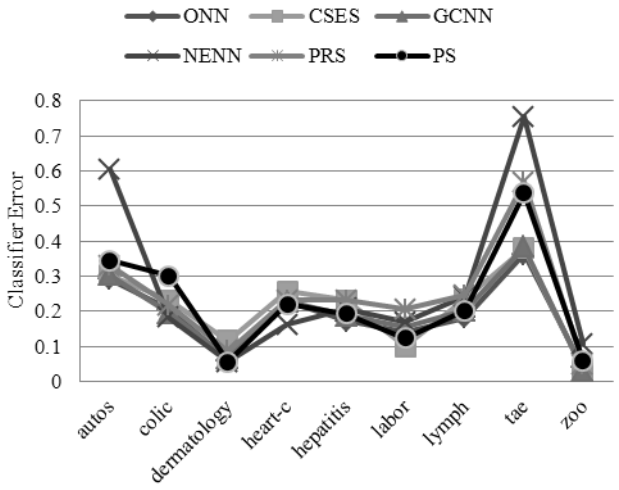


Fig.9. Results of Classifier Error of the methods over the UCI databases

The proposal was able to outperform classifier accuracy in three databases, as well as NENN, and does not have a significant increase of classifier error in the remaining databases.

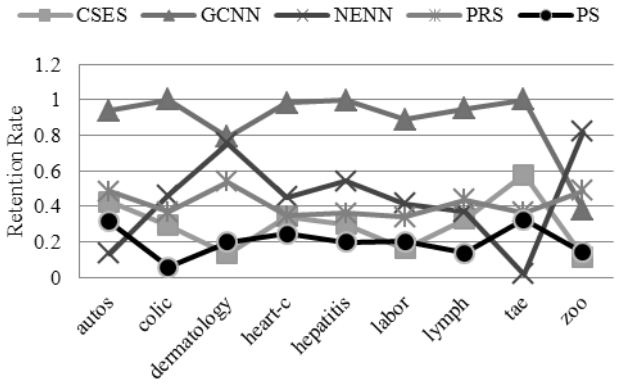


Fig.10. Results of Retention Rate of the methods over the UCI databases.

The proposal gets the lower object retention rates in four databases, and keeps it lower than 35% in the remaining. These results are due to the selected amount of prototypes, established to be 50.

TABLE X
RETENTION RATE OF PROTOTYPE SELECTION METHODS OVER REPOSITORY DATA

Databases	CSES	GCNN	NENN	PRS	PS
autos	0.4277	0.9393	**0.1328**	0.4856	0.3155
colic	0.2936	1.0000	0.4616	0.3702	**0.0590**
dermatology	**0.1345**	0.7905	0.7562	0.5395	0.1982
heart-c	0.3447	0.9817	0.4518	0.3487	**0.2472**
hepatitis	0.3011	0.9971	0.5426	0.3606	**0.2000**
labor	**0.1638**	0.8887	0.4174	0.3392	0.2046
lymph	0.3304	0.9504	0.3686	0.4369	**0.1381**
tae	0.5798	1.0000	**0.0184**	0.3642	0.3252
zoo	**0.1166**	0.3851	0.8196	0.4873	0.1430
Times Best	**3**	**0**	**2**	**0**	**4**

Best results are shown in bold.

Although the above results are very promising, again the Wilcoxon test (table XI) was used to establish the differences between the proposed approach and other algorithms, according to classifier error and object retention rates. Again, it is define the null hypothesis as no differences in performance, and the alternative hypothesis as the proposed method outperforms the other method. It is used an alpha value of 0.05, with a 95% confidence level.

TABLE XI
RESULTS OF WILCOXON TEST FOR PAIR WISE COMPARISON OF PROTOTYPE SELECTION METHODS OVER REPOSITORY DATA

Asymptotical Significance	Our method vs.				
	CSES	*GCNN*	*NENN*	*PRS*	*ONN*
Classifier Error	0.678	0.263	0.327	0.314	0.051
Retention Rate	0.051	**0.008**	0.051	**0.008**	**0.008**

According to classifier error, the proposed Prototype Selection (PS) ties with other prototype selection algorithms, and with the original classifier. In addition, this approach has a significant better performance than two other methods according to object retention rates, according to a 95% of confidence. These results reflect that the proposed method is able to maintain classifier accuracy, using only a reduced number of prototypes. In addition, the nature of the PS algorithm makes it suitable for dealing with quantitative and qualitative features, absences of information and non-symmetric dissimilarity functions.

VII. CONCLUSION

A conclusion might elaborate on the importance of the work or suggest applications and extensions. In Cuban special schools, the family orientation process has two stages: family clustering and family classification. This paper proposed a novel method for clustering and Nearest Prototype Classification. The proposed approach has its bases on hierarchical compact sets and handles mixed type data as well as non-symmetric similarity functions. It is compared the performance of the proposal with respect to existing clustering and prototype selection algorithms over repository and real Cuban special schools data. The proposal successfully clusters and classifies the families of children in Cuban special schools. This leads to a better orientation process, spending less time to correct the children deficiencies.

REFERENCES

[1] J. Ruiz-Shulcloper and M. A. Abidi, "Logical combinatorial pattern recognition: A Review," in *Recent Research Developments in Pattern Recognition*, S. G. Pandalai, Ed. USA: Transword Research Networks, 2002, pp. 133-176.

[2] A. K. Jain and R. C. Dubes, *Algorithms for clustering data*. New Jersey, USA: Prentice Hall, 1988.

[3] T. M. Cover and P. E. Hart, "Nearest Neighbor pattern classification," *IEEE T Inform Theory*, vol. 13, pp. 21-27, 1967.

[4] P. Perner, "Prototype based classification," *ApplIntell*, vol. 28, pp. 238-246, 2008.

[5] Y. Villuendas-Rey, Y. Caballero-Mota, and M. M. García-Lorenzo, "Using Rough Sets and Maximum Similarity Graphs for Nearest Prototype Classification**,"** *Lecture Notes in Comp Sci* 7441, pp. 300-307, 2012**.**

[6] Y. Villuendas-Rey, C. Rey-Benguría, Y. Caballero-Mota, and M. M. García-Lorenzo, "Nearest prototype classification of Special School families *based on hierarchical compact sets clustering," Lecture* Notes in Artif Intell 7637, pp. 662–671, 2012.

[7] M. Brun, C. Sima, J. Hua, J. Lowey, B. Carroll, E. Suh, and E. R. Dougherty, "Model-based evaluation of clustering validation measures," *Pattern Recogn*, vol. 40, pp. 807-824, 2007.

[8] Z. Huang, "Clustering large data sets with numeric and categorical values," in *Proc. 1stPacific - Asia Conference on Knowledge discovery and Data Mining*, 1997, pp. 21-34.

[9] A. Ahmad and L. Dey, "A k-means clustering algorithm for mixed numerical and categorical data," *Data KnowlEng*, vol. 63, pp. 503-527, 2007.

[10] A. Ahmad and L. Dey, "A k-means type clustering algorithm for subspace clustering of mixed numeric and categorical data," *Pattern RecognLett*, vol. 32, pp. 1062-1069, 2011.

[11] R. A. Ahmed, B. Borah, D. K. Bhattacharyya, and J. K. Kalita. (2005).*"HIMIC: A Hierarchical Mixed Type Data Clustering Algorithm,"*Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.61.6369&rep=rep1&type=pdf

[12] D. K. Roy and L. K. Sharma, "Genetic k-means clustering algorithm for mixed numeric and categorical datasets," *International Journal of Artificial Intelligence & Applications (IJAIA*), vol. 1, 2010.

[13] X. Cui, J. Gao, and T. E. Potok, "A Flocking Based Algorithm for Document Clustering Analysis," *J Syst Architect*, vol. 52, pp. 505-515, 2006.

[14] I. Triguero, J. Derrac, S. García, and F. Herrera, "Prototype generation for Nearest Neighbor classification: Taxonomy and Experimental Study," *IEEE TPatternAnal*, vol. 34, pp. 417-435, 2012.

[15] S. García, J. Derrac, J. R. Cano, and F. Herrera, "Prototype Selection for Nearest Neighbor Classification: Taxonomy and Empirical Study," *IEEE TPatternAnal*, vol. 34, pp. 417-435, 2012.

[16] D. L. Wilson, "Asymptotic properties of nearest neighbor rules using edited data," *IEEE T Syst Man Cy B*, vol. 2, pp. 408-421, 1972.

[17] P. A. Devijver and J. Kittler, "On the edited nearest neighbor rule," in *Proc. 5th International Conference on Pattern Recognition*, Los Alamitos, California, 1980, pp. 72-80.

[18] K. Hattori and M. Takahashi, "A new edited k-nearest neighbor rule in the pattern classification problem," *Pattern Recogn*, vol. 33, pp. 521-528, 2000.

[19] G. Toussaint, "Proximity Graphs for Nearest Neighbor Decision Rules: Recent Progress," in *Proc. 34 Symposium on Computing and Statistics INTERFACE-2002*, Montreal, Canada, 2002, pp. 1-20.

[20] Y. Caballero, R. Bello, Y. Salgado, and M. M. García, "A method to edit training set based on rough sets," *International Journal of Computational Intelligence Research*, vol. 3, pp. 219-229, 2007.

[21] P. E. Hart, "The condensed nearest neighbor rule," *IEEE TInform Theory*, vol. 14, pp. 515-516, 1968.

[22] G. W. Gates, "The reduced nearest neighbor rule," *IEEE T Inform Theory*, vol. IT-18, pp. 431-433, May 1972.

[23] C. H. Chou, B. A. Kuo, and F. Cheng, "The Generalized Condensed Nearest Neighbor rule as a data reduction technique," in *Proc. 18th International Conference on Pattern Recognition*, 2006, pp. 556-559.
[24] J. A. Olvera López, J. A. Carrasco-Ochoa, and J. F. Martinez Trinidad, "Prototype selection via prototype relevance," Lecture Notes in Comp Sci 5197, pp. 153-160, 2008.
[25] M. García-Borroto, Y. Villuendas-Rey, J.A. Carrasco-Ochoa, and J.F. Martínez Trinidad. "Finding Small Consistent Subset for the Nearest Neighbor Classifier Based on Support Graphs," Lecture Notes in Comp Sci 5856, pp. 465–472, 2009.
[26] B. D. Dasarathy, "Nearest unlike neighbor (NUN): an aid to decision confidence estimation," *OptEng*, vol. 34, pp. 2785-2792, 1995.
[27] C. J. Merz and P. M. Murphy, "UCI Repository of Machine Learning Databases," Department of Information and Computer Science, University of California at Irvine,1998.
[28] A. Rosenberg and J. Hirschberg, "V-Measure: A conditional entropy-based external cluster evaluation measure," in *Proc. 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, Prague, 2007, pp. 410-420.
[29] Y. Zhao and G. Karypis, "Criterion functions for document clustering: Experiments and analysis," Department of Computer Science, University of Minnesota, Technical Report TR 01-40, 2001.
[30] R. D. Wilson and T. R. Martinez, "Improved Heterogeneous Distance Functions," *J ArtifIntell Res*, vol. 6, pp. 1-34, 1997.

Y. Villuendas-Rey was born in Ciego de Ávila, Cuba, in 1983. She obtains her B.Sc on Computer Science Engineering in 2005, and her M.Sc. degree on Applied Informatics in 2007, both from the University of Ciego de Ávila, Cuba. She is currently working towards the PhD degree in the Computer Science Department of the University of Las Villas, Cuba.

She has worked as a lecturer in the Computer Science Department of the Engineering faculty of University of Ciego de Ávila, since 2006. She is currently the head of the Artificial Intelligence Research Group of the Computer Science Department of the University of Ciego de Ávila. Her research interests include supervised classification, data preprocessing, clustering and swarm and evolutionary computation .Prof. Villuendas-Rey is member of the Cuban Society of Mathematics, Physics and Computers, the Cuban Society of Pattern Recognition, and the Cuban Society of Pedagogies.

Confidentiality of 2D Code using Infrared with Cell-level Error Correction

Nobuyuki Teraura[1] and Kouichi Sakurai[2]
[1] *Terrara Code Research Institute , Tokai, Japan*
[2] *Information Science and Electrical Enginerring, Kyushu University, Fukuoka, Japan*

***Abstract* — Optical information media printed on paper use printing materials to absorb visible light. There is a 2D code, which may be encrypted but also can possibly be copied. Hence, we envisage an information medium that cannot possibly be copied and thereby offers high security. At the surface, the normal 2D code is printed. The inner layers consist of 2D codes printed using a variety of materials, which absorb certain distinct wavelengths, to form a multilayered 2D code. Information can be distributed among the 2D codes forming the inner layers of the multiplex. Additionally, error correction at cell level can be introduced.**

***Keywords* — confidentiality, error correction, multilayer, 2D code .**

I. NTRODUCTION

THIS paper discusses the advancement of confidentiality by hiding information within an optical information medium based on a paper medium.

A. Background

In optical information media, one-dimensional codes (barcodes) and two-dimensional codes (2D codes) are used as information symbols. Since little data volume can be accommodated, a barcode is applicable when accommodating only identification numbers, such as product numbers. When accommodating comparatively many data, a 2D code is used. The same quantity of data can be printed within a smaller area, so 2D codes have found use in various fields. Moreover, a Web address can be accommodated by a 2D code, making it readable with a mobile cellular telephone. Thus, it is becoming easy for someone to read a 2D code and be guided to a Web site.

B. Motivation

The usual 2D code was developed for ease of reading, like in the abovementioned invocation of a Web site. However, the usual 2D code is not suitable for a use that requires confidentiality, like a credit card transaction. In particular, the usual 2D code is easy to copy, and the fact that the duplicate can be read instead the original is a great defect for such use.

C. Previous Work

In order to prevent reading by a third party, a 2D code that has a secret fill area has been developed [1]. Although this also has an open fill area that can be read as a normal 2D code, the secret fill area can only be read by a unit that possesses the encryption key.

For detecting a manipulation attack and forgery, inserting a digital watermark into a 2D code has been proposed [2]–[5]. Similarly, hiding input data by using steganography has been proposed [6]. Further, embedding data in the 2D code of a multi-stage barcode format by using spectrum spreading has been proposed [7]–[9].

Although a 2D code with the abovementioned secret fill area offers confidentiality and cannot be read by a third party, making a copy of the 2D code is easy. Since it is possible to make a copy that reads like the original, this remains a subject of concern in applications.

A 2D code with a digital watermark inserted can disclose a manipulation attack and forgery at the time of reading, and it can prevent data input. However, the content will probably be accessible to a third party and hence lack confidentiality.

D. Our Contribution

This paper considers improvements in confidentiality and copy protection, which are challenges faced by current optical information media. Furthermore, a basic configuration for a 2D code that solves these two problems simultaneously is proposed.

In the present research, a 2D code using ink that absorbs ordinary visible light is printed on the surface of a paper medium. Two or more 2D codes using inks that absorb different wavelengths of the infrared region are printed in a pile as the inner layers at the bottom, and so a multilayer 2D code is formed. A high degree of confidentiality is realized by distributing the information among the 2D codes forming the inner layers of the multiplex. The encoding table used for distributing information plays the role of an encryption key. The number of combinations for an encoding table is immense. In fact, decryption by a round-robin attack is impossible.

Copying a 2D code of such a configuration by using visible light is also impossible, since the inks of the inner layers are penetrated by visible light. Moreover, a third party who does not know the encoding table used for distributing the information cannot restore the original 2D code. Thus, a 2D

DOI: 10.9781/ijimai.2013.213

code that simultaneously affords great confidentiality and copy protection is realizable. This enables its safe employment on theatre tickets, parking stubs, highway toll coupons, betting slips, etc.

E. Comparison with Related Work

In order to embed discriminable information visibly in a 2D code, an encoding of three bits is utilized. The encoding adds 1 bit of visual data to 2 bits of regular data and arranges these in the shape of an L character. A visual check is made of whether the data are correct [10]. An encoding table approach is proposed for encoding the module in the shape of the L character.

In the present research, the virtual 2D code that incorporates an actual data bit is distributed between two or more real 2D codes using an encoding table, and the pixels of the virtual 2D code are distributed as many small bits using an encoding table.

Whereas the related work mentioned above was aimed at the authentication of the 2D code through a visual inspection by a person, the present research is aimed at improvement of confidentiality through encryption by data variance.

F. Related Work on Identification using an Image

Information media and persons can be objects of automatic identifications using image data. As examples of information media that are identified from images there are barcodes and two-dimensional codes. The main aims of identifying a person are to maintain security and perform an authentication of the person. In the case of information media, such as a two-dimensional code, the data memorized according to the rule of the structure is read. However, the approach to identifying a person involves learning an individual characteristic beforehand and judging the degree of similarity to the person by comparing with the learned pattern. On occasions when the similarity is high, the algorithm will judge the person to be authentic. The body trait used for this identification can be a fingerprint [11], an iris [12], [13], the face [14], etc., and is put to practical use. Since these approaches use a biological feature of the person, they are called biometrics.

II. HIDING OPTICAL INFORMATION

A. Optical Information Media

Although optical information media were originally conceived for identification of alphanumeric characters, this was difficult with the processing capability in those days. Then, symbols representing characters were conceived for computing devices. Even though interpretation would be difficult for people, a symbol easily discerned by a computing device was conceivable. At a time when the microprocessor unit (MPU) had not quite been invented yet, the blue-eye code was devised. Then, the barcode as shown in Fig. 1 was invented with the advent of the MPU and the appearance of a one-dimensional linear sensor. Based on the class of data (numbers, letters, symbols, etc.) accommodated in a barcode, various kinds of barcodes have been devised and currently are in practical use. Moreover, these follow an international standard fixed by the ISO/IEC.

Fig. 1. Examples of EAN-13 [15] and Code 39 [16].

A barcode contains information only in the transverse direction, not in the longitudinal direction. Thus, the lengthwise direction of the symbol functions to provide redundancy. When the central part cannot be read, due to dirt etc., it may be possible to read the upper or lower part instead. However, a large area was needed for printing, and there arose the problem that little data volume was contained. Although the barcode was initially put into practical use by encoding only identification, the need for memorizing a larger volume of data has gradually evolved.

The matrix type and stack type of 2D code, as shown in Figs. 2 and 3, were devised in order to meet that need. These have various characteristics, depending on the specific needs.

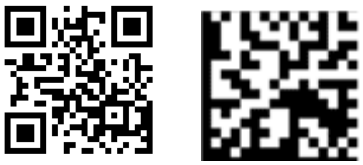


Fig. 2. QR code [17] and data matrix [18] as examples of the matrix type.

Fig. 3. PDF417 code [19] as an example of the stack type.

B. Security of 2D Code

The following are important for the security of a 2D code: confidentiality of data, impossibility of copying, and impossibility of forgery. The proposed scheme not only realizes confidentiality and copy protection but also prevents forgery.

C. Improvements in Confidentiality

The specifications of 2D codes, such as the QR code, are disseminated and readers are put on the market by assorted manufacturers. Nowadays, codes can also be read with a mobile cellular telephone. The 2D code provides confidentiality in comparison with characters, in that a person can see it but cannot understand it. However, for those who have a reader, the content can be read easily. In addition, the following two approaches are used to guarantee confidentiality: decryption at application level and decryption at system level.

Decryption at Application Level

The decryption at application level involves a system that performs data encryption with an application, creates a 2D code, and then performs data decryption with an application at the time of reading. Since it becomes impossible for a third party to interpret data without the decryption key, a guarantee of confidentiality is feasible. However, the facility to process both encryption and decryption is needed for every individual application, and there is the disadvantage of complexity.

Decryption at System Level

The decryption at system level involves both a system that performs data encryption with an application and a reader that performs data decryption using the encryption key chosen beforehand. A 2D code is created by the application to memorize the data encryption. This may include not only the case where all the storage areas are encryption data areas but also the case where the usual non-enciphering fill area combines with an encryption fill area. Although encryption with an application is required by this system, there is the advantage that decoding is unnecessary. Moreover, in the case of a non-enciphering fill area, reading as with a normal 2D code is possible. It can be said that decryption at system level enabled users easily to keep data secret from a third party.

D. Copy Protection

It is impossible to prevent copying of common optical information media. This is because the reading unit operates by receiving the waves reflected from the medium and a copying machine receives a reflection in the same way. Hence, for copy protection, a theoretically different system is required.

Incidentally, optical information media are very easy to reproduce. Thus, in the case of application to a credit card, the account number is not easily read by others, but uses of replicas cannot be prevented.

III. PHYSICAL PRINCIPLE

For secure optical information media, prevention of reading and protection from copying are important. Thus, an information hiding method for simultaneously realizing both objectives has been developed.

With the usual optical information media, as shown in Fig. 4, a wavelength of the visible light region is used for reading. Conversely, with multilayer optical information media, as shown in Fig. 5, there is a characteristic present so that wavelengths of the infrared light region can be used for reading. On materials that reflect infrared light (e.g., paper), a 2D code is printed using a material that absorbs infrared light as shown in Fig. 6. Therefore, a printed segment and unprinted segment that are irradiated with infrared light become possible to discriminate in a manner similar to that for black and white in visible light. Moreover, different segments can be printed using materials with different peak absorption wavelengths in the infrared light region as shown in Fig. 7. It thus becomes possible to acquire the image of a 2D code by using a luminous source that matches the absorption wavelength of each layer at the time of reading, even if the code is printed on top of a multilayer.

The 2D code is printed on the surface of the information medium in a material that transmits infrared light but absorbs visible light. Since only a superficial 2D code can be seen when this configuration is irradiated with visible light, it becomes impossible to copy a lower layer.

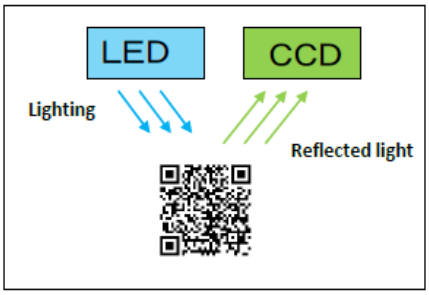


Fig. 4. Principle for reading an optical information medium.

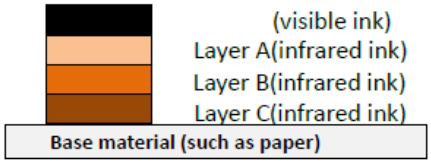


Fig. 5. Multilayer structure.

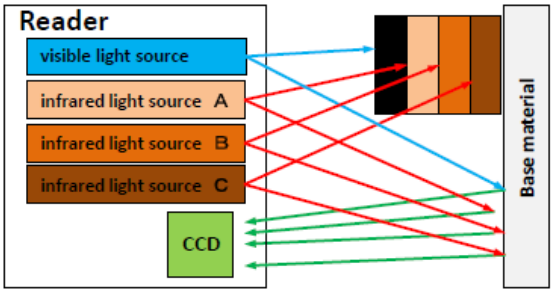


Fig. 6. Principle for reading an optical information medium.

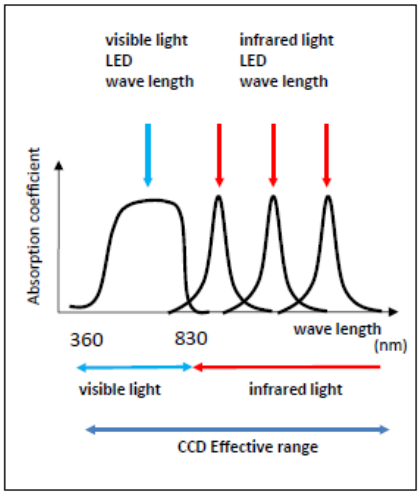


Fig. 7. Light absorption wavelengths of optical multilayer.

IV. Proposed Security System

As mentioned above, copy protection is realized by the hardware for the information media. In multilayer optical information media with the prescribed structure, the number of layers and the absorption wavelengths of the printing materials are used as the means of confidentiality. The information is memorized by distributing it among two or more layers. The confidentiality of the data is realized by introducing an information variance into the hardware configuration.

Various software approaches to confidentiality may be followed. Here, application of the visual secret sharing scheme is examined. With the visual secret sharing scheme the original image is disassembled into two or more images. Consequently, the image data cannot be identified through human vision unless those images are superposed.

In the decryption of a 2D code, the image (symbol) identification is done by image sensors, so the image editing and identifying capabilities of humans cannot be used. However, since the image sensing capability of an image sensor is greater than that of a human, this identification approach is employed profitably.

A. Encryption and Decoding at Cell Level

A data distribution at cell level is realized by distributing the monochrome data of the original 2D code among the 2D codes of the inner layers. For example, the case of three inner layers shown in Fig. 8 will be examined. The data distribution follows the logical table listed as Table I. Four alternative data encodings are possible for white and black, respectively. These are chosen with a random number.

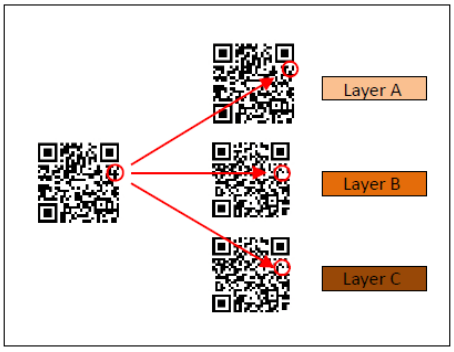


Fig. 8. Distribution to each layer.

TABLE I
Coding Table

Coding			Decoding
W	W	W	B
W	W	B	W
W	B	W	W
W	B	B	B
B	W	W	W
B	W	B	B
B	B	W	B
B	B	B	W

W:White B:Black

Even if a third party is able to read all three of the inner layers, decryption is impossible unless the logical table used for the encoding is known. For three inner layers, there are eight merged colors to distribute, from white-white-white to black-black-black, and four of these are chosen for black (or white). The number of encoding tables is ${}_8C_4$, which becomes 70.

When the number of inner layers, N, is equal to n, the number of hue combinations to distribute, D, is

$$D = 2^n. \tag{1}$$

Since the black (white) half is chosen from these combinations, the number of cell encoding tables, T_S, is

$$T_S(n) = DCD/2 = (2^n)C(2^{n-1}). \tag{2}$$

The number of encoding tables for each number of inner layers is listed in Table II. This is introduced here in order to apply several different encoding tables to each cell. The number of different encoding tables is denoted by L.

TABLE II
Number of Encoding Tables

N	T_S			
	L = 1	L = 2	L = 4	L = 8
1	2	4	8	16
2	6	36	12966	1.68E+08
3	70	4900	2.4E+07	5.76E+14
4	12870	1.65E+08	2.74E+16	7.53E+32
5	6.01E+08	3.61E+17	1.30E+35	1.70E+70

Since a layout is chosen with a random number, the inner layers that carry out a cell distribution may become a hue layout that is greatly inclined toward black or white. In that case, since identification of an optical cell becomes difficult, filtering to generate a uniform layout of white and black is needed. However, in the subcell distribution described below, white and black are represented by a subcell that undergoes the same number of occurrences for each, and so this polarization is corrected.

B. Encryption and Decoding at Subcell Level

A data distribution (information hiding) at subcell level is realized by dividing a cell into several squares (subcells) and distributing those bits among the subcells of the inner layers as shown in Fig. 9.

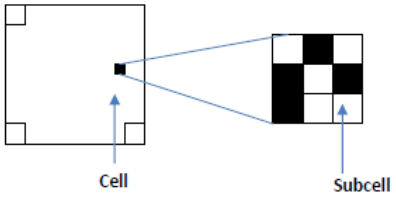


Fig. 9. Units of cell and subcell.

There exist both a horizontal manner and a vertical manner for data distribution at subcell level. For a horizontal distribution system as shown in Fig. 10, after first distributing a cell among the cells of an inner layer, the subcells of the same layer are made to redistribute the distributed cell. Conversely, for a vertical distribution system as shown in Fig. 11, after first distributing a cell among virtual subcells, the subcells of the same layer are made to redistribute the distributed subcells.

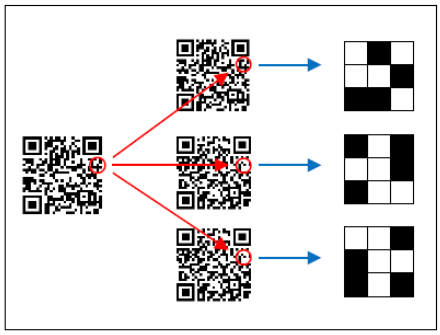


Fig. 10. Horizontal distribution.

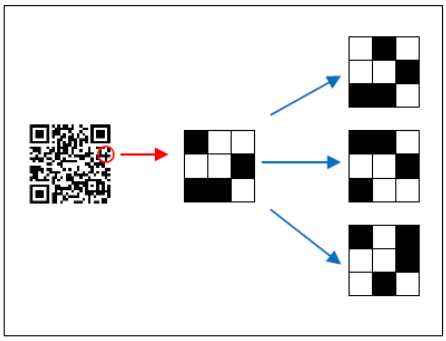


Fig. 11. Vertical distribution.

Here, we argue by using the horizontal manner of distribution for the case in which the number of inner layers is three, and so a cell is decomposed into 3×3 subcells. Moreover, in order to employ the same number of white and black subcells, have symmetry, and make identification of a subcell easy, a central subcell is removed from an encoding area and is always considered white. Then, the number of subcells is set to eight. The distribution to the subcells of a cell is performed using the distribution table (encoding table) listed as Table III. This is the same as the case of the previous cell distribution.

TABLE III
CODING TABLE

Coding								Decoding
B	B	B	B	W	W	W	W	B
B	B	B	W	B	W	W	W	W
B	B	B	W	W	B	W	W	W
B	B	B	W	W	W	B	W	B
B	B	B	W	W	W	W	B	W
B	B	W	B	B	W	W	W	B
...								B
W	W	W	W	B	B	B	B	W

W:White B:Black

The number of hue combinations to distribute is ${}_8C_4$ (i.e., 70). Since these 70 are assigned equally to black and white, the number of bit encoding tables, T_B, becomes

$$T_B = {}_{75}C_{35} \fallingdotseq 1.12 \times 10^{20}. \quad (3)$$

In general, the bit count to distribute is M^2. When the number M^2 is even, the number of hue combinations to distribute, D, becomes

$$D = (M^2)C(M^2/2), \quad (4)$$

Since these combinations are assigned equally to black and white, the number of bit encoding tables, T_B, becomes

$$\begin{aligned} T_B &= DCD/2 \\ &= ((M^2)C(M^2/2))C((M^2)C(M^2/2) / 2). \end{aligned} \quad (5)$$

Moreover, since a central subcell is eliminated from an encoding area when the number M^2 is odd, the number of blacks, B, becomes

$$B = M^2 - 1. \quad (6)$$

Since the black (white) half is chosen, the number of hue combinations of a subcell, D, becomes

$$D = {}_BC_{B/2}. \quad (7)$$

Since these combinations are assigned equally to black and white, the number of encoding tables, T_B, becomes

$$T_B = {}_DC_{D/2}. \quad (8)$$

For M from 1 to 3, the number of hue combinations, D, and the number of encoding tables, T_B, are listed in Table IV.

TABLE IV
NUMBER OF ENCODING TABLES

M	T_B
1	2
2	20
3	1.12E+20

C. Number of Effective Encoding Tables

When the extent of the target 2D code is small and there are few cells, the rule used for the encoding is evident in many of the encoding tables, and so the number of effective encoding tables can be considered to decrease.

If the case of $M = 3$ is examined, the hue combinations amount to 224. Thus, when the cells exceed this number, the previous argument holds. In the case of a cell size of not more than 15 × 15, if the number of cells is set to j, the number of effective encoding tables becomes ${}_jC_{j/2}$

V. VERIFICATION OF PROPOSED SECURITY SYSTEM

When the number of layers N increases, the number of combinations of cell distributions and subcell distributions (horizontal or vertical) will increase. Here, the number of combinations in the case of the encoding table is examined in the cases of from one to three layers and in the general case of n layers. When the number of inner layers is set to N and the number of subcell distributions is set to M^2, the distribution state can be expressed as $P(N, M)$. Below, the number of cases in a distribution is examined using this expression.

A. In the Case of N = 1

Since there is one inner layer in the case of $N = 1$, there is no distribution to the orientation of an inner layer and only a horizontal subcell variance is possible, as shown in Fig. 12.

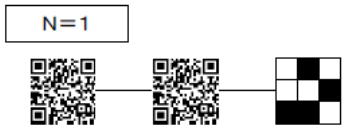


Fig. 12. The case $N = 1$.

If the number of encoding tables for a subcell distribution is set to $T_B(M)$, the number of cases becomes

$$P(1,M) = T_B(M). \tag{9}$$

This is equal to the number of encoding tables for a subcell distribution.

B. In the Case of N = 2

In the case of $N = 2$ there are two kinds of distributions, as shown in Fig. 13. The first distribution is one in which a horizontal subcell distribution of each cell is carried out after performing a cell variance. The other distribution is one in which a vertical subcell distribution is carried out.

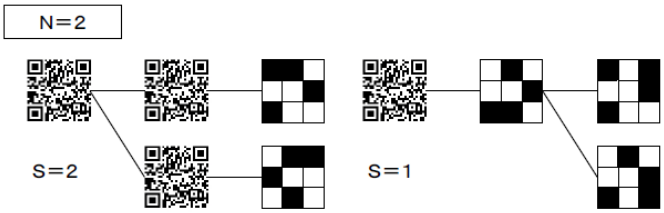


Fig. 13. The case $N = 2$.

If the number of cases at the ith subcell of a virtual cell is set to $S(k)$, the total number of cases is obtained as follows:

$$\begin{aligned} P(2, M) &= S(1) + S(2), \\ S(1) &= T_S(2)\cdot T_B(\mathrm{M}), \\ S(2) &= T_B(\mathrm{M})\cdot T_S(2). \end{aligned} \tag{10}$$

Here, the encoding table of the subcell distribution is assumed identical for each layer. Hereinafter, the same is true.

C. In the Case of N = 3

In the case of $N = 3$ there are four kinds of distributions, as shown in Fig. 14. A characteristic is that not only the horizontal subcell distribution and vertical subcell distribution appear but also the pair in combination.

The number of cases is obtained as follows:

$$\begin{aligned} P(3,M) &= S(1) + S(2) + S(3), \\ S(1) &= T_B(M)\, T_S(3), \\ S(2) &= 2T_S(2)\, T_S(2)\, T_B(M), \\ S(3) &= T_S(3)\, T_B(M). \end{aligned} \tag{11}$$

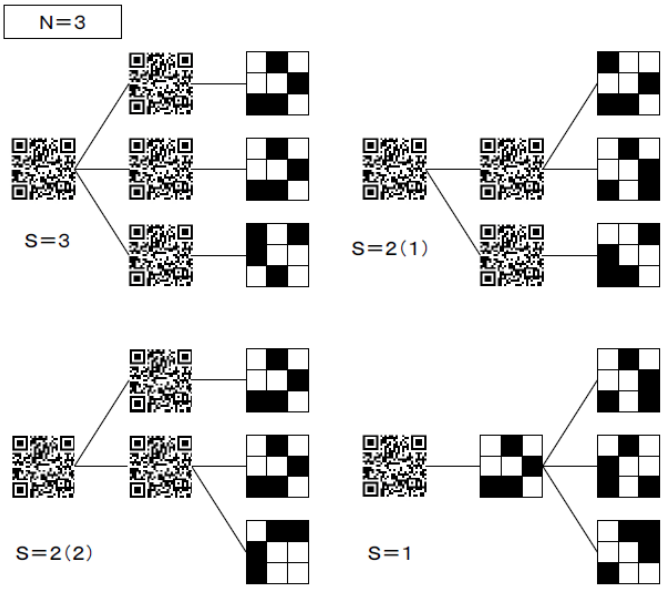


Fig. 14. The case $N = 3$.

D. In the Case of $N = n$

The general case of $N = n$ cannot be illustrated like the examples above, but it can be evaluated. If k is the number of virtual cells, the number of cases is obtained as follows:

$$P(n,M) = \Sigma S(k) = \Sigma E(n,k)\cdot T_S(k)\cdot T_B(M). \quad (13)$$

Here, $T_S(k)$ and $T_B(M)$ are given by (2) and (4), respectively, while $E(n,k)$ is a coefficient listed in Table V. The number of patterns is determined from the numbers for the case of one fewer layer and the case of one fewer layer with one fewer virtual cell. Therefore, the coefficient obeys the following recurrence formula:

$$E(n, k) = E(n–1, k–1) + E(n–1, k). \quad (14)$$

TABLE V
COEFFICIENT VALUES

	E(n,k)					
n	1	2	3	4	5	6
1	1					
2	1	1				
3	1	2	1			
4	1	3	3	1		
5	1	4	6	4	1	
6	1	5	10	10	5	1

VI. CONFIDENTIALITY ESTIMATION

Based on the formula given in the preceding section, the number of calculated encoding tables is listed in Table VI. The subcell distribution table and the table in which the cell encodings differ for each number of inner layers were calculated with the same table.

TABLE VI
NUMBER OF ENCODING TABLES

N	M		
	1(L=8)	2(L=8)	3(L=1)
1	128	1280	7.17E+21
2	3.36E+08	3.36E+09	1.88E+28
3	1.15E+15	1.15E+16	6.45E+34
4	1.50E+33	1.50E+34	8.43E+52
5	3.40E+70	3.40E+71	1.90E+90

This result means that for $N = 1$, $M = 3$, and $L = 1$, or for $N = 3$, $M = 1$, and $L = 8$, the configuration is easily realized and is effective. Furthermore, the number of cases increases with an encoding table, so the number of subcell slices, M, is large and confidentiality becomes great. However, a subcell image becomes difficult for a person to read, because its resolution falls. Thus, confidentiality and legibility form the basis of a trade-off in terms of M.

VII. INTRODUCTION OF ERROR CORRECTING CODE

The size of the subcells introduced above is comparatively small, and so the possibility of a mistaken reading as a result of dirt or other blemishes on the 2D code is larger than with a regular cell. Accordingly, we now discuss the introduction of an error-correction function for the data in the subcells.

A. Extended Hamming Code

A Hamming code corresponding to the whole number m is constituted with a code length $n = 2^m – 1$ and data length $k = n – m$. The data length is the bit count of the original data, and the code length is the bit count of the whole code that is generated. Furthermore, an extended Hamming code has a parity bit added, in order to distinguish between 1- and 2-bit errors.

In the case of $m = 3$, each eight bits consists of four data bits, three error-correction bits, and one parity bit. We now discuss performing error correction using this 8-bit extended Hamming code.

TABLE VII
EXAMPLE OF AN EXTENDED HAMMING CODE

DATA BITS				CORRECTION			PARITY
0	0	0	0	0	0	0	0
0	0	0	1	0	1	1	1
0	0	1	0	1	1	0	1
0	0	1	1	1	0	1	0
0	1	0	0	1	1	1	0
0	1	0	1	1	0	0	1
0	1	1	0	0	0	1	1
0	1	1	1	0	1	0	0

DATA BITS				CORRECTION			PARITY
1	0	0	0	1	0	1	1
1	0	0	1	1	1	0	0
1	0	1	0	0	1	1	0
1	0	1	1	0	0	0	1
1	1	0	0	0	1	0	1
1	1	0	1	0	0	1	0
1	1	1	0	1	0	0	0
1	1	1	1	1	1	1	1

The 3 × 3 subcell configuration has eight subcells because a central subcell is eliminated from the encoding area. We now consider layouts that correspond to the bit array of the extended Hamming code listed in Table VII, where 0 and 1 represent opposite hues. During decryption, subcells with the same hue as the cell are assigned 0 while subcells of the opposite hue are assigned 1, and the extended Hamming code is checked. In this case, on the occasions that a subcell is incorrectly identified due to dirt or blemishes on the 2D code, errors in two subcells can be detected and those in one subcell can be corrected. In addition to this error checking at the subcell level, the 2D code has an error-correction function at the cell level. Thus, correction may occur even when not possible at the subcell level.

The encoding examples shown in Table VIII correspond to the extended Hamming code shown in Table VII.

TABLE VIII
ENCODING EXAMPLES CORRESPONDING TO AN EXTENDED HAMMING CODE

Coding								Decoding
B	B	B	W	B	W	W	W	B
B	B	W	B	W	W	B	W	W
B	B	W	W	W	B	W	B	W
B	W	B	B	W	W	W	B	B
B	W	B	W	W	B	B	W	W
B	W	W	B	B	B	W	W	B
...								...
W	W	W	B	W	B	B	B	W

W:White B:Black

B. Number of Encoding Tables

In the full encoding list, there are 14 lines (Table VIII gives 8 examples) in which four subcells become of the opposite hue. This is adopted as an encoding bit stream. Therefore, the number of encoding patterns D is

$$D = 14. \tag{15}$$

Then, T_B becomes

$$T_B = {}_DC_{D/2} = {}_{14}C_7 \fallingdotseq 1.7 \times 10^7. \tag{16}$$

This value is comparatively small in order to correspond with a round-robin attack. When two encoding tables are used (L = 2), T_B becomes

$$T_B \fallingdotseq 3.0 \times 10^{14}. \tag{17}$$

This result means that for N = 1, M = 3, and L = 2 the configuration is suitable for introduction of error correction.

VIII. EXAMPLE OF VERTICAL DISTRIBUTION

A detailed image of the 2D code that introduces the error-correction function using the extended Hamming code is shown in Fig. 15. The left-hand side is the 2D code before a data distribution, and the right-hand side is the image after the data distribution. A random number was used when encoding each cell of the original 2D code into subcells.

Fig. 15. Images of 2D code.

IX. CONSIDERATION OF PRACTICAL DIFFICULTIES

A. Anti-copying Prerequisites

This paper has discussed the prerequisites that the following three be kept secret: the contents of the inner layers of the 2D code, an infrared wavelength, and the encoding table at the time of data distribution. Therefore, if these pieces of information are known and printing material corresponding to the infrared wavelength is prepared, then creating a replica becomes possible, even though simple copying is impossible.

B. Printing Material and LEDs

The infrared absorption properties of the printing materials now on the market are insufficient to realize multilayer information media [23]. This is because the printing materials now offered have wavelength windows that are wide for absorption of infrared light, so multilayering is not easy. Hence, the development of a marking material with a narrow wavelength window for infrared absorption is expected.

Various LEDs that emit infrared rays have been developed, put on the market, and used for lighting. However, the wavelength range of their infrared light is limited, so infrared LEDs corresponding in wavelength to printing materials still are needed for multilayering.

C. Camera Shake and Focus

In a horizontal distribution, a cell is divided into small subcells and the data are distributed. In order to identify the small subcells, the picturized data need to reveal these as either white or black. However, at the time of image capture, the adjoining subcells of the image may overlap under the effect of camera shake and a clear image may not be obtained. Moreover, identification becomes impossible when the image is not assembled in sharp focus. Camera shake and resultant blurriness that do not pose a concern a cell level may pose a concern at subcell level.

D. Image Resolution

The reading unit of the present 2D code uses about 4 million CCD pixels. Assuming that image formation is carried out with one-quarter of the imaging device, this becomes 1 million pixels or 1000 pixels in each direction. Conversely, when each of the 50 × 50 cells in a 2D code is divided into nine, there will be only 150 bits in each direction. Hence, 6 × 6 pixels will generally be assigned to each bit. According to conventional wisdom, stable identification is possible when there are

3 × 3 pixels in the identification of a so-called atomic unit. Thus, in the identification of the subcells mentioned above, there will be sufficient image resolution.

X. Conclusion

In this paper, multilayer information media using inks that absorb infrared light were introduced, and information media that cannot be copied were proposed. By distributing data among layers, information media with a high degree of confidentiality are realizable. Indeed, the confidentiality has been verified, since combinations of the number of layers and the number of bit distributions that ensure good confidentiality were ascertained.

References

[1] M. Hara, "Method for producing 2D code reader for reading the 2D code," US patent application 20090323959, by Denso Wave Inc., Patent and Trademark Office, 2009.

[2] J.-J. Shen and P.-W. Hsu, "A fragile associative watermarking on 2D barcode for data authentication," *Int. J. Network Security*, vol. 7, no. 3, pp. 301–309, Nov. 2008.

[3] L. Li, R.-L. Wang, and C.-C. Chang, "A digital watermark algorithm for QR code," *Int. J. Intell. Inform. Process.*, vol. 2, no. 2, pp. 29–36, 2011.

[4] F. Ming, H. Ye, and L. Hong, "A method of 2D BarCode anti-counterfeit based on the digital watermark," *J. Changsha Commun. Univ.*, vol. 24, no. 2, pp. 85–89, 2008.

[5] M. Sun, J. Si, and S. Zhang, "Research on embedding and extracting methods for digital watermarks applied to QR code images," *New Zealand J. Agric. Res.*, vol. 50, pp. 861–867, 2007.

[6] W.-Y. Chen and J.-W. Wang, "Nested image steganography scheme using QR-barcode technique," *Opt. Eng.*, vol. 48, no. 5, 057004, May 2009.

[7] N. Xiamu, H. Wenjun, W. Di, and Z. Hui, "Information hiding technique based on 2D barcode," *Acta Sci. Natur. Univ. Sunyatseni*, vol. 43, pp. 21–25, 2004.

[8] L. Xue, Y. Chao, L. Liu, and X. Zhang, "Information hiding algorithm for PDF417 barcode," in *Proc. 5th Int. Conf. Natural Computation (ICNC '09)*, Tianjian, 2009, vol. 6, pp. 578–581.

[9] Z. Bo and H. Jin, "Information hiding technique based on PDF417 barcode," *Comput. Eng. Design*, vol. 19, pp. 4806–4809, 2007.

[10] C. Fang and E.-C. Chang, "Securing interactive sessions using mobile device through visual channel and visual inspection," in *Proc. 26th Annu. Computer Security Appl. Conf. (ACSAC '10)*, Austin, TX, 2010, pp. 69–78.

[11] M. Jampour, M. M. Javidi, A. S. Nejad, M. Ashourzadeh, and M. Yaghoobi, "A new technique in saving fingerprint with low volume by using chaos game and fractal theory," *Int. J. Interact. Multimed. Artif. Intell.*, vol. 1, no. 3, pp. 28–32, 2010.

[12] V. L. Hidalgo, L. M. Garcia, and M. T. Lorenzo, "Iris recognition using the JavaVis library," *Int. J. Interact. Multimed. Artif. Intell.*, vol. 1, no. 1, pp. 43–48, 2008.

[13] M. F. Hurtado, M. H. Langreo, P. M. de Miguel, and V. D. Villanueva, "Biometry, the safe key," *Int. J. Interact. Multimed. Artif. Intell.*, vol. 1, no. 3, pp. 33–37, 2010.

[14] J. A. Dargham, A. Chekima, and E. G. Moung, "Fusing facial features for face recognition," *Int. J. Interact. Multimed. Artif. Intell.*, vol. 1, no. 5, pp. 54–60, 2012.

[15] *Information technology — Automatic identification and data capture techniques — EAN/UPC bar code symbology specification*, ISO/IEC 15420:2009.

[16] *Information technology — Automatic identification and data capture techniques — Code39 bar code symbology specification*, ISO/IEC 16388:2007.

[17] *Information technology — Automatic identification and data capture techniques — QR Code 2005 bar code symbology specification*, ISO/IEC 18004:2006.

[18] *Information technology — Automatic identification and data capture techniques — Data Matrix bar code symbology specification*, ISO/IEC 16022:2006.

[19] *Information technology — Automatic identification and data capture techniques — PDF417 barcode symbology specification*, ISO/IEC 15438:2006.

[20] R. Villan, S. Voloshynovskiy, O. Koval, and T. Pum, "Multilevel 2D bar codes: Towards high-capacity storage modules for multimedia security and management," *IEEE Trans. Inf. Forens. Security*, vol. 1, no. 4, pp. 405–420, 2006.

[21] N. Degara-Quintela and F. Perez-Gonzalez, "Visible encryption: Using paper as a secure channel," in *Proc. SPIE-IS&T Electronic Imaging 2003: Security and Watermarking of Multimedia Contents V*, Santa Clara, CA, 2003, vol. 5020, pp. 413–422.

[22] S. Ono, "Fusion of interactive and non-interactive evolutionary computation for two-dimensional barcode decoration," in *Proc. 2010 IEEE Congress on Evolutionary Computation (CEC)*, Barcelona, 2010, pp. 1–8.

[23] TOSCO Co. Available: http://www.tosco-intl.co.jp/cgibin/imgchemi/general_display_sensitive.php?id=85&code=SDA5677.

Nobuyuki Teraura received B.S. and M.S. degrees in engineering from the Faculty of Engineering, Nagoya University, Japan. He is a Researcher at the Terrara Code Research Institute, Tokai, Japan. He is also an external PhD student of the Department of Informatics, Kyushu University, Fukuoka, Japan. Mr. Teraura is a member of the IEICE and the Information Processing Society of Japan.

Kouichi Sakurai received a B.S. degree in mathematics from the Faculty of Science, Kyushu University, Fukuoka, Japan, and an M.S. degree in applied science from the Faculty of Engineering, Kyushu University. He received his doctorate in engineering from the Faculty of Engineering, Kyushu University. He works in the capacity of Professor for the Department of Information Science and Electrical Engineering at Kyushu University. Prof. Sakurai is a member of the IEICE, Information Processing Society of Japan, Mathematical Society of Japan (IPSJ), IEEE, and ACM.

Mining Social and Affective Data for Recommendation of Student Tutors

Elisa Boff[1] and Eliseo Berni Reategui[2]
[1]*Universidade de Caxias do Sul (UCS) – Brazil*
[2]*Universidade Federal do Rio Grande do Sul (UFRGS) – Brazil*

***Abstract* — This paper presents a learning environment where a mining algorithm is used to learn patterns of interaction with the user and to represent these patterns in a scheme called item descriptors. The learning environment keeps theoretical information about subjects, as well as tools and exercises where the student can put into practice the knowledge gained. One of the main purposes of the project is to stimulate collaborative learning through the interaction of students with different levels of knowledge. The students' actions, as well as their interactions, are monitored by the system and used to find patterns that can guide the search for students that may play the role of a tutor. Such patterns are found with a particular learning algorithm and represented in item descriptors. The paper presents the educational environment, the representation mechanism and learning algorithm used to mine social-affective data in order to create a recommendation model of tutors.**

***Keywords* — Collaboration, Learning Environment, Recommender Systems, Social-Affective Data.**

I. INTRODUCTION

MINING data in educational environments is often used with two main purposes:

(1) to give educators a better understanding of how users learn with the system;

(2) to define different paths of study according to students' profiles learned from data.

The first goal may be achieved by using mining algorithms to identify patterns and represent them in a scheme that is easy to understand. The second goal can be pursued by employing a mechanism capable of using the patterns found to suggest topics related to the subjects being studied.

We used mining algorithms here in order to accomplish both purposes (1 and 2), and also to identify suitable student tutors that may help other students needing assistance. The use of data mining in Education has expanded considerably in the last decade mostly because of the growing number of systems that store large databases about students, their accesses to material available, their assignments and grades. Such expansion in the field yielded the establishment of a community concerned mostly with the development of methods for exploring data coming from educational settings, and employing those methods to better understand students and learning processes [4].

Current research has shown the potentiality of cooperative learning, demonstrating that group work is fundamental for the cognitive development of the student [7] [8]. It is known that knowledge composition occurs on an individual basis, but cooperation (subjects acting together over the same topic, with common goals, interacting and exchanging ideas) is capable of involving all participants in learning [18]. In this perspective, motivating the students to interact can lead to an effective learning practice.

The recommendation service of tutors works in the sense of motivating group formation among the students. According to Andrade [1], a group can be formed due to similarity and empathy of its members or to the necessity of support for the accomplishment of some task. The latter can be motivated by prestige or status, economic benefits or the necessity and desire of contribution. [1] also says that the affective states of the individuals have significant importance in the interaction process. The author complements affirming that some dimensions of the personality seem to have certain connections with the social performance in the interaction, but establishing an accurate relationship between them seems to be a complex task.

Our tutor recommendation service explores the social-affective dimension through the analysis of emotional states and social behavior of the users. A recommender system analyses students' interactions and finds suitable tutors among them as well as contents to be recommended. A specific algorithm was built to identify behavioral patterns in the students interaction, and to store this knowledge in structures called item descriptors [19]. The method proposed shows a good performance with respect to processing time and accuracy, and has an advantage over other techniques when it comes to understanding the knowledge elicited and letting users modify it. The first section of the paper gives an overview of the types of data collected from the interaction with the users. Then, the mechanism employed to represent knowledge is explained, in addition to its learning algorithm and recommendation process. Finally, preliminary results are discussed, as well as conceptual advantages and drawbacks of the approach. The last section of the paper offers conclusions and directions for future work.

DOI: 10.9781/ijimai.2013.214

II. Collecting Interaction Data

When students navigate in our learning environment (Fig.2), different types of data are collected from their interaction. By keeping the navigation history of every student, for example, we are able to identify navigation patterns and to use them in real-time recommendation of contents. For the recommendation of tutor colleagues, six other types of data are collected: Social Profile; Acceptance Degree; Sociability Degree; Mood State; Tutorial Degree and Performance.

The Social Profile (SP) is built during the communication process among students. The following information is collected during the interaction of the students through an instant message service:

• *Initiatives of communication*: number of times that the student had the initiative to talk with other pupils.

• *Answers to initial communications*: in an initial communication, number of times that the student answered.

• *Interaction history*: individuals with whom the student interacts or has interacted, and number of interactions.

• *Friends Group*: individuals with which the student interacts regularly, and number interactions.

Based on Maturana [15] we defined the Acceptance Degree (AD), which measures the acceptance a student has for another one. Such data is collected through a graphical interface that enables each student to indicate his/her acceptance degree for other students. This measure may also be considered from a point of view of Social Networks, which constitutes one of the most popular approaches for the analysis of human interactions. The most important concept in this approach is centrality. If an individual is central in a group, he/she is popular and gets a great amount of attention from the group members. As the AD is indicated by the students themselves based on their affective structures, the measurement can indicate diverse emotions, such as love, envy, hatred, etc. The average of all AD received by a student influences his/her Sociability Degree (SD).

The Mood State (MS) represents our belief in the capability of a student to play the role of a tutor if he/she is not in a positive mood state (although the student may have all the technical and social requirements to be a tutor). We consider three values for the MS: "bad mood", "regular mood" and "good mood". These states are indicated by the students in a graphical interface through corresponding clip-arts.

After a helping session, a small questionnaire is submitted to the student who got assistance. The goal of this questionnaire is to collect information about the performance of the tutor. The questions made are based on concepts from Social Networks and Sociometry, and may be answered by four qualitative values: "excellent", "good", "regular", and "bad". They are:

• *How do you classify the sociability of your class fellow?*

• *How do you classify the help given by your class fellow?*

The answer to the first question together with the average of the ADs of a student, form his/her Sociability Degree (SD). This measure indicates how other individuals see the social capability of this student.

The Tutorial Degree (TD) measures a student's pedagogical capacity to help, to explain and teach. This value is obtained from the answers given for the second question of the questionnaire above and from the marks the tutor got when he/she studied the contents for which he/she was asked for help. These marks were called Performance (P) and were used in the computation of the TD because when a tutor is not able to help another student it does not necessarily mean that the student is a bad tutor. He/she may simply not know very well the content for which his/her help was requested. Therefore, the answers of the students have to be "weighted".

A mining process determines relationships among these factors, and represents such relationships in item descriptors, which are later used for recommendation purposes.

III. The Item Descriptors

An item descriptor represents knowledge about when to recommend a particular item (a topic of study, an exercise, or a tutor) by listing other items found to be related to it. Users have features that may be classified as:

• *demographic*: data describing an individual, such as age, gender, occupation, address;

• *behavioral*: data describing tutoring and social capacity, navigation and study patterns.

It has been shown that both types of data are important when building a user profile [13] and inferring user's needs [5] [6]. Demographic material is represented here in attribute-value pairs. Behavioral information is represented by actions carried out by the user, such as the selection of a topic for reading. Emotional states and social behavior can either be inferred or collected explicitly in questionnaires.

While attributes used to define demographic features are typically single-valued, behavioral data is usually multi-valued. For instance, a person can only belong to one age group (demographic), but he/she may be friendly and patient at the same time (behavioral). Nevertheless, both types of information are represented in our model in a similar way. Let us examine an example of an item descriptor and its related items (Table 1).

TABLE I
ITEM DESCRIPTOR AND RELATED ITEMS

DESCRIPTOR D_N	
Correlated terms	**Confidence**
t_a	0.92
t_e	0.87
t_c	0.85
t_d	0.84
t_b	0.77

The descriptor has a *target* (d_n), i.e. an item that may be recommended in the presence of some of its correlated terms. Each term's class and *confidence* (the strength with which the

term is correlated with the target item) is displayed next to its identification.

We use *confidence* as a correlation factor in order to determine how relevant a piece of information is to the recommendation of a given item. This is the same as computing the conditional probability $P(dj|e)$, i.e. the probability that the item represented by descriptor d_j is rated positively by a user given evidence e. Therefore, the descriptors can be learned through the analysis of actual users' records. For each item for which we want to define a recommendation strategy, a descriptor is created with the item defined as its target. Then, the confidence between the target and other existing demographic features and behavioral data is computed. This process continues until all descriptors have been created. For the recommendation of tutors, descriptors are built indicating the features of good and bad instructors.

IV. The Recommendation of Tutors

Collaborative Filtering, one of the most popular technologies in recommender systems [15], has been used in the past in several research projects, such as Tapestry [13], GroupLens [27], and more recently in related research focusing on the extraction of information from social networks [9][21]. The technique is based on the idea that the active user is more likely to prefer items that like-minded people prefer [28]. To support this, similarity scores between the active user and every other user are calculated. Predictions are generated by selecting items rated by the users with the highest degrees of similarity.

Here, a different approach has been followed, as the main idea in the project was not to keep track of users' interests, but to evaluate their willingness to collaborate. This task, called here recommendation of tutors, is explained below.

Given a list of possible tutors $U=\{u_1, u_2,..., u_m\}$, the recommendation process starts with the gathering of demographic and behavioral information about each of them. Next, the data collected for each user is matched against a descriptor d_j which lists the most important features of good instructors, according to the terms $T=\{t_1,t_2,...,t_k\}$ stored in the descriptor. The system computes a score for each student that ranges from not similar (0) to very similar (1), according to the formula:

$$Score\ (d_j) = 1 - \prod_{kji} (Noise\ (t_p))$$

where $Score(d_j)$ is the final score of the descriptor d_j; $Noise(t_p)$ is the value of the noise parameter of term t_p, a concept used in noisy-OR probability models (Pradhan et al., 1994) and computed as $1 - P(d_j \mid t_p)$. The individual with the highest score is selected to assist the student needing assistance.

That expression contains an assumption of independence of the various t_p - which the designer of a practical system should be trying to achieve in the choice of terms. Ultimately the test of the assumption is in the users' perception of the quality of a system's recommendations: if the perception is that the outputs are fully satisfactory, this is circumstantial evidence for the soundness of the underlying design choices. The situation here is the same as in numerical taxonomy [21], where distances between topics id in a multidimensional space of attributes are given by metric functions where the choice of distinct dimensions should obviously aim to avoid terms that have mutual dependences. If the aim fails, the metric cannot - except occasionally by accident - produce taxonomic clusters C (analogous to sets of topics offered by a recommender system once a user has selected one member of C) that satisfy the users. This method is based on the assumption that any term matching the user's terms should increase the confidence that the descriptor holds the most appropriate recommendation. In a real-life example, let us suppose that we have a certain degree of confidence that a student who has shown a good ability in answering factorial exercises is our best bet to help another student who is having problem with the subject. Knowing that that same student is friendly and is in a good mood should increase the total confidence on his recommendation as a tutor, subject to not exceeding the maximum value of 1.

The Virtual Character is the interface element that delivers to student the result of recommendation process in natural language (Fig.1).

The knowledge base of the Virtual Character stores knowledge about Algorithms, enabling the character to assist students mainly in theoretical questions. The Artificial Intelligence Markup Language (AIML) is used to represent the character's conversational knowledge [30], employing a mechanism of stimulus-response. The stimuli (sentences and fragments which may be used to question the agent) are stored and used to search for pre-defined replies. The most important AIML tags are:

- **<aiml>**: indicates the beginning of a document.
- **<category>**: the simplest knowledge unit in AIML. Each category consists of an input question, an output answer and an optional context. The question, or stimulus, is called the pattern, while the answer is called the template.
- **<pattern>:** keeps a set of words which is searched for in sentences which the user may enter to communicate with the virtual character. The language that may be used to form the patterns includes words, spaces, and the wildcard symbols _ and *;
- **<template>**: when a given pattern is found in the input sentence, the corresponding template is returned and presented to the user. In its simplest form, a pattern is a word and the template consists of plain text. However, the tags may also force the conversion of the reply into a procedure which may activate other programs and recursively call the

pattern matcher to insert the responses from other categories.

The optional context of a category enables the character to remember a previous statement. This feature, together with the possibility of launching particular programs when a certain pattern is found, makes the AIML communication mechanism very distinct from a simple retrieval of questions and answers from a database.

Fig. 1. Recommendation example.

The user's affective state is also considered in order to choose the type of language the character uses to talk at a given moment. The affective state is entered as a pattern which has to be matched for the selection of a given sentence. For instance, the pattern *RECURSION* is modified into *RECURSION CHEERFUL* if the user is in a cheerful mood.

In addition to the existing AIML tags, new ones were created to manage the agents' emotional appearance. For instance, we created the tag <*humor*> to control the image changes reflecting different moods of the virtual character (happy, receptive, annoyed, etc).

Therefore, when the user poses a question (stimulus), the character starts the AIML Retrieval Mechanism in order to build an appropriate reply using the information, patterns and templates from the AIML database. A suitable picture of the character is picked from the Image Database to match the sentence retrieved according to the humor tag.

In addition to being able to answer questions in natural language, our character is also able to monitor the actions of each student and notice, for instance, that a particular topic is related to a given exercise. Such a behavior is achieved through the use of the template tag to launch the recommender system, which looks for appropriate activities and contents to each student.

V. Validation and Discussion

An Environment for the Learning of Algorithms (A3), Fig. 1, has been developed at the Department of Computer Science of the University of Caxias do Sul with the main goal of making the courses more dynamic, increasing the interest and participation of the students and providing an environment where students may interact in order to improve their knowledge. The environment presents students with the regular contents of algorithms (central area of Fig.2), it proposes exercises, provides a forum for discussion and a tool for the testing and running of algorithms. All website functions can be accessed by the left menu on the detail 3 of Fig.2. Having been developed as a dynamic website, the system enables teachers and administrators to modify contents easily. Online users are shown in the interface (detail 2 of Fig.2). And most importantly, the system promotes the communication among students by suggesting individuals that may help others showing difficulty in learning a given topic. The recommendation is present in the detail 4 of Fig.2, below the image of Virtual Character.

Fig. 2. Environment for the Learning of Algorithms (A3).

The Affective States of students describe social-affective data which is used to recommend students tutors. The system does not try to infer social-affective states, but the user deliberately informs it about how he/she feels at login time (detail 1 in Fig.2). This information is used to define the type of language and stimuli that our Virtual Character has to show in order to communicate better with the user.

The A3 environment started to be tested in 2 courses at the Department. Descriptors were built manually in order to get the system to recommend contents and tutors. The data collected so far has not been sufficient for us to carry out conclusive experiments as to whether the system is making tutoring recommendations appropriately. However, initial experiments carried out and reported in Reategui [19] show that the item descriptors have a good performance in terms of

processing time and accuracy, when compared with collaborative filtering, one of the most popular approaches in recommender systems.

For the MovieLens database[1], for example, storing anonymous ratings of 3900 movies assigned by 6040 users, the item descriptors show an accuracy rate that is 6 points higher than that of the k-nearest neighbor algorithm. The Table 2 summarizes the results obtained.

The experiments were carried out considering neighborhoods with sizes 1, 20 and 40 (we did not observe any significant improvement in accuracy for the nearest-neighbor algorithm with neighborhoods larger than 40). The topic descriptors performed better than the k-nearest-neighbor algorithm, no matter what size of the neighborhoods was chosen.

Sarwar [20] have carried out a series of experiments with the same data set, employing the Mean Absolute Error (MAE) method to measure the accuracy of item-based recommendation algorithms. The results reported could not be compared directly with our own as the authors computed their system's accuracy using the MAE and considering integer ratings ranging from 1 to 5 (reaching values around 75%). In our experiment, we only took into account whether a user rated (1) or did not rate (0) a topic.

TABLE II
SCORING RESULTS FOR THE MOVIELENS DATA SET

Method	Scoring
Item Descriptors	65,7
k-nearest-neighbor (k=1)	39,3
k-nearest-neighbor (k=20)	54,9
k-nearest-neighbor (k=40)	59,7

In order to evaluate the system's performance, we monitored how much time was spent by the system in order to recommend the 2114 topics in the test data set[2]. For k=1, the nearest-neighbor approach needed less time than the topic descriptors to perform the tests, though showing a lower rate of accuracy. However, for larger values of k (or simply larger numbers of users) the performance of the nearest-neighbor algorithm degrades, while that of the topic descriptors remains stable. Table 3 summarizes the results of the experiment.

In more realistic situations where the nearest-neighbor algorithm may have to access a database containing actual users' transactions, the nearest-neighbor approach may become impractical. For the same experiment described above, we tested the nearest-neighbor through access to an actual database, using k=10. A few hours was needed for the system to make the whole set of recommendations. Further validation results may be found in Reategui [19].

Another popular approach applied to recommender systems is association rules [14] (Mombasher, 2001). This technique use well-known inductive learning algorithms, such as *a priori* [2], to extract knowledge and represent them in "if ... then ..." rules format. The main advantage of such learning method relies on the robustness and stability of the algorithms available. Although being successfully applied in innumerable application areas, association rules are hard to modify while keeping the rule base consistent (e.g. adding new rules without contradicting existing ones). Keeping track of and trying to understand the large number of generated rules for each topic is another difficulty of this approach.

TABLE III
PERFORMANCE RESULTS FOR THE MOVIELENS DATA SET

Method	**Time spent in secs.**
Topic Descriptors	32
k-nearest-neighbor (k=1)	14
k-nearest-neighbor (k=20)	43
k-nearest-neighbor (k=40)	86

The item descriptor approach is different in that it represents knowledge in the form of descriptors and correlation factors. When compared with the other approaches in this respect, descriptors are interesting because they make it easy for users to understand as well as modify the knowledge represented. This is particularly important when the user wants to make the system respond in a certain way in given circumstances, e.g. if the teacher wants the system to recommend a certain reading when the student is viewing a particular topic.

The learning mechanism used on the item descriptors also exploits well-known methods to compute correlation factors and define the strength of the relationships among features and topics. The option to use term confidence instead of conditional probability to describe the model comes from the fact that other correlation factors that are not supported by probability theory are computed by the system, such as interest and conviction [4]. However, at present these are provided only to let the user analyze and validate the knowledge extracted from the database. We are currently testing different variations on the combination of these factors in the reasoning process.

Although the system learns and updates its descriptors in an offline process (therefore not critical for the application to recommend topics in real time), our learning algorithm is fairly simple and fast. Above all, it is faster than algorithms that group evidence and try to compute the relevance of each topic and then of each group of evidence.

Our model may also be compared with Hidden Markov Models (HMM), employed in tasks such as the inference of grammars of simple language [10], or the discovery of patterns in DNA sequences [3]. The two models are similar in that both

1 MovieLens is a project developed in the Department of Computer Science and Engineering at the University of Minnesota (http://movielens.umn.edu).

2 The tests were performed on a PIII 500MHZ PC with 128Mb of RAM.

use probability theory to determine the likelihood that a given event takes place. However, the actual methods used to compute probabilities of events are different: while HMM considers the product of the probabilities of individual events, we consider the product of noise parameters. Both models are based on the assumption that an output is statistically independent of previous outputs. This assumption may be limiting in given circumstances, but for the type of application we have chosen, we do not believe this to be a serious problem (e.g. as we have remarked above in our comments on independence). To take one practical example, the probability that a user studies topic C is very rarely dependent on the order in which users have read other topics (e.g. B before A, or A before B).

The recommendation method we use has the peculiarity of computing the correlation of individual terms initially, and then combining them in real time. This is analogous to finding first a set of rules with only one left-side term, followed at run time by finding associations between the rules. This is a good technique to avoid computing the relevance of all possible associations among terms in the learning phase.

Gomes [11] proposes a different recommendation strategy to identify tutors based on the computation of a utility function. Their strategy combines features in a mathematical expression to determine how effective a student can be for a given tutoring task. Compared to this approach, our mining and recommendation mechanism is more interesting in that it uses learning algorithms to learn a model from the available data automatically, identifying the importance of each utility function variable.

VI. Conclusion

One important contribution of this work has been the definition of the types of data to be used in the mining and in the recommendation process of student tutors. Using the descriptors to calculate the relevance of terms individually, and then combining them at recommendation time through the use of the noisy-OR is also a novel approach. A similar use of the function can be found in research on expert systems [9], but not in applications for recommender systems. Initial results have shown that the approach can be very effective in large-scale practice for personalization purposes.

The use of social-affective information to promote the communication and collaborative learning among students is starting to be tested in the environment A3. The results obtained so far show that the use of Social Profile, Mood State, Performance Acceptance, Sociability and Tutorial Degree in tutor recommendation, is a promising alternative.

Although the data collected from students' interactions so far are not sufficient for us to draw assertive conclusions about the use of item descriptors to recommend tutors, other experiments have shown the adequacy of the approach in item recommendation.

The possibility to represent different types of information (demographic or behavioral) in a similar way seems to be advantageous when it comes to practical implementation issues. Previous work in the field has shown the importance of dealing with and combining such types of knowledge in recommender systems [17]. Current research on the identification of implicit user information also shows that recommender systems will have to manipulate different sorts of data in order to infer users' preferences [6].

One of our biggest challenges now concerns the automatic inference of students' affective states. At present we are using questionnaires and graphic interface controls to let the users indicate such states. Thus, little is done to automatically infer the social-affective information necessary for tutor recommendation. This will be one of our main research efforts in the near future.

This project should also be integrated with the JADE/MAIDE platform [11] [22] and have its knowledge used in the MACE platform [1].

References

[1] Andrade, A.; Jaques, P.; Viccari, R.; Bordini, R. and Jung, J. 2001. A Computational Model of Distance Learning Based on Vygotsky's Socio-Cultural Approach. In Proceedings of Multi-Agent Based Learning Environments Workshop. International Conference on Artificial Intelligence on Education. Antonio, Texas, May 19-23.

[2] Agrawal, R. and Srikant, R. 1994. Fast Algorithms for Mining Association Rules. In Proceedings of the 20th International Conference on Very Large Databases, Santiago, Chile.

[3] Allison, L.; Stern, L.; Edgoose, T. and Dix, T. I. 2000. Sequence Complexity for Biological Sequence Analysis. Computers and Chemistry 24(1):43-55.

[4] Baker, R. S. J. D., Yacef, K. 2009. The State of Educational Data Mining in 2009: A Review and Future Visions, Journal of Educational Data Mining, 1(1): 3-17.

[5] Brin, S.; Motwani, R.; Ullman, J. D.; and Tsur, S. 1997. Dynamic topicset counting and implication rules for market basket data. SIGMOD Record (ACM Special Interest Group on Management of Data), 26(2):255.

[6] Buchner A. and Mulvenna, M. 1998. Discovering Behavioural Patterns in internet files. Proceedings of Intelligent Tutoring Systems '98, San Antonio, Texas, USA. 16-19. Lecture Notes in Computer Science, 1452.

[7] Claypool, M.; Brown, D.; Le, P.; and Waseda, M. 2001. Inferring User Interest. IEEE Internet Computing, 5(6):32-39.

[8] Echeita, G., Martin, E. 1995. Interação Social e Aprendizagem. Porto Alegre: Artes Médicas.

[9] Finkelstein, A., Lim, S. L. 2012. Using Social Networks and Collaborative Filtering for Large-Scale Requirements Elicitation. IEEE Transactions on Software Engineering, 38(3): 707-735.

[10] Gallant, S. 1988. Connectionist Expert Systems. Communications of the ACM, Vol. 31, Issue 2.

[11] Georgeff, M. P. and Wallace, C. S. 1984. A General Selection Criterion for Inductive Inference. European Conference on Artificial Intelligence – ECAI 84, Pisa, 473-482.

[12] Gomes, E. R; Boff, E. and Viccari, R. 2004. Social, Affective and Pedagogical Agents for the Recommendation of Student Tutors. Proceedings of The Seventh Intelligent Tutoring Systems, Workshop on Social and Emotional Intelligence on Learning Environments, Maceió, Brazil, August, 31.

[13] Goldberg, D., Nichols, D., Oki, B. M. & Terry, D. 1992. Using collaborative filtering to weave an information tapestry. Communications of the ACM, 35(12).

[14] Gomes, E. R.; Silveira, R. A. and Viccari, R. M. 2003. Utilização de agentes FIPA em ambientes para Ensino a Distância. In: XI Congresso Iberoamericano de Educação Superior em Computação (CIESC), Bolívia.

[15] Herlocker, J., Konstan, J. & Riedl, J. 2000. Explaining Collaborative Filtering Recommendations. In Proceedings of ACM Conference on

Computer Supported Cooperative Work. Philadelphia, Pennsylvania, USA.

[16] Krulwich, B. 1997. Lifestyle Finder: Intelligent User Profiling Using Large-Scale Demographic Data, Artificial Intelligence Magazine 18(2). 37-45.

[17] Lin, W.; Alvarez, S. A.; and Ruiz, C. 2000. Collaborative Recommendation via Adaptive Association Rule Mining. KDD-2000 Workshop on Web Mining for E-Commerce, Boston, MA, USA.

[18] Maturana, H. 1998. Emoções e Linguagem na Educação e na Política. Belo Horizonte: Ed. UFMG.

[19] Mobasher, B., Dai, H., Luo, T. and Nakagawa, M. 2001. Effective Personalization Based on Association Rule Discovery from Web Usage Data. Proceedings of the ACM Workshop on Web Information and Data Management. Atlanta, Georgia, USA.

[20] Nitzke, J., Carneiro, M. L., Franco, S. 2002. Ambientes de Aprendizagem Cooperativa Apoiada pelo Computador e sua Epistemologia. Informatica na educação: Teoria & Prática. Porto Alegre, v.5, n.1, p.13-23.

[21] Noel, J., Sanner, S., Khoi-Nguyen, T., Christen, P., Xie, L. Bonilla, E., Abbasnejad, E., Penna, N. D. 2012. New Objective Functions for Social Collaborative Filtering. ACM International World Wide Web, Lyon, France, 859-868.

[22] Pazzani, M. 1999. A Framework for Collaborative, Content-Based and Demographic Filtering. Artificial Intelligence Review. 13(5-6): 393-408.

[23] Piaget, J. 1973. Estudos Sociológicos. Rio de Janeiro: Companhia Editora Forense.

[24] Pradhan, M., Provan, G. M., Middleton, B., and Henrion, M. 1994. Knowledge engineering for large belief networks. In Proceedings of Uncertainty in Artificial Intelligence, Seattle, Washington. Morgan Kaufmann.

[25] Reategui, E.; Campbell, J. A. and Torres, R. 2004. AAAI - The Nineteenth National Conference on Artificial Intelligence, Workshop on Semantic Web Personalization, San Jose, USA, July.

[26] Sarwar, B. M.; Karypis, G.; Konstan, J. A. and Riedl, J. 2001. Topic-based collaborative filtering recommendation algorithms. In Proceedings of the 10th International World Wide Web Conference - WWW10, Hong Kong.

[27] Sarwar, B., Konstan, J., Borchers, A., Herlocker, J., Miller, B. & Riedl, J. 1998. Using Filtering Agents to Improve Prediction Quality in the GroupLens Research Collaborative Filtering System. Proceedings of the Conference on Computer Supported Cooperative Work. Seattle, USA, 345-354.

[28] Shardanand, U. & Maes, P. 1995. Social information filtering: Algorithms for automating "word of mouth". In Proceedings of Human Factors in Computing Systems - CHI, Denver, Colorado, USA. 210-217.

[29] Sneath, P. A. and Sokal, R. R. 1973. Numerical Taxonomy: The Theory and Practice of Numerical Classification. San Francisco, CA: Freeman.

[30] Wallace, R. The Elements of AIML Style, ALICE A. I. Foundation. 2003.

Elisa Boff has her Ph.D. in Computer Science at the Federal University of Rio Grande do Sul (UFRGS), 2008. Currently is full professor and researcher at the Caxias do Sul University (UCS). She has experience in Computer Science with emphasis in the following topics: Artificial Intelligence applied to Education, Computers in Education, Recommender Systems and Distance Education. Dr. Boff is member of Brazilian Computer Society and she has publications at IEEE Journals and Proceedings, such as ITEE, ICTAI, ICALT, and Computers and Education.

Eliseo Berni Reategui has a PhD degree in Computer Science from the University of London, England, and MS in Computer Science from Universidade Federal do Rio Grande do Sul, Brazil. After finishing his PhD and working in the industry for 5 years, Dr Reategui held a lecturer position at the University of Caxias do Sul, Brazil, for a few more years. Nowadays, he works as a lecturer and researcher at the Federal University of Rio Grande do Sul. His research interests are related to the use of computers in education, involving areas such as artificial intelligence and human-computer interaction.

Efficient Measurement of the User Experience of Interactive Products. How to use the User Experience Questionnaire (UEQ).Example: Spanish Language Version

Maria Rauschenberger, *MSP Medien-Systempartner, Germany*; Martin Schrepp, *SAP AG, Germany*
Manuel Pérez Cota, *University of Vigo, Spain*; Siegfried Olschner, *DATEV eG, Germany*
Jörg Thomaschewski, *University of Applied Science Emden/Leer, Germany*

***Abstract* — Developer, manager and user feedback is needed to optimize products. Besides the basic Software qualities – usability and user experience are important properties for improving your product.**

Usability is well known and can be tested with e.g. a usability test or an expert review. In contrast user experience describes the whole impact a product has on the end-user. The timeline goes from before, while and after the use of a product. We present a tool that allows you to evaluate the user experience of a product with little effort. Furthermore the tool is available in different languages and we are using the new Spanish Version. We show how this tool can be used for a continuous user experience assessment.

***Keywords* — Software Quality, User Experience, Questionnaire, Usability, Test, Development**

I. Introduction

IS your redesign of the website better than the old version? Has the development effort spent to increase user experience really paid off? If you want to answer such questions you need a quantitative method to measure user experience [1]. An efficient and inexpensive method to do such measurements is the usage of rigorously constructed and validated questionnaires.

The concept of user experience combines well-known aspects like efficiency and effectiveness with additional criteria like aesthetics, joy-of-use or attractiveness. The first group of criteria is often referred as pragmatic quality aspects [2], while the second group is called hedonic quality aspects. Another often-used terminology to distinguish both classes of quality criteria is usability goals versus user experience goals [3]. The dependency of pragmatic and hedonic quality is presented in Fig. 1.

One well investigated research question is the relationship of pragmatic and hedonic quality. Empirical evidence proves that products, which are perceived to show a high level of hedonic quality, are also perceived as easy to use [4], [5], [6]. These and similar observations cause some authors [7] to state that 'What is beautiful is usable'. In contrast other studies point out [8], [9] an opposite dependency. The perception of the aesthetic value of a user interface increased when the number of concrete usability problems decreased. Thus, in this study a 'What is usable is beautiful' effect was observed.

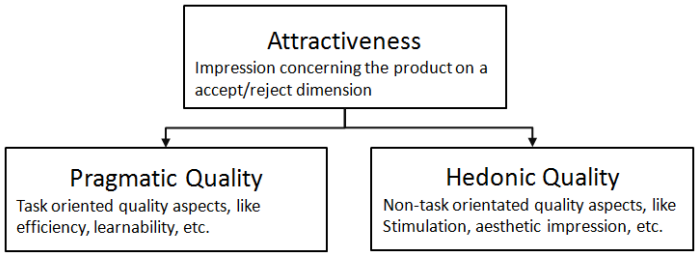


Fig. 1. Grouping of different quality attributes.

Why are perceived hedonic and pragmatic quality aspects associated? As possible explanation for this connection halo-effects [10], mediation by the mood of the user [11] or mediation by other variables [6] have been suggested. Since it is quite difficult to separate these effects experimentally [8] it is currently unclear which of these hypotheses are able to explain this effect.

These results indicate that it is necessary to consider both pragmatic and hedonic aspects if we want to measure how satisfied users are with a given product.

This is the underlined idea of constructing the User Experience Questionnaire (UEQ) [12], [13] that is described in this paper. In the context of the questionnaire user experience is understood as the overall impression of a user when he or she interacts with a product, i.e. covers both pragmatic and hedonic quality aspects.

The UEQ allows a quick assessment of the user experience for any interactive product. The scales of the questionnaire are designed to cover a comprehensive impression of user experience. The questionnaire format supports the user response to immediately express feelings, impressions, and attitudes that arise when they use a product.

If a new product is rolled out or if an existing product is evaluated the first time typical questions are '*Does the product create a positive user experience?'* or '*How do users feel about the product?'*. To answer such questions it is sufficient that a representative sample of users of the new product fill out the UEQ. 30 answers are usually enough to get a valid impression. For example, the answers can come from participants of a usability test or pilot users.

DOI: 10.9781/ijimai.2013.215

Another application is the continuous quality assessment of a software product within a development process [14]. In this approach a measurement with the UEQ is collected with each new version of the software. Thus, we can directly see if new versions bring an improvement in user experience if the scale values for the six scales of the UEQ increase with the new version (for an example on the concrete implementation of such a process, see [14]). An application of the UEQ in the process of idea and innovation management is described in [15].

User experience is not only a snapshot of the present usage a product has. It is an entire impression a product makes on the user. Even more, the user's judgement starts before touching and using a new product. In addition the change of impression carries on during and after the usage [1]. The UEQ is able to present the distinct results over time for the result analysis.

The UEQ is a semantic differential. For such questionnaires it is especially important that users see the items in their native language. So far the UEQ was available in German, English, French and Italian. We present in this paper the Spanish language version of the questionnaire.

We describe in the following how the UEQ was constructed and validated. In addition, the structure of the questionnaire and the meaning of the subscales are explained. We then show, how the UEQ should be applied in a company and how the results can be analyzed. Besides, the DATEV eG a big business software company is presenting their design process with the UEQ. Finally, we describe the creation of the Spanish language version of the UEQ.

II. Construction and Validation of the User Experience Questionnaire (UEQ)

The items and scales of the UEQ were created by a data analytical approach. First, a set of 229 potential items was built as a result of several brainstorming sessions with usability experts. Second, this set was reduced to an 80 items raw version by an expert evaluation. Third, the eighty items raw-version of the questionnaire was used in several studies focusing on the quality of interactive products, including e. g. a statistics software package, cell phone address book, online-collaboration software, or business software. In total the data of 153 participants were collected for the initial data set. Finally, the scales and the items representing each scale were extracted from the data by factor analysis (principal components, varimax rotation). Six factors resulted from this analysis. Details concerning the process can be found in [12], [13].

The reliability (i.e. the scales are consistent) and validity (i.e. the scales do really measure what they intend to measure) of the UEQ scales was investigated in several studies (in 11 usability tests with a total number of 144 participants and an online survey with 722 participants). A review of all available studies showed that reliability (Cronbach's Alpha was used for an estimation of internal consistency) of the scales was sufficiently high. In addition, the validity of the scales was investigated in a number of studies [12], [13], [14]. Results indicate good construct validity.

III. Structure of the Questionnaire

The user experience questionnaire contains 6 scales with 26 items in total:

1) **Attractiveness**: General impression towards the product. Do users like or dislike the product? This scale is a pure valence dimension. *Items: annoying / enjoyable, good / bad, unlikable / pleasing, unpleasant / pleasant, attractive / unattractive, friendly / unfriendly*
2) **Efficiency**: Is it possible to use the product fast and efficient? Does the user interface looks organized? *Items: fast / slow, inefficient / efficient, impractical / practical, organized / cluttered*
3) **Perspicuity**: Is it easy to understand how to use the product? Is it easy to get familiar with the product? *Items: not understandable / understandable, easy to learn / difficult to learn, complicated / easy, clear / confusing*
4) **Dependability**: Does the user feel in control of the interaction? Is the interaction with the product secure and predicable? *Items: unpredictable / predictable, obstructive / supportive, secure / not secure, meets expectations / does not meet expectations*
5) **Stimulation**: Is it interesting and exciting to use the product? Does the user feel motivated to further use the product? *Items: valuable / inferior, boring / exiting, not interesting / interesting, motivating / demotivating*
6) **Novelty**: Is the design of the product innovative and creative? Does the product grab users attention? *Items: creative / dull, inventive / conventional, usual / leading edge, conservative / innovative*

The dependency of the UEQ scale is presented in Fig. 2.

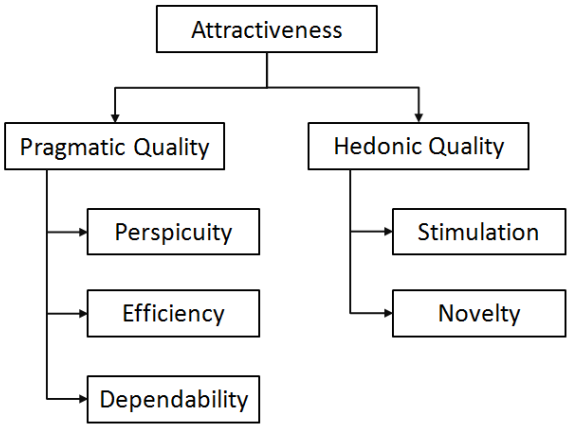


Fig. 2. Scale structure of the UEQ questionnaire.

For the specific questionnaire the order of the items and their orientation (starting with the positive or the antonym statement) is randomized. The specific English questionnaire is shown in Fig. 3 and the Spanish questionnaire is shown in Fig. 7.

	1	...	7		
annoying	○	...	○	enjoyable	1
not understandable	○	...	○	understandable	2
creative	○	...	○	dull	3
easy to learn	○	...	○	difficult to learn	4
valuable	○	...	○	inferior	5
boring	○	...	○	exciting	6
not interesting	○	...	○	interesting	7
unpredictable	○	...	○	predictable	8
fast	○	...	○	slow	9
inventive	○	...	○	conventional	10
obstructive	○	...	○	supportive	11
good	○	...	○	bad	12
complicated	○	...	○	easy	13
unlikable	○	...	○	pleasing	14
usual	○	...	○	leading edge	15
unpleasant	○	...	○	pleasant	16
secure	○	...	○	not secure	17
motivating	○	...	○	demotivating	18
meets expectations	○	...	○	does not meet expectations	19
inefficient	○	...	○	efficient	20
clear	○	...	○	confusing	21
impractical	○	...	○	practical	22
organized	○	...	○	cluttered	23
attractive	○	...	○	unattractive	24
friendly	○	...	○	unfriendly	25
conservative	○	...	○	innovative	26

Fig. 3. English version of the UEQ

IV. How to Apply the Questionnaire

For a successful application of the UEQ the acceptance of following two groups are needed: Users and Managers. To achieve a high user acceptance of the UEQ you should take following points into account:

- Background and benefits of the method should be clear to the user
- A personal contact should be available for the responders
- The time interval between repeated measurements should be long enough

To achieve acceptance by product managers consider to:

- Provide help during the interpretation of the UEQ factor values. Define your range of good, medium and bad and explain the theoretical background
- Combine old and new UEQ values in one picture and show the changes during the development in order to increase the intelligibility of the measurements
- Search for other user feedback that supports the interpretation of the UEQ outcome and integrate it into your report
- Enhance the UEQ results with concrete enhancement suggestions based on user experience expertise and use this as a base for further discussions about the next development goals

V. Analyzing Result

After collecting the answers from the users a three step analysis as presented can follow. To reduce the effort for data analysis an MS Excel file is created, doing all the necessary calculations. Only the raw data of the questionnaire results have to be entered into the tool. The tool then calculates the scale values, creates a bar chart to visualize the results and calculates some basic statistical indicators necessary for an interpretation of the data, for example confidence intervals for the scales. Fig. 4 presents an example of a result and Fig. 45 shows an example of a comparison of two product versions.

A. Verifying the validation

The first step is to confirm the Cronbach's Alpha data, which describes the consistency of the items of the scales (i.e. if all items in the scale measure the same quality). It is calculated automatically for each study in the excel sheet which can be downloaded from www.ueq-online.org.

If the Alpha value for a scale is small this is an indication that some of the items in this scale are possibly misinterpreted or interpreted in a direction that does not reflect their intention in the context of the UEQ. In this case it is questionable if this specific scale can be interpreted for the final result.

There are two well-known effects that can cause a small value of the Alpha-Coefficient for a scale. First, it is possible that the context in which the questionnaire is applied yields to a misinterpretation of some items in the scale. For example, in a study with informatics students the item 'secure/not secure' was referred from the users to the security (i.e. absence of malware or spyware) of the web-service and not to the dependability of the interaction.

Second, a scale may be simply irrelevant in the context in which the questionnaire is applied. Thus, the participants may have problems to interpret the items of the scale properly, which lowers the correlations between the items of the scale and thus decreases the Alpha-Coefficient.

If the alpha coefficient is higher or equal than 0,7 the scales show high consistency, i.e. all items in a scale measure the same aspect and it is unlikely that one of the items is misinterpreted in the given context.

But it can also happen that all items in a scale are influenced by a context specific effect, i.e. one of the scales differs highly from the other scales due to a special target group.

In a study with 20 participants the scale novelty had low results caused by a target group with different age. The VoIP-Software Skype was evaluated. The younger group had no enthusiasm about the technology, because they had known it for a long time. It was not exciting anymore. Elsewise the older group did not know Skype or any similar product. It was their first contact with this technology and they found it very fascinating. The consequence was that one group perceived Skype very stimulating and the other not.

After exanimating the Alpha value next step is the interpreting of the overall result as descripted in Chapter B.

B. Intepretate the overall result

The items are scaled from -3 to +3. Thus, -3 represents the most negative answer, 0 a neutral answer, and +3 the most positive answer. When analyzed the following aspect should be considered. Scale values above +1 indicate a positive impression of the users concerning this scale, values below -1 a negative impression. Due to well-known answer effects, like the avoidance of extremes, observed scales means are in general in the range of -2 to +2. More extreme values are rarely observed, so a value near +2 represents a very positive near optimal impression of participants.

Fig. 4 shows an example for an overall result for a product. The graphic is automatically generated by the data analysis sheet (Excel) that can be downloaded together with the questionnaire.

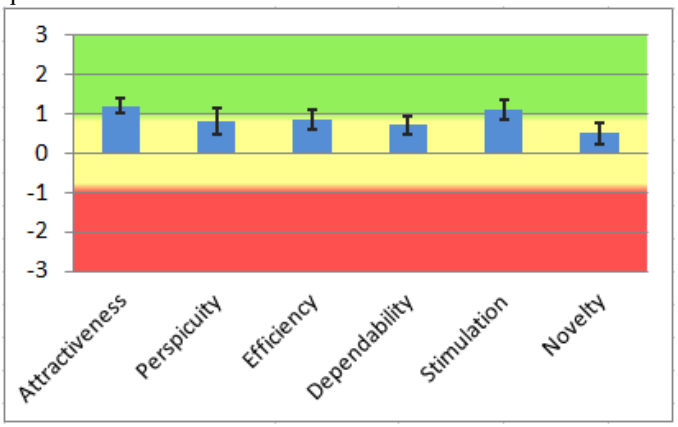


Fig. 4. Example of an overview result.

Thus, this particular product created a slightly positive impression concerning Attractiveness and Stimulation, but is judged neutral concerning the other 4 scales. The error bars represent the 5% confidence intervals for the scale means, i.e. the probability that the true value of the scale mean lies outside this interval is less than 5%. The width of the error bars depend on the number of respondents and on the level of agreement between the respondents. Thus, the more the participants that filled out the questionnaire agree concerning their evaluation of the product the smaller are typically the width of the error bars. Thus, if there are many respondents to the questionnaire and the error bars are still wide, this can be an indication that there are different sub-groups of participants with quite opposite options about the product.

Two different products or product versions can thus easily be compared concerning their user experience by comparing the scale means. See Fig. 5 for a comparison of two product versions concerning the observed scale means.

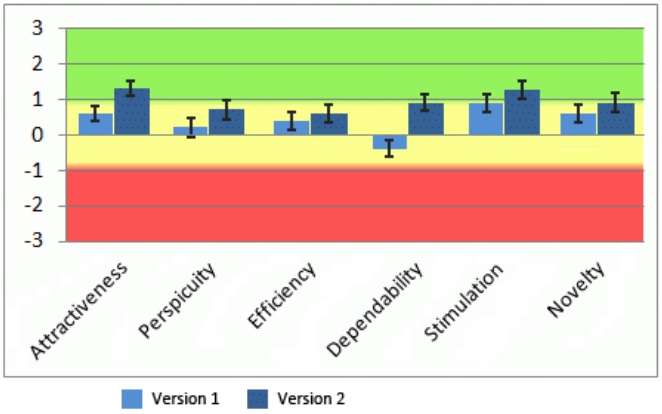


Fig. 5. Example of a comparison of two product versions concerning the UEQ scales.

In this example version 2 is much better concerning Attractiveness, Perspicuity Efficiency and Dependability. Concerning the hedonic scales Stimulation and Novelty both versions seems to be comparable.

To find out if the difference concerning the scale values is significant on the 5% level (or any other level you choose) it is necessary to apply a statistical test that compares the scale means (for ex. a t-test). It is not sufficient to check if the error bars do not overlap. If they do not overlap it can be concluded that the difference is significant at 5% level. But the opposite is not true. The error bars can overlap and the difference may still be significant!

The scales can be grouped into three categories. Atractiveness is a pure valence dimension. The scales efficiency, perspicuity and dependability describe the pragmatic quality of the product. The scales stimulation and novelty describe the hedonic quality of the product.

C. Analyzing the results of the individual items

After the overview the details have to be examined. First if you have two software versions with the UEQ results the items results are placed opposite each other. Items with extreme differences give a hint which areas have been improved or not. These way product versions can be compared easily and exact with one another. Also the detail analyzing shows, which areas should be improved for the next release (See Fig. 6). If it is the first product release see if some items show extreme results compared to other in the same UEQ results.

While analyzing each item the target group could give hints about what caused the significant distinction. Therefore the basic demographic data has to be collected with the UEQ results as well.

The UEQ exists in different languages which are tested reliably. Nevertheless, because of the complexity of language, it is also possible that translation deviance the results.

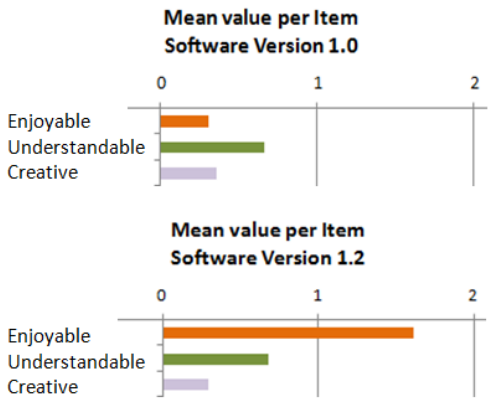


Fig. 6. Example for the detail analyzing of the results from the UEQ-Excel-Sheet (a specimen of the first three items)

VI. Application in the Company DATEV eG

This part presents an example how the UEQ is applied for benchmarking in a big business software company. A general impression of a process is presented in [16].

A. About DATEV eG

The cooperative DATEV eG, Nuremberg (Germany), is a software company and IT service provider for tax consultants, auditors and lawyers as well as their clients. Roughly 5800 employees produce more than 220 applications and provide service for about 39800 cooperative members.

B. Usage of UEQ within a defined Design Process

The concept of user centered design is meanwhile part of the official DATEV eG software development model and the UEQ is an integral component among other UCD methods like classical usability testing, focus groups, persona development and heuristic evaluation. The questionnaire is used to get user feedback at different development stages and all UEQ data are collected in one database.

C. Scenarios of use

One major goal is to perform a regular standardized survey with our users in consultant companies and enterprises. The challenge here is the integration into software release plans and market research activities. The UEQ is currently used successfully in three scenarios:

- Evaluation of new beta versions by selected beta testers
- Assessment of released software by randomly selected users
- At the end of a classic usability test to evaluate a new prototype

In the last scenario it is not the primary goal to get an accurate assessment, but the outcome will give an orientation whether the new software design will bring a significant improvement compared to the DATEV eG benchmark and previous measurements for the tested application. Of course one must be cautious, the tasks in a laboratory test do not represent the entire application and the demonstrated improvements in some parts will perhaps have no effect on the overall user experience of the complete application.

A current project is the test of the combination of online questionnaire and focus group. The outcome of the online-UEQ should be the base for questions in asynchronous online focus groups. Another example how to use the UEQ is described in an article concerning user experience for business software [16].

Because of the special form of the UEQ it is important that participants fill out the questionnaire in their natural language. Thus, it is for companies that use the UEQ on multi-national level important to have language versions of the questionnaire available.

VII. Creation of a Spanish Language Version

First, the German version of the UEQ was translated into Spanish by a native speaker and a bilingual person. After that the Spanish version had been retranslated into German. If the words turned out to match the original words the translation was declared to be successful. Otherwise the process was repeated until all words matched. To demand a one-to-one translation from one language into another is not entirely possible. The reason for that are the different meanings of one word, which make it difficult to find synonym in any language.

The translator was open minded and didn't know the questionnaire before. For more information see [17].

	1	...	7		
desagradable	○	...	○	agradable	1
no entendible	○	...	○	entendible	2
creativo	○	...	○	sin imaginación	3
fácil de aprender	○	...	○	difícil de aprender	4
valioso	○	...	○	de poco valor	5
aburrido	○	...	○	emocionante	6
no interesante	○	...	○	interesante	7
impredecible	○	...	○	predecible	8
rápido	○	...	○	lento	9
original	○	...	○	convencional	10
obstructivo	○	...	○	impulsor	11
bueno	○	...	○	malo	12
complicado	○	...	○	fácil	13
repelente	○	...	○	atractivo	14
convencional	○	...	○	novedoso	15
incómodo	○	...	○	cómodo	16
seguro	○	...	○	inseguro	17
activante	○	...	○	adormecedor	18
cubre expectativas	○	...	○	no cubre expectativas	19
ineficiente	○	...	○	eficiente	20
claro	○	...	○	confuso	21
no pragmático	○	...	○	pragmático	22
ordenado	○	...	○	desordenado	23
atractivo	○	...	○	feo	24
simpático	○	...	○	antipático	25
conservador	○	...	○	innovador	26

Fig. 7. Spanish version of the UEQ

VIII. Validation Studies for the Spanish Version

The Spanish Language version of the UEQ is already validated in two bigger studies.

In the first study 94 students evaluated the user experience of the Amazon web-shop (www.amazon.es). The scale means and confidence intervals are shown in Fig. 8.

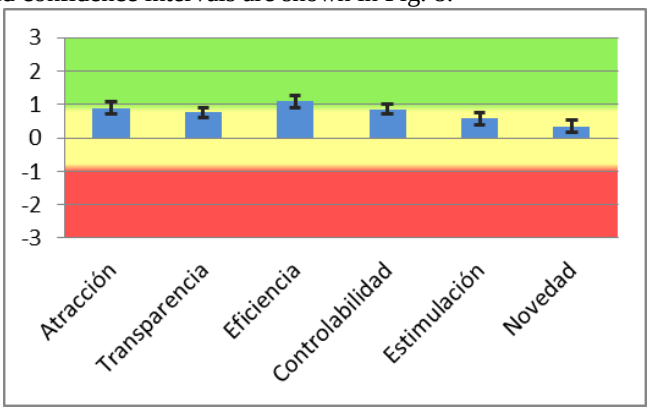


Fig. 8. Result for Amazon web-shop.

Thus, overall the participants had a slightly positive or neutral impression concerning the user experience of the Amazon web-shop. The impression concerning the pragmatic quality (Perspicuity, Efficiency and Dependability) is clearly higher than the impression concerning the hedonic quality (Stimulation, Novelty).

An analysis of the Cronbach Alpha coefficient showed that the single scales showed high consistency values (Attractiveness: 0.85, Perspicuity: 0.59, Efficiency: 0.74, Dependability: 0.48, Stimulation: 0.75, Novelty: 0.64). This is an indicator that the scales are sufficiently consistent.

In a second study 95 students evaluated the user experience of Skype. Again scale means and confidence intervals are shown in Fig. 9.

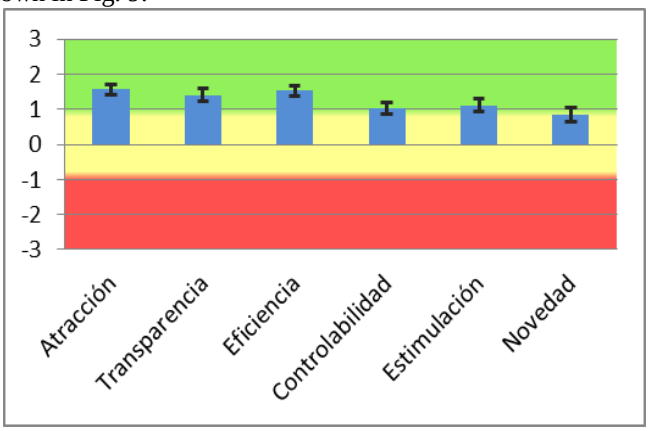


Fig. 9. Result for Skype.

The impression concerning the Skype user experience is quite positive. Again pragmatic quality is judged better than hedonic quality aspects. If we compare these evaluations to the results for the Amazon web-shop we clearly see that Skype creates a better user experience.

As in the first study alpha coefficient for the scales shows high values (Attractiveness: 0.83, Perspicuity: 0.71, Efficiency: 0.72, Dependability: 0.55, Stimulation: 0.78, Novelty: 0.71) again indicating sufficient scale consistency.

Of course further studies are necessary to finally judge if the psychometric properties of the Spanish version are identical to the existing and well-evaluated German and English version. But these first results are positive.

IX. Availability

The UEQ questionnaire can be used free of charge. The questionnaire itself, a data analysis tool and literature describing the construction of the questionnaire can be downloaded from www.ueq-online.org. The questionnaire and the analysis tool are available in several languages. Currently German, English, French, Italian and the Spanish version are available. It is worked on a Portuguese Version as well.

X. Summary

We described the construction, the result analyzing and the validation studies of the Spanish language version of the User Experience Questionnaire. This questionnaire allows a fast evaluation of the user experience of interactive products. It measures not only usability aspects like efficiency, perspicuity and dependability, but also user experience aspects like stimulation or originality.

Since the UEQ has the form of a semantic differential, it is quite important that participants can rate a product in their natural language. Thus, the new language version allows the application of the UEQ in Spanish speaking target groups.

The first available validation studies suggest that the scale quality of the Spanish version is sufficient to apply the questionnaire in projects to collect feedback about user impressions.

References

[1] DIN EN 9241-210, 2011-01, Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems, Beuth, Berlin, 2011.

[2] Hassenzahl, M., "The effect of perceived hedonic quality on product appealingness", Internatinal Journal of Human-Computer Interaction, 13, pp. 481-499, 2001.

[3] Preece, J., Rogers, Y., Sharp, H., "Interaction design: Beyond human-computer interaction". Wiley, New York , 2002.

[4] Kurosu, M., Kashimura, K., "Apparent usability vs. inherent usability: experimental analysis of the determinants of the apparent usability". In: Conference Companion of human factors in computing systems, pp. 292--293. Denver, Colorado, 1995.

[5] Tractinsky, N., "Aesthetics and Apparent Usability: Empirical Assessing Cultural and Methodological Issues". CHI'97 Electronic Publications http://www.acm.org/sigchi/chi97/proceedings/paper/nt.htm 1997.

[6] Hassenzahl, M., Monk, A., "The influence of perceived usability from beauty". Human-Computer Interaction, 25(3), pp. 235-260, 2010.

[7] Tractinsky, N., Katz, A.S., Ikar, D., "What is beautiful is usable", Interacting with Computers, 13, pp.127--145, 2000.

[8] Ilmberger, W.; Schrepp, M. , Held, T., "What kind of cognitive process causes the relationship between aesthetics and usability". In: Holzinger, A. (Ed.): USAB 2008, LNCS 5298, pp. 43-54, 2008.

[9] Tuch, A.N., Roth, S.P., Hornbaek, K., Opwis, K. & Bargas-Avila, J., "Is beautiful really usable? Toward understanding the relation between usability, aesthetics, and affect in HCI". Computers in Human Behavior 28 (5), pp. 596–1607, 2012.

[10] Hartmann, J., Sutcliffe, A., DeAngeli, A., "Towards a theory of user judgement of aesthetics and user interface quality". ACM Transactions on Computer-Human Interaction, 15, 2008.

[11] Norman, D., "Emotional design. Why we love (or hate) everyday things". New York: Basic books, 2004.

[12] Laugwitz, B.; Held, T., Schrepp, M. "Construction and evaluation of a user experience questionnaire". In: Holzinger, A. (Ed.): USAB 2008, LNCS 5298, pp. 63-76, 2008.

[13] Laugwitz, B.; Schrepp, M., Held, T. "Konstruktion eines Fragebogens zur Messung der User Experience von Softwareprodukten [Construction of a questionnaire to measure the user experience of software products] ". In: A.M. Heinecke, H. Paul (Eds.): Mensch & Computer 2006 – Mensch und Computer im Strukturwandel. Oldenbourg Verlag, pp. 125 – 134, 2006.

[14] Laugwitz, B.; Schubert, U.; Ilmberger, W.; Tamm, N.; Held, T., Schrepp, M., "Subjektive Benutzerzufriedenheit quantitativ erfassen: Erfahrungen mit dem User Experience Questionnaire UEQ [Quantitative measurement of subjective user satisfaction: Experiences with the user experience questionnaire UEQ]". In: Brau, H.; Diefenbach, S.; Hassenzahl, M.; Kohler, K.; Koller, F.; Peissner, M.; Petrovic, K.; Thielsch, M.;Ulrich, D., Zimmermann, D. (Eds.); Usability Professionals 2009, pp. 220 – 225, 2009.

[15] Winter, D., Pietschmann, J., "UX in den frühen Phasen des Innovationsprozesses. User Experience von Anfang an bedenken[UX in the early phases of the innovation process. Consider User Experience from scratch.]". In: Brau, H.; Lehmann, A.; Petrovic, K.; Schroeder, M. (Eds.); Usability Professionals 2012 , pp. 274-279, 2012.

[16] Rauschenberger, M., Hinderks, A.; Thomaschewski, J., "Benutzererlebnis bei Unternehmenssoftware: Ein Praxisbericht über die Umsetzung attraktiver Unternehmenssoftware [User experencing of Busnisssoftware: A practical report how attractive Busnisssoftware could be developt.]". In: Brau, H; Lehmann, A.; Petrovic, K.; Schroeder, M.C. (Eds.); Usability Professionals 2011, pp. 158 - 163, 2011.

[17] Rauschenberger, M., Schrepp, M.; Olschner, S.; Thomaschewski, J.; M. P. Cota, "Measurement of user experience: A Spanish Language Version of the User Experience Questionaire (UEQ) ". In: Information Systems and Technologies (CISTI), Vol. 1, pp. 471-476, 2012

Maria Rauschenberger has a Bachelor of Engineering in Media technology and is studying the Master of Science in Media Informatics from the University of Applied Science Emden/Leer. Furthermore is she working as a Project manager for MSP Medien-Systempartner, Germany. Currently, she pursues the finals of Master degree and taking part of the regional user experience group Bremen (UPA). She is involved in research activities dealing with interaction design, process optimizing, usability and user experience since 2011.

Dr. Martin Schrepp works since 1994 as user interface designer for SAP AG. He finished in 1990 his studies in Mathematics with a Diploma from the University of Heidelberg (Germany). In addition, he received in 1993 a PhD in Psychology (also from the University of Heidelberg). His research interests are the application of psychological theories to improve the design of software interfaces, the application of Design for All principles to increase accessibility of business software, measurement of usability and user experience and the development of general methods for data analysis. He published several papers in these research fields.

Prof. Dr. Manuel Pérez Cota is Professor and Researcher at the state-owned university Universidade de Vigo (UVIGO) in Vigo, Spain. He is graduated with honors in Electrical Engineering (Universidad La Salle) and Electronics and Communications Engineering (Universidad Nacional Autónoma de México – UNAM, 1980), a PhD with honors in Industrial Engineering (Universidad de Santiago de Compostela, 1990). He is the director of the international research group SI1-GEAC (http://cuautla.uvigo.es/si1-geac/). He was the first director and developer of the Informatics (Computer Science School of the University of Vigo) and he was director, also, of the Informatics Department. He widely collaborates in different Master and PhD programs in Spain, Portugal, Germany, Argentina and Bolivia, and he has been supervisor in several PhDs, with some others in progress. He is part in different European and International projects. He has published quite extensively and has a lot of publications (including books, book chapters, Scientific Citation Index journal articles, and international journal articles, as well as publications in refereed conference proceedings). He is member of different international committees and associations (ACM, IEEE, AISTI, AIPO, ANALCT).

Prof. Dr. Jörg Thomaschewski. He became full professor at the University of Applied Sciences Emden/Leer, Germany from September 2000. He received the Doctor degree in Physics from the University of Oldenburg, Germany, in 1996. His research interests are in the fields "Internet Applications" focusing on human computer interaction, e-learning and software engineering. He is author of various online modules, e.g. "Human Computer Communication", which is used in the Virtual University (Online) at six university sites. He has wide experience in usability training, analysis and consulting.

Dr. Siegfried Olschner is a senior member of the User Experience and Design team of DATEV eG, Nuremberg, Germany. His principal activities are dialogue design, workflow design and general user experience consulting for colleagues from the development departments. He also deals with the research and development of standardized test methods that are used as parts within the DATEV User Centered Design Process. One example is the improvement of the UEQ.

Engineering Education through eLearning technology in Spain

Juan Carlos Fernández Rodríguez[1], José Javier Rainer Granados[1], Fernando Miralles Muñoz[2]
[1]*Bureau Veritas Centro Universitario, Madrid, Spain.*
[2]*Universidad San Pablo CEU, Madrid, Spain*

***Abstract* — eLearning kind of education is stirring up all the disciplines in the academic circles, especially since it provides an access to educational areas that are uneasy and traditionally in-person, such as Engineering. Even though it had an outbreak in some of the most prestigious American universities, eLearning has being a reality in Spain for some years now, changing educational and teaching habits. To ensure a proper education is not an easy task with it comes to engineering fields, therefore this article shows an update on the works developed on this issue and the technologies they used. In this report it is given a perspective of the intimate relationship between the eLearning method of learning and the studies of Engineering in Spain, through the TIC development and the current educational legislation. In this regard, teaching examples are given on several subjects of different engineering studies, emphasizing the good results obtained in the abovementioned experiences. Below here is a evaluation on the results obtained in the analyzed studies.**

***Keywords* — eLearning, engineering, competencies, evaluation.**

I. Introduction

ROYAL Decree 1393/2007 establishes the new regulation on official university educations. Through this decree, the Spanish university system is integrated in the European Higher Education Area (EHEA). In accordance to this new model, there has been a major swift, going from a teaching model based in the transmission of knowledge directly from the professor, to a model based on acquiring different competencies and the alumni learning process. Engineering studies, both from a grade and postgraduate point of view, have not been an exception.

Concisely, a competency, in the engineering environment, refers to the activities the engineer must be able to perform, using the information, skills and tools required to accomplish his professional practice [1]. This way, an education based on competencies is referred basically to expected and visible performances, meaning analysis, synthesis and evaluation.

The necessary learning process is not oblivious to the present technological changes, the use of TIC has lead to a big leap in evolution of distance learning, creating the eLearning method of education. eLearning provides the opportunity to create virtual learning environments focused on the student, as EHEA demands. The aim of these virtual learning environments is to move the emphasis from the teaching to the learning [2].

These scenarios are characterized by being basically interactive, efficient, distributed and easily accessible. According Khan, quoted by Boneu [3], an ELearning scenario must consider eight facts: instructional design, educational model, technology, interface development, evaluation, management, support and use ethics. The platform in which the eLearning is developed is the software server responsible for the users management, courses management and communication services. These platforms are not isolated systems, given that they can rely on tools developed by third parties or on integrations developed by the designers or administrators.

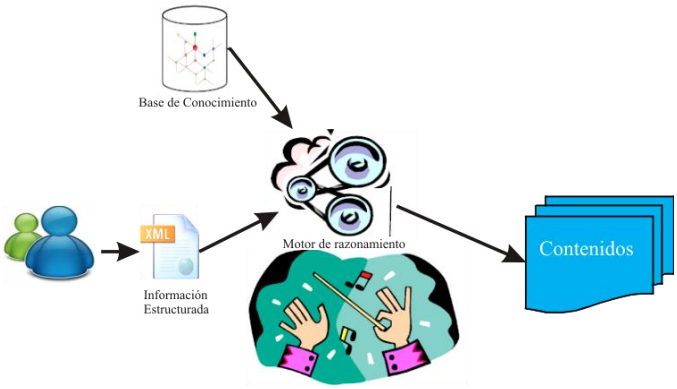


Fig.1. General Diagram of Contents Generation

eLearning can enhance some of the competencies that an engineer must develop, such as: identification, formulation and resolution of problems, the ability to design experiments or the ability to apply math, among others.

The specifications development, standard and tools regarding cognitive systems, ontologies and semantic maps are also helping in the formative systems enhacement.

In essence, the systems are based in showing some contents after receiving some request from the user. To this effect, the knowledge basis provides all the available information on this issue and some restrictions will select the definitive contents to show. Work on Intelligent Educational Systems (IESs) is traditionally divided into two main paradigms [4]: Intelligent Tutoring Systems (ITSs) [5] [6] and Adaptive Hypermedia Systems (AHSs) [7] [8].

DOI: 10.9781/ijimai.2013.216

According Caravantes and Galán [9], the essential part of the teaching-learning model is the educational domain, which is responsible for the stimuli generated toward the student to properly modify his cognitive state. All intelligent educational systems incorporate an educational domain that determines its scope and the process effectiveness. Usually the educational domain is divided into two parts: one represents the target knowledge of the learning process and is called domain model, and the other defines rules or procedures governing the process and is called the pedagogical/adaptive/operational model.

Teachers use the educational knowledge of a particular domain together with a meta-knowledge that encodes pedagogical skills for process controlling. This pedagogical knowledge represents instructional principles such as positive reinforcement, variability, action, etc., that facilitate the proper selection and sequencing of contents. It is a type of procedural knowledge based on Pedagogical Regulators (PRs) that controls the process, reading information from the emotional, characteristic and instructional domains. Common IESs implement an explicit pedagogical knowledge just as expert systems do using PRs called logical rules. However, there are other valid implicit representations such as neural networks, in which the PRs are called nodes, artificial neurons, and so on.

Learning content specifications like ADL-SCORM [10] or IMS-CC [11] are based on independent and interchangeable objects or packages that encode one or more ways to teach-learn something.

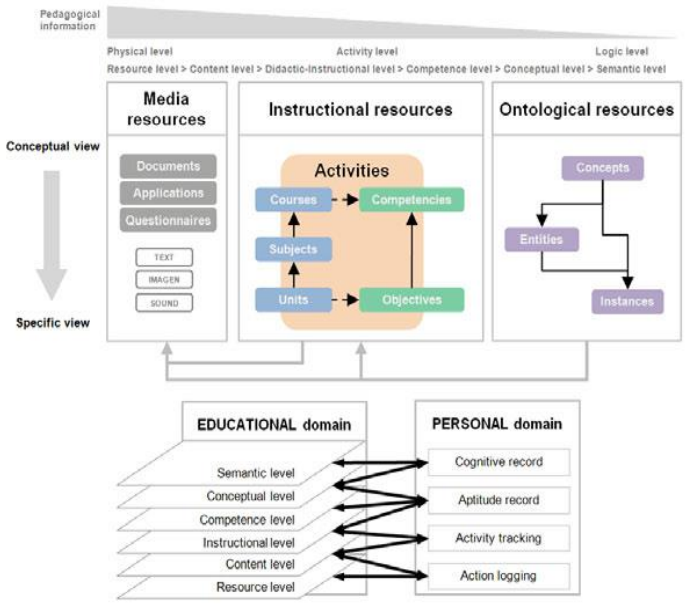


Fig.2. Levels of the educational knowledge representation [9].

Distance learning has usually made available to the student a set of continuous learning resources (documents, presentations, videos, animations and simulations) and discrete resources (texts, graphics, links and images) [12]. Resources are located at the lower material level of the educational knowledge structure to interact with the student to transfer knowledge. Educational resources are increasingly specified by soft links [13] that allow them to be searched, filtered and selected from large and dynamic repositories using metadata such as type (exercise, questionnaire, diagram, graph, table, text...), format, language, difficulty, etc. (see IEEE-LOM).

In this article, we will go through some of the experiences performed in the engineering area from the eLearning point of view, since just from the beginning eLearning and engineering have got along very well. Due to this connection, engineering learning becomes a higher quality learning, therefore it provides engineering a wider coverage regarding its education and extension [14].

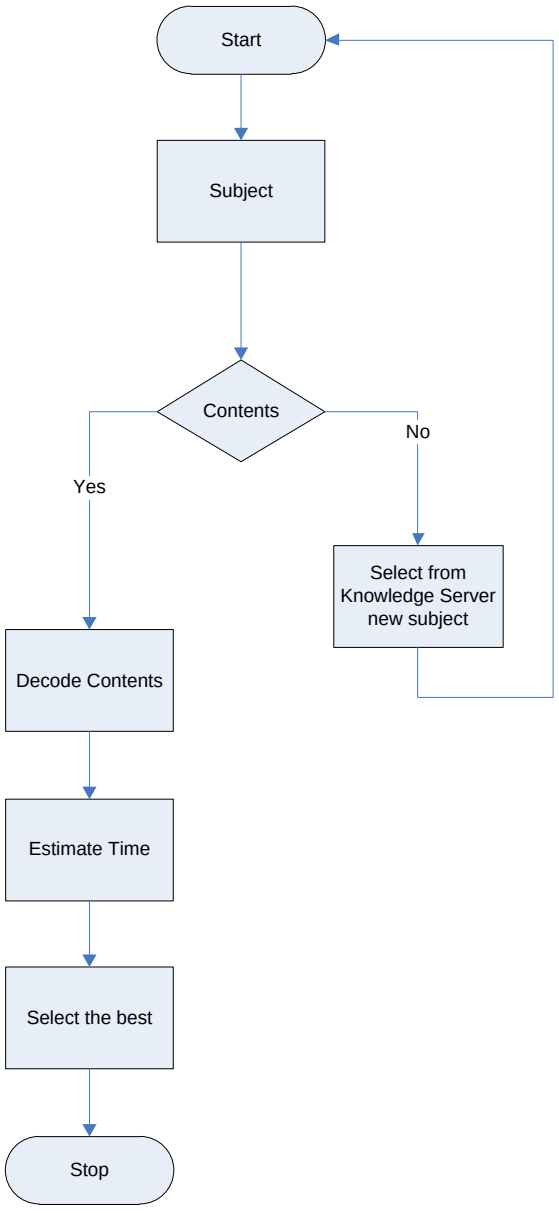


Fig. 3. General Diagram of Contents Management.

II. ELEARNING AND ENGINEERING

Engineering education has always been liked to attendance, especially because the help of a tutor easies the comprehension of the concepts and the assimilation of contents. It has been required the correct performance of the tools that integrate

eLearning education in order to make sense out of engineering eLearning education. Therefore, the first works are somehow recent, as shown in some of the most important works, summarized below.

One of the first approaches in literature we can find was in 2000. In that year, there were already many American universities that would offer a wide range of courses with the eLearning teaching method [15], proving that, right from the start, eLearning was meant to become one of the strongest lines of high education.

Databases knowledge is one of the paramount subjects in the high courses of computer engineering. Back in 2006, and through the use of eLearning platforms, it was achieved not only to increase the motivations but to improve the academic results of the alumni in the Technical Engineering in Computer Managements and Systems in Gerona University [16]. The most relevant achievement of this study was the tool that auto corrects the exercises related to the subject. This tool allowed the professor to acknowledge the learning level of the students and realize their possible deficiencies. Also, there are significant gains for the students, since the passing rate increased around a 9%. This increase might be caused by a sizeable boost of students that have taken on tutorial classes on the course.

In courses like "Databases" and "Business and Commercial Management Techniques" the auto correcting system had a great acceptance both within alumni and teachers, saving the later a significant work load: having to correct several exercises [17].

A very important aspect of this learning method is all that is related with the use and development of tools that will allow us to evaluate the assignments of the students in every subject automatically. In other words, a field that refers to the application of the different information and communication technologies in order to establish learning platforms and environments to make to evaluation of the alumni easier.

This way, the Computer Based Assessment (CBA) represents one of the eLearning technologies, distinguished by the automation of every teaching/learning feature of the student, integrating advanced functions held jointly, i.e., correcting databases structure diagrams and related consults [18]. With this CBA, there is an interaction between teacher and student all along the evaluation process. In this process, the turn in of the course exercises, its correction and the feedback generated is performed by the system automatically.

Put this into practice, CBA has been used in educational subjects of the science/technical area, mainly in test or multiple choice questions, with a limited scenarios feedback. The reason of these limitations is the disaffection toward these kinds of questions, since they can be perceived as distant and only acceptable in some low cognitive leveled tasks, which require fixed answers.

With this tool (CBA), the author scored better academic results and a higher satisfaction on the environment developed in students of the Databases subject of the Technical Computer Management Engineer of the Gerona University.

One of the tools that contributes to improve the learning process in multimedia environments are the Learning Objects, which are digital resources based on the Web, whose main feature is they can be used and reused to support the learning. A Learning Object (LO) is the minimal learning unit that has sense by itself, regardless of the context. Furthermore, it is a unit with an educational content and reusable in its digital format. Since this format must be standardized, it can be reused in different platforms, ensuring its reusability and usefulness [19]. Among these digital resources you can incorporate images, videos or pre-recorded audios, small text pieces, animations, little Web apps or even entire Web sites combining texts, images and other communication media.

One good example of the reusability and usefulness of the Learning Objects is the one performed on the Industrial Engineer School of Vigo University. Through this action, it was expected to encourage and motivate the engineering learning of the alumni, even though it can be conducted in any other discipline [20]. The authors designed a learning pill (very short course the summarizes the main concepts of the subject) in order to use it as a tool the students must use before the master lectures, but due to the specific features of the learning pills, they can be reached and used at any time and any place, thanks to their free access. The learning pill was effectively used by the whole alumni and it was verified that it was used worldwide.

The incorporation of multimedia technology and eLearning education has been applied equally to the Project Direction learning [21]. These authors instituted the eLearning technology in the "Projects" subject on the Industrial Engineering degree at Escuela Politécnica Superior de Ingeniería in Gijón. In order to achieve this, in this occasion it was selected a platform different from the traditional ones, such as Moodle. In this case, Microsoft Sharepoint was used, a specific project management tool, favoring the approach of the teacher's work towards the professional environment. The experience, besides enhancing the cooperative work between students, contributed to preparing the alumni towards the analysis, evaluation and alternatives choice, leading them to an educated decision making, something of paramount importance in their future engineering works. Nowadays, the eLearning method is widely spread in Projects Direction and Management and programs exist in most of the Spanish universities.

In engineering teaching, on-line methodologies have also been applied for evaluation and formation in generic competences. In the Technical Computer Engineering School of the Universidad Politécnica de Madrid it has been developed a competence approach for the "Telecommunications History" and "Telecommunications Policy for the Information Society" subjects, which were given through the Moodle platform. It is worth mentioning that the chosen teaching method was bLearning which, unlike eLearning, it combines online contents with in-person given contents. The results, evaluated through an online survey to the alumni, were highly satisfactory, pointing out several technical aspects, mainly the simplicity in the use of the Moodle platform [22].

The eLearning app can be used in any kind of subject, being the Logic subject of Computer Engineering both a particular

and relevant example of application of this educational method [23]. In the Computer Engineering degree at UOC (Universidad Abierta de Cataluña), there is this Basic Logics subject. Said subject has very low performance and very high abandon ratings and by the application of eLearning with the idea of improving these rates in mind, eLearning was implemented. The results scored may be considered as discreet, being the reduction in abandoning rate just a 5%, same as academic results improvement. It is important to point out that the measurement tools used to calculate these rates does not allow ensuring the reason of these changes is only the eLEarning based tool. The web sites of the tools have been developed using PHP for the server program, as well as HTML, CSS and Javascript. It is also worth mentioning that for the data lodging and abidance a MySQL database was used, granting an efficient session management.

The engineering laboratory practices have not been an exception on the eLearning educational method [24]. Particularly [25], the alumni were offered in their programming practices to use robots with eLearning method. In said practices, the student could book the robot for 30 minutes, and download and execute the program, being able to visualize it through an IP camera. Meanwhile, the web server was programmed to offer use statistics.

This project constantly seeks for the flexibility, therefore searches for a system design that is easily adaptable to any kind of experiment where the student used to utilize hardly reachable hardware, along making the high costs of an industrial robot profitable.

The main thing is that the learning curve is ascendant, even if it is the student who adjusts the intensity, leading to cases where the performance of the student behaves as shown in the following figure. The figure initials correspond to IN: I, initial phase of the course and E, end of the course.

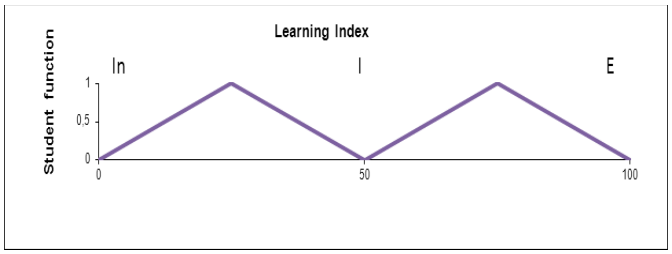


Fig. 4. Learning intensity along the course.

In a similar line [26], two traditional Systems and Automatic Engineering laboratories (three tanks control system and robot programming) of different universities have been transformed into two virtual and remote laboratories. This way, students can perform their lab exercises virtually, organized by an automatic booking system. The experience evaluation, gathered by a survey to the alumni, pointed out the achievement of an elevated satisfaction rate, existing among the students a high rate of motivation to use the technology in their learning activities.

III. Conclusion

In every experience analyzed in the present article is possible to detect the efficiency of the eLearning educational method in all the Engineering areas. In the abovementioned projects and their corresponding evaluations, is obvious the good results scored and the satisfaction of the students. The alumni feedback is very important for every learning method, through a satisfaction survey or test, but especially in this kind of method, since the student judges only the knowledge he has acquired, withdrawing any personal relationship he may have with his tutor, which can get even closer that in-person classes.

When it comes to eLearning education, the behavior of the system relies mainly in two factors: architecture and contents. In every cases previously analyzed, both factors are combined. The potential of these systems increases remarkably when Artificial Intelligence techniques are integrated, leading the system to a teacher-cognitive system combination that ensures not only the contents transmission but the proper choice of content within the wide knowledge database, supported by teaching techniques that permit an effective contents transmission as well as a certain knowledge acquisition.

Therefore it seems to be proof enough to state the validity of eLearning when it regards to engineering, as The American Accreditation Board for Engineering and Technology (ABET) claims [27]. Accordingly, not only the students score better results, but there is a possible chance to widely extend the education while cutting expenses, two important features when it comes to practice subjects to perform in laboratories.

However, there is a certain lack of more specific experiences, with a better educational design, and a lack of experiences that affect the paramount subjects of every engineering university title, both degree and masters. Besides, it would be convenient to expand the evaluation to teachers and apply a more severe criteria to evaluate de possible academic scores improvement, establishing the criteria required in order to ensure the students the correct acquisition of competencies.

Also it would be desirable a future research on the feelings of the student towards his presence and involvement in certain subjects and the improvements this might bring, being this an aspect where is pivotal the development and application of every multimedia technology available.

References

[1] M. Duque, “Competencias, aprendizaje activo e indagación: un caso práctico en ingeniería”. Revista de Educación en Ingeniería, 2, 2006, pp.7-18.

[2] J.E. Silva, “Diseño y moderación de entornos virtuales de aprendizaje”. Barcelona: UOC, 2011.

[3] J.M. Boneu, “Plataformas abiertas de e-learning para el soporte de contenidos educativos abiertos”. Revista de Universidad y Sociedad del Conocimiento (RUSC), (4) 1, 2007, pp. 36-47.

[4] A. Nicholas, B. Martin. “Merging adaptive hypermedia and intelligent tutoring systems using knowledge spaces”. Adaptive Hypermedia (AH2008), 2008, pp. 426–430

[5] T. Murray. “Authoring Intelligent Tutoring Systems: An analysis of the state of the art”. International Journal of Artificial Intelligence in Education, 98, 1999, pp. 98-129.

[6] E. Wenger. "Artificial Intelligence and Tutoring Systems: Computational and Cognitive approaches to the communication of knowledge". San Francisco, USA, Morgan Kaufman, 1987.

[7] P. Dolog. "Designing Adaptive Web Applications". In SOFSEM 2008: Theory and Practice of Computer Science, pp. 23-33.

[8] P. Brusilovsky. "Methods and techniques of adaptative hypermedia". User Modeling and User Adapted Interaction, Special issue on adaptive hypertext and hypermedia, 6 (2-3), 1996,pp-87-129.

[9] A. Caravantes y R. Galán. "Generic educational knowledge for adaptive and cognuitive systems". Educational Technology & Society, 14 (3), 2011, pp-252-266.

[10] ADL. "Sharable Content Object Reference Model (SCORM)". Retrieved June December, 11, 2012, from http://www.adlnet.gov/Technologies/scorm/default.aspx.2004.

[11] IMS. "Common Cartridge Specications". Retrieved December, 12, 2012, from http://www.imsglobal.org/cc. 2008.

[12] K. Verbert, E. Duval.. "Towards a Global Component Architecture for Learning Objects: A Comparative Analysis of
Learning Object Content Models". World Conference on Educational Multimedia, Hypermedia and Telecommunications 2004, pp. 202–208.

[13] P Brusilovsky, N. Henze, N." Open corpus adaptive educational hypermedia. Adaptive Web 2007" (AW2007). pp. 671–696. Heidelberg: Springer.

[14] J. Bourne, D. Harris and F. Mayadas, (Janaury, 2005) "Online Engineering Education: Learning Anywhere, Anytime". Journal of Engineering Education. (94) 1, pp. 131-146 Avaible http://digitalcommons.olin.edu/facpub_2005/1.

[15] R. Ubell. "Engineers turn to e-Learning". IEEE Spectrum, 37, 2000, pp. 59-63.

[16] J. Soler, F. Prados, I. Boada and J. Pocho. "Utilización de una plataforma eLearning en la docencia de bases de datos". In Actas de las XII Jornadas de la Enseñanza Universitaria de la Informática, JENUI. Bilbao, 2006, pp. 581-586.

[17] J. Soler, I. Boada, F. Prados, J. Poch and F. Fabregat. "Experiencia docente en diseño de bases de datos con la ayuda de herramientas de eLearning" In Actas de las XII Jornadas de la Enseñanza Universitaria de la Informática, JENUI. Barcelona, pp. 63-70

[18] J. Soler. "Entorno virtual para el aprendizaje y la evaluación automática en bases de datos". Doctoral Thesis. University of Gerona, Spain, 2010.

[19] J.C. Fernández, J.J. Rainer and F. Miralles. "Aportaciones al diseño pedagógico de entornos tecnológicos eLearning". Madrid, Lulu, 2012.

[20] R.Maceiras, A. Cancela and V. Goyanes. "Aplicación de Nuevas Tecnologías en la Docencia Universitaria". Formación Universitaria (3) 1, 2010, pp. 21-26.

[21] J.M. Mesa, J.V. Álvarez, J.M. Villanueva and F.J. de Cos. "Actualización de métodos de enseñanza-aprendizaje en asignaturas de Dirección de Proyectos de Ingeniería". Formación Universitaria (1) 4, 2008, pp. 23-28.

[22] R. Herradón, J. Blanco, A. Pérez and J.A. Sánchez.). "Experiencias y metodologías "bLearning" para la formación y evaluación en competencias genéricas en Ingeniería". La Cuestión Universitaria, 3, 2009, pp. 33-45.

[23] M. Hertas, E. Mor and A. Guerrero-Roldan. "Herramienta de apoyo para el aprendizaje a distancia de la Lógica en Ingeniería Informática". Revista de Educación a Distancia, 24, 2010, pp. 1-10.

[24] L. Gomes and J. García-Zubia. "Advances on remote laboratories and eLearning experiences". Bilbao, Universidad de Deusto, 2007.

[25] Rosado, Muñoz-Marí and Magdalena. "Herramienta eLearning para la programación de robots mediante entorno web". Revista d´Innovació Educativa, 1, 2008, pp. 45-48.

[26] H. Vargas, J. Sánchez, C.A. Jara, F.A. Candelas, O. Reinoso and J.L. Díez. "Docencia en automática. Aplicación de las TIC a la realización de actividades prácticas a través de Internet". Revista Iberoamericana de Automática e Informática Industrial, (7) 1, 2010, pp. 35-45.

[27] ABET, Acreditation Board for Engineering and Technology. "Acreditation policy and procedure manual". USA, Engineering Accreditation Commission Publication, 2009.

Juan Carlos Fernandez Rodríguez. He has been awarded a doctorate degree (PhD) from the Universidad Complutense de Madrid UCM (Spain) and he is a bachelor of Philosophy and Sciences of Education from Universidad de Oviedo. At present he is Head of Studies and teacher in Bureau Veritas Centro Universitario. Actually he is investigating eLearning topics. He has great experience in the private sector and in the educational sector. He is author of several books and scientific articles. In addition, he has taken part as a speaker in congresses and symposiums.

Javier Rainer (PhD). Received a PhD in Robotics and Automation from the Universidad Politécnica de Madrid, UPM (Spain) in 2011. He is Director of Research and Director of Engineering Area at Bureau Veritas Centro Universitario, and researcher at the Intelligent Control Group of the UPM. He is Industrial Engineer, from Málaga University. He has been co-author of published papers in several journals and recognized international conferences and symposiums. He received the best paper Award in IARIA Cognitive 2010**.**

Fernando Miralles Muñoz.. Received a PhD in Phsicology from the Universidad Complutense de Madrid, UCM (Spain) in 2001. He is awarded honoraty doctoradtes from the Bircham International University since 2011. Currently combines his research with teaching at the University San Pablo CEU in Madrid

A multi-agent system model to integrate Virtual Learning Environments and Intelligent Tutoring Systems

Giuffra P., Cecilia E., Silveira Ricardo A.
Departament of Informatic and Statistics
Federal University of Santa Catarina, Brasil

***Abstract* — Virtual learning environments (VLEs) are used in distance learning and classroom teaching as teachers and students support tools in the teaching–learning process, where teachers can provide material, activities and assessments for students. However, this process is done in the same way for all the students, regardless of their differences in performance and behavior in the environment. The purpose of this work is to develop an agent-based intelligent learning environment model inspired by intelligent tutoring to provide adaptability to distributed VLEs, using Moodle as a case study and taking into account students' performance on tasks and activities proposed by the teacher, as well as monitoring his/her study material access.**

***Keywords* — Agents, Intelligent Tutoring Systems, Multi-Agent System, Virtual Learning Environments.**

I. INTRODUCTION

THE number of students with computer access has increased substantially in recent years. A qualitative change in the teaching–learning process happens when we can integrate within an innovative view all technologies, including telematics, audiovisual, textual, oral and physical [11]. The fact of students seeking information on the computer converts them into more active students. "There are activities that can be performed with the computer, forcing the student to seek information, process it and use it to solve problems, allowing the understanding of what makes and the construction of their own knowledge" [18].

Virtual learning usually offer the same learning experience, during the course, for all students, without considering their specific needs. The problem is that the students are treated as if they always had the same profile, the same goals and the same knowledge [12].

In order to provide adaptability to learning environments, according to student characteristics, and to allow a greater interactivity degree between the learning environment and the users, the research points to the use of resources provided by artificial intelligence (AI) and in particular the use of multi-agent system-based architectures [15].

In agreement with this emerges the motivation of this research: to enhance the teaching–learning process in virtual learning environments using artificial intelligence techniques to make the environments more adaptive and more interactive. This paper proposes the use of agent-based intelligent tutoring systems architectures to get personalized teaching strategies, taking into account the student profile and his/her performance, exploring their skills as best as possible, in order to have better and more effective learning in an intelligent learning environment.

This paper is structured as follows: the second section presents the theoretical reference and related works, the third section presents the definition of the model, the fourth section presents an explanation about the model implementation and the last section presents the conclusions.

II. BACKGROUND

Virtual learning environments are technological tools and resources using cyberspace to lead content and enable pedagogical mediation through the interaction between the educational process actors [14]. The use of these environments has increased significantly by the strong possibility of interaction between student and teacher that they offer, and by easy access anywhere and anytime. The virtual learning environments provide tools for interaction, such as forums and chats, and enable the provision of materials by teachers about the content of the course.

For Dillenbourg [6], virtual learning environments are not only restricted to distance learning. Web-based education is often associated with distance learning; however, in practice it is also widely used to support classroom learning. The author also comments that the difference between these two types of education is disappearing. Many students in distance courses do not live far from school, but have time constraints. Often they work. In addition, there are courses that combine distance and presence, which makes for more robust learning environments.

Virtual learning environments, at first, were used primarily in distance learning; now they also serve as support in classroom courses, as a teacher's tool to provide materials, to review tasks, to keep track of the students on course (activity

DOI: 10.9781/ijimai.2013.217

logs) and also to evaluate them. For students, the environment facilitates the delivery of tasks, the obtaining of materials for the course and the monitoring of their evaluation.

Virtual learning environments can be enhanced with artificial intelligence techniques, using intelligent agents, having intelligent learning environments as result. An agent is an abstraction of something that can perceive its environment through sensors and can act upon that environment through actuators [16]. Intelligent agents are those that have at least the following characteristics: autonomy, reactivity, proactivity and social ability[21].

In practice, systems with only one agent are not common. The most common are the cases of agents that inhabit an environment containing other agents. There are two major types of multi-agent systems: reactive and cognitive. The reactive acts under a stimulus-response scheme; the cognitive has, in general, few agents because each agent is a complex and computationally heavy system[4].

A rational agent is one who chooses his/her actions according to their own interests, given the belief that he/she has about the world. The Belief, Desire, Intention (BDI) model recognizes the importance of beliefs, desires and intentions in rational actions [20].

The BDI model represents a cognitive architecture based on mental states, and has its origin in the human practical reasoning model. An architecture based on the BDI model represents its internal processes through the mental states: belief, desire and intention, and defines a control mechanism that selects in a rational way the course of actions [7].

In the context of this work, an agent is considered as an autonomous entity, able to make decisions, respond in a timely manner, pursue goals, interact with other agents, and has reasoning and character. This agent is a of type BDI, with beliefs, desires and intentions, and operates in a virtual learning environment as an intelligent tutor.

Intelligent tutoring systems (ITS) are complex systems involving several different types of expertise: subject knowledge, knowledge of the student's knowledge, and pedagogical knowledge, among others. According to Santos et al. [17], an ITS is characterized for incorporating AI techniques into a development project and acts as a helper in the teaching–learning process.

According to Conati [5], intelligent tutoring systems are an interdisciplinary field that investigates how elaborate educational systems provide adapted instructions to the needs of students, as many teachers do.

ITS research has been investigating how to make computer-based tutors more flexible, autonomous and adaptive to the needs of each student by giving them explicit knowledge of the relevant components of the teaching process and reasoning skills to convert this knowledge into intelligent behavior.

To Giraffa and Viccari [9], ITS developments consider a cooperative approach between student and system. The goal of ITS is to complement or replace a human tutor, with the advantage of monitoring the student in each learning step [13].

Research in intelligent tutoring systems is concerned about the construction of environments that enable more efficient learning. [8].

Intelligent tutoring systems offer flexibility in the presentation of material and have the major ability to respond to students' needs. They seek, in addition to teaching, learning relevant information about the student, providing an individualized learning. Intelligent tutoring systems have been shown to be highly effective in improving performance and motivation of students [10].

Intelligent tutoring systems in virtual learning environments potentiate the teaching–learning process, making the virtual environment into an intelligent learning environment. Intelligent learning environments use AI techniques to respond to students' needs, making that learning personalized [10].

According [15], the intelligent learning environment must build and update the student model in terms of what he/she already knows, which can vary significantly from one student to another.

Related Works

In order to know the current status of recent research about virtual learning environments and the use of intelligent agents as tutors in these environments, we performed a systematic literature review. Among them, there were three items which were most closely related to the purpose of this study.

"Approach to an Adaptive and Intelligent Learning Environment" [1], which proposes an agent-oriented approach for the design and implementation of an adaptive and smart component for a virtual learning environment. The adaptivity in the model is defined as a system's ability to create and, during the learning process, uniformly upgrade the curriculum that satisfies the student's needs. The proposed model has three parts that describe the main features of intelligent and adaptive component. First, the student chooses the courses based on his/her needs, the level of excellence that he/she wants to achieve, and his/her preference concerning the type of study material. Second, the system will decide how to act – for example, show the material to the student, based on the belief (student model) that the system has about it. In the last part it is decided when to propose an evaluation test for the student or any other activity that can evaluate any specific knowledge in relation to the curriculum. After completing the evaluation, the belief about the student is updated.

"Cluster Analysis in Personalized E-Learning" [22]: this is a proposed system architecture in which teaching techniques and appropriate layouts are set to groups of students with similar preferences, created by applying clustering techniques. Teaching materials and content can be adapted to the needs of each group and different learning paths can be created. New students fill out a questionnaire to determine their learning style and their choices of usability and, according to this, the appropriate group is chosen for them. The idea of the proposed solution is to divide the process into two steps: first, to look for groups of students with a great similarity and detect those isolates. In the second step the groups are mixed in larger

groups if necessary and the isolates are indicated. The objective of the experiment was to examine the performance of the proposed clustering technique for different student's data sets, depending on the choice of parameters.

"Supporting Cognitive Competence Development in Virtual Classrooms" [19]. The approach described in this article implements a mechanism to adaptively create self-assessment questionnaires in a Moodle environment. The Learning Management System (LMS) is capable of saving all online activities of the students in log files. This information can be used also to automatically generate intelligent feedback to the student. The questions are derived from an ontology of skills that is also used for indexing learning materials. The student traces through the learning materials used to determine the current state of "expected knowledge" or skills. The system includes two main agents: the goals manager agent, which guides the student in planning activities; and the content manager agent, which guides the student during the resources review. In this paper, an extension of the Moodle LMS – in which ontologies are used to structure the learning process by providing resources and generating questionnaires automatically for self-assessment of the students – is presented.

In the conducted research to analyze the state of the art, papers were found dealing with adaptability in virtual learning environments, taking into account the students' needs, learning styles, usability preferences and their activities report (log).

In the first related paper, adaptability is based on the preferences of students regarding the study material, where the agent provides this material according to the information that he/she has about the student's preferences. In the second paper, the proposal is to provide different layouts to the students, taking into consideration similar preferences with regard to learning styles and usability choices. In the last related paper, self-assessment questionnaires are used, adaptively created in the virtual learning environment Moodle, and ontologies are used to index the learning materials. This paper proposes to join data that can be obtained from the database (information of student performance and logs) that most virtual learning environments widely used usually have, in order to set the discipline in a personalized way for each student, with regards to the available material for the student as the activities proposed to him/her, exploring his/her skills and bypassing disabilities, always having a baseline with material and compulsory activities, and activities outside of this line divided into different levels of difficulty.

III. Definition of the Model

The aim of this work is to create an agent architecture and a knowledge base of these agents that compose an intelligent tutoring system with information obtained from the database of a teaching–learning virtual environment. For this, a case study is done based on the Moodle platform architecture, chosen because it is a platform widely used today, in addition to being consolidated from the standpoint of operation, and also to be formally used in the institution where the research is performed.

The classical model of ITS contains the pedagogical model, the student model, the domain base and the control. In the proposed model, two types of agents, called "Bedel" and "tutor" are used. The Bedel Agent and all their knowledge and interaction structure corresponds in the classical model of intelligent tutoring to the Pedagogical Model. The Tutor Agent and all their structure corresponds in the model of intelligent tutors to the Student Model. The content of the discipline, in turn, may be associated with the abstraction of the Domain Base (Fig. 1).

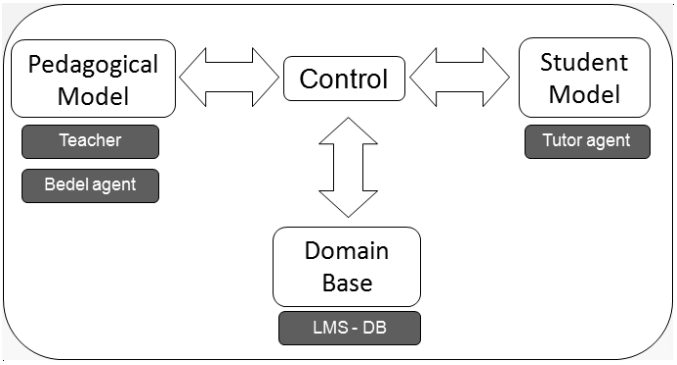


Fig. 1 Classical model with proposed model

The model of agents "Bedel" (agent of the discipline) and "Tutor" (agent of the student) is defined; they are connected with the learning virtual environment through the database. The interface in which the teacher sets the priority and levels of the resources and tasks in the environment is developed. The database is adapted with the creation of the table of grade profiles of the students and the table of dependencies of resources and activities, configured by the teacher by means of a Moodle block into his discipline, and is made the integration of the actions of agents with the virtual learning environment Moodle. The agents model (Fig. 2) shows the agents "Tutor" and "Bedel" their actions, messages and perceptions, as well as their connection to the database. The actor "teacher" is the figure of the discipline teacher who inserts the resources and activities into the learning environment, sets the type of profile and sequence, and this information is stored in the database, so the agent "Bedel" knows how to show them to students.

The database has information concerning the student, such as personal data, performance data and data from student interaction in the system. Every student interaction in the environment is saved in the base in the form of a log. Similarly, the student's performance in each of the activities and tasks is stored in the database and updated constantly every access and interaction of the student, providing rich material for the agent's performance. The agents share the information from the database. The "Tutor" updates the student profile and, if necessary, shows a message to him/her about his/her performance. The "Bedel" obtains from the database the configuration of resources and activities in the

discipline made by the teacher; it sets their preview and verifies if tasks were evaluated to send a message to the "Tutor" who, upon receiving the message, updates the data of the student profile.

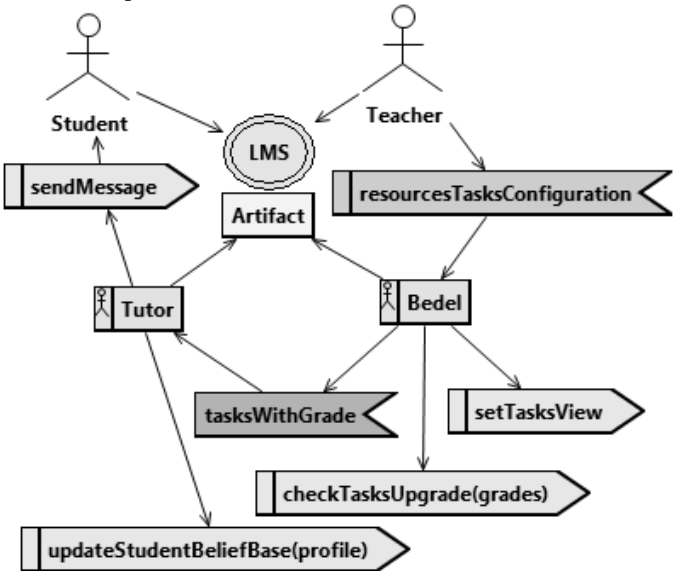


Fig. 2. Agent System Model

The model works as follows:

1) The teacher inserts the resources and creates tasks in the Moodle environment as usual. After that, he/she adds the tutor block in his/her discipline and configures the Bedel setting the first resource and activity (Fig. 3), the dependencies (Fig. 4) of activities and resources, and their level (basic, intermediate and advanced), also by means of the Moodle environment. In addition, the teacher has the option to choose resources that must be shown to all students in a general way. The first reading (resource) and the first activity are shown for all students; therefore the teacher needs to indicate which they are. This information is stored in the database, where some tables which are necessary for the model are added. With this information, the "Bedel" knows about resources and activities of the discipline and knows how the course should be developed for each type of student.
2) In the environment's database the grade_profile table (Fig. 5) is also added, which contains a numeric value, calculated with the grades of the activities performed by students and the access made by them in different files provided by the teacher. This table is updated each time a teacher updates the worksheet with the grades of some of the tasks that he/she provides to students.
3) The grade_profile average of all students is computed and students are separated by profile into groups – basic, intermediate or advanced – according to their grade_profile. Whoever has average grades is in the average profile, whoever has grades below average is in the basic profile, and those who have grades above average are in the advanced profile.
4) Tasks are provided independently for each student, according to their performance in the previous tasks, and their access to previous reading material. The availability of tasks and resources is made by the "Bedel" using the conditional access resource of Moodle.
5) Each time the Bedel calculates the grade_profile it updates the belief base of the current profile of the student, which can go from basic to intermediate or advanced and vice-versa during the time that the course is offered.

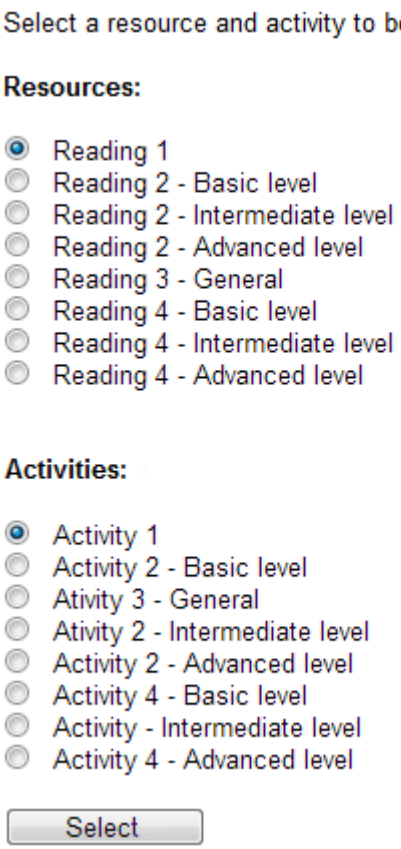


Fig. 3 First resource and activity

The calculation of grade_profile is done as follows: The student grade of the last activity assessed by the teacher and the grade of the student access on reading that is a pre-requisite for activity are summed. If the student accesses the reading, two points in the activity grade are added. If he/she does not access, four points are added. This difference is given for increasing the possibility to the student who does not access the reading, to have a higher grade_profile and then go to a higher level task than the profile that he/she would belong to if he/she had a lower grade, stimulating him/her to read before accomplishing future activities.

After this, the average value of the grade_profile field is computed for all students. The lowest value considered for the profile average is 6; if the profile average is lower, it automatically becomes 6. For the student to be at an average profile, his/her grade_profile must be between 5.5 and 6.5; if the student has a grade_profile less than 5.5 he/she is on a basic profile and if he/she has grade_profile greater than 6.5

he/she is on an advanced profile.

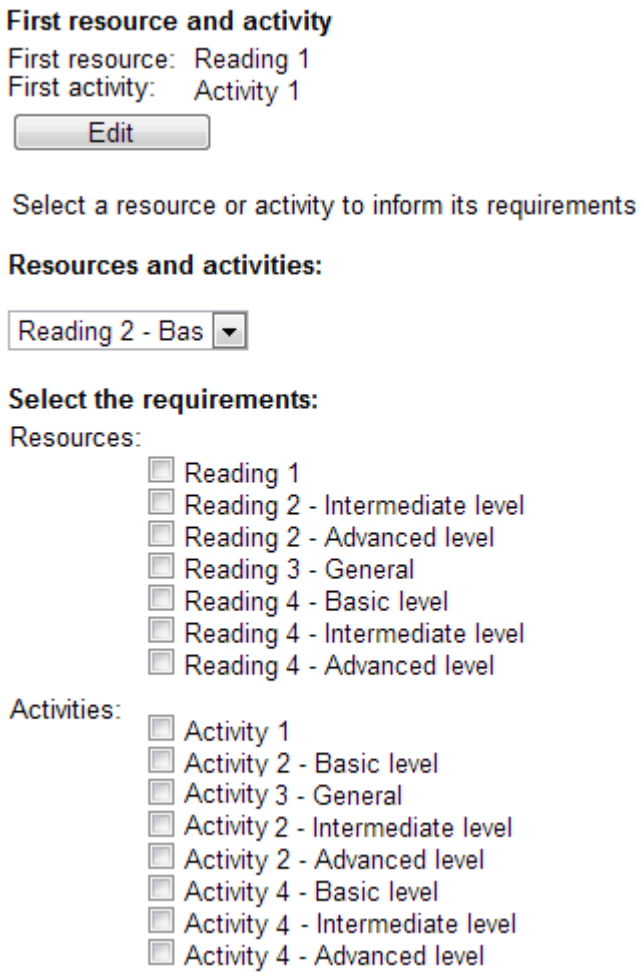


Fig. 4 Dependencies configuration

The maximum value considered for the profile average is 8. If the average profile is greater than that, it automatically becomes 8. For the student to be at average profile, his/her grade_profile must be between 7.5 and 8.5, if the student has a grade_profile less than 7.5 he/she is on the basic profile and if he/she has a grade_profile greater than 8.5 he/she is on the advanced profile.

The student belongs to the average profile if his/her grade_profile is 0.5 less or more than the average grade_profile in his/her class; for example, if the grade_profile average is 7.5, he/she will be in the intermediate profile if he/she has a grade_profile between 7 and 8. The student who has a higher grade with more than 0.5 of diference with the average will be in the advanced profile and the student who has a lower grade with more than 0.5 of diference will be in the basic profile.

The student will have access to the material of their profile (basic, intermediate, advanced), according the configuration of resources and tasks made by the teacher.

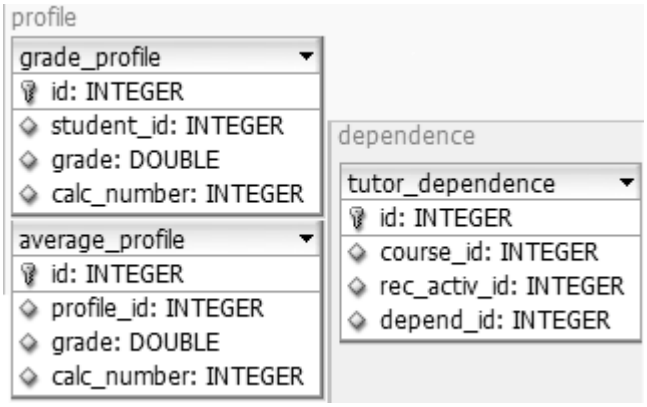


Fig. 5. Profile and dependence tables

IV. IMPLEMENTATION OF THE MODEL

The model integrates concepts of intelligent tutoring systems architectures with VLE's that have their use consolidated as Moodle, which are not adaptive for itself only, and can be potentiated with artificial intelligence techniques, resulting in intelligent learning environments which are shown to be adaptive and more suitable to the implementation of teaching defiant methodologies for the student.

The use of agents in the implementation of this model is important because of the agent's ability to adapt to environment changes, showing resources and activities to students in a personalized way, according to their performance in the discipline, and taking into account the teacher's initial settings.

For the agent implementation the Jason tool was used, which is an interpreter for an extended version of AgentSpeak, oriented agent programming language, implemented in Java. The basic idea of AgentSpeak is to define the know-how (knowledge about how to do things) of a program in the form of plans [3].

One of the most interesting aspects of AgentSpeak is that it is inspired by and based on a model of human behavior that was developed by philosophers. This model is called the Belief Desire Intention (BDI) model. The language interpreted by Jason is an extension of AgentSpeak, based on BDI architecture. A component of the agent architecture is a beliefs base and an example of what the interpreter does constantly, without being specifically programmed, is to perceive the environment and update the beliefs base accordingly with this [3].

The teaching–learning virtual environments are designed to enable the knowledge-building process. Different to conventional software, which seeks to facilitate the tasks achievement by user, learning environments incorporate the complexity to more flexible different forms of users (students), relations, to learn and to practice content, and to collaborate. These environments are used by students of various cognitive profiles [2].

The version of virtual learning environment Moodle used for this work is 2.2, where the task condition resource is

available, which allows the provision of content and activities with a restriction. This feature must be activated by the administrator of Moodle in the environment advanced settings, enabling the option "Enable tracking of completion" and "Enable conditional access." Moreover, in course settings, in "student progress" topic, the teacher must enable the completion tracking option.

With this feature enabled, tasks can be made available only to students who perform the pre-requisites set, which can be: a grade on a specific activity; the viewing of a resource; or his/her grade.

In this work, the availability of resources and activities is done taking into account the student's performance and his/her access in the system and is made available to the student depending on his/her grade_profile, computed according to his/her performance and participation in the discipline. The information between the agent system and the virtual learning environment are exchanged through the database of the learning environment which contains information about the pre-requisites and profiles of tasks and resources, defined by the teacher at a time to configure the Bedel. The development was made in 4 steps. On Fig. 6 we can see the Bedel sequence diagram, with the Artifact working like an interface between the database and the Bedel Agent.

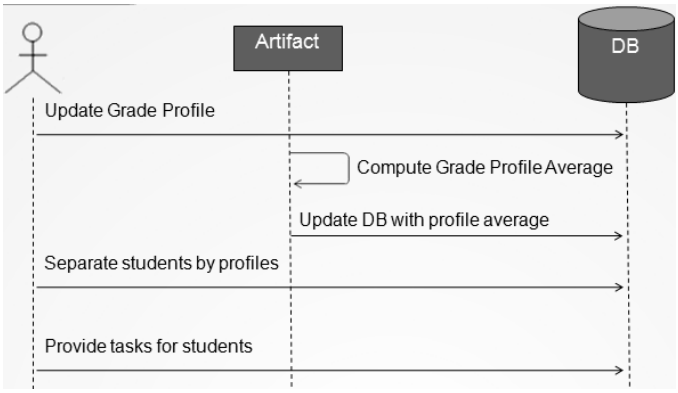


Fig. 6 Bedel Sequence Diagram

A. Development of the tutorblock in Moodle LMS.

It was developed a moodle block for the teacher to configure the agent. This block was created following the Moodle standard programming for creating blocks. The name was defined as "Tutor Block". This block have to be added by the teacher and is used by him/her to set the Bedel Agent, after the insertion of resources and activities in the environment.

After the teacher configures the agent through the block, he/she can see the dependency graph (Fig. 7) generated after setting all prerequisites. This graph shows the relations between the resources and activities setting by the teacher.

B. Development of the Bedel Agent. Programming student profile

In this step was implemented the code for the Bedel Agent to calculate de student profile, using the information available on database. The Bedel Agent access the database trough the Artifact and execute the plans it has for calculating this profile. Then, the Bedel Agent calculate the profile average of all the students and separate them into profiles (basic, intermediate and advanced).

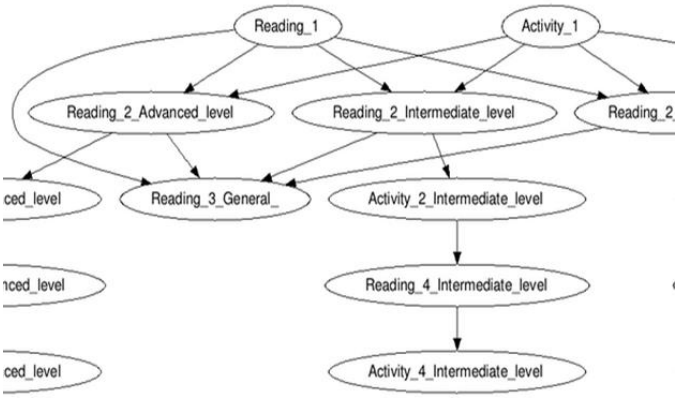


Fig. 7 Dependency graph

The Bedel Agent implementation has a definition file, where is specified the infrastructure, the Cartago environment used to program the artifacts (Fig. 8) and the Tutor and Bedel Agents.

C. Development of the code to display resources and activities

For the availability of resources and activities according to the student profile was created the tutor_profile_availability table in the database, to store the information of the minimum and maximum grades of the intermediate profile. Verifying this information the resources and activities can be provided, according to the students profile.

```
package artifact;

import java.sql.Connection;

//import cartago.OpFeedbackParam;

public class BD_Artifact extends Artifact {

    Connection conn;
    int idCourse = 2;
    String contextid;
    boolean counting;
    final static long TICK_TIME = 5000;//86400000; //24 horas em milisegundos
    ArrayList<String> id_grade_item_Tarefas_entregues;
    ArrayList<String> lista_tarefas_avaliadas;

    void init(int initialValue) {
        counting = false;
        // defineObsProperty("idCourse", 2);
    }

    ResultSet select(String string) throws SQLException, ClassNotFoundException {
        ResultSet result = null;
        try {
            Statement stm = conn.createStatement();
            result = stm.executeQuery(string);
        } catch (Exception e) {
            System.out.println(e);
        }
        return result;
    }
```

Fig. 8 Artifact Source Code

D. Development of the Tutor Agent to send messages to the students

It was implemented the Tutor Agent, that is responsible for

sending messages of encouragement to the student, taking into account his/her performance. This agent receives the student profile information from the Bedel Agent, updates the belief it has about this student and sends some message to him/her according to the student situation. If the student had a better performance than in the previous activity the Tutor Agent send a congratulation message (Fig. 9), if not, it sends a message to encourage him to be better the next time.

At this stage, also, the Bedel Agent is updated, inserting methods that enable communication between it and the Tutor Agent. Sending the student information to the Tutor Agent.

Fig. 9 Congratulation message (in Portuguese)

V. Conclusions

In this study is proposed a solution for virtual learning environments to assist teachers to provide activities and resources in a personalized way depending on the student's performance and his/her behavior in the discipline.

Students are assessed by their interaction in the discipline and the grades obtained in tasks, creating different profiles for groups of students with the same behavior. More advanced tasks are available for students who have improved performance, enabling more efficient learning, exploring students' skills, and maintaining a basic level for learning the discipline content.

Works related to virtual learning environments and adaptivity in general differentiate students by learning style – for example, a student who learns better with pictures than with reading lots of text. In this work students are distinguished by their performance, taking into account the grades obtained, and their participation (access) in the various resources available in the discipline, creating an adaptive environment that constantly updates the profile of students, and therefore, a student with a basic profile, at the end of the course may have an average profile. These profile changes can be studied and displayed to the teacher, in an extension of this model.

References

[1] Baziukaité, D.: Approach to an Adaptive and Intelligent Learning Environment. In: Elleithy, K., Sobh, T., Mahmood, A., Iskander, M., Karim, M. Advances in Computer, Information, And Systems Sciences, And Engineering, pp. 399–406. Springer Netherlands (2006)

[2] Boff, E.: Collaboration in Learning Intelligent Environments mediated by a Social Agent Probabilistic. In: Thesis (Ph.D.) Computer Science Course. Federal University of Rio Grande do Sul, Porto Alegre. (2008) (in Portuguese)

[3] Bordini, R.H., Hubner, J.F., Wooldridge, M.: Programming Multi-Agent Systems in AgentSpeak using Jason. Editora Wiley. England (2007)

[4] Bordini, R., Vieira, R., Moreira, A.F.: Fundamentos de Sistemas Multiagentes. In: Ferreira, C. E. (ed.). Jornada de Atualização em Informática (JAI'01), vol. 2, pp. 3–44. SBC, Fortaleza, Brasil (2001)

[5] Conati, C.: Intelligent Tutoring Systems: New Challenges and Directions. Paper Presented at the Proceedings of the 21st International Joint Conference on Artificial intelligence (2009)

[6] Dillenbourg, P.: Virtual Learning Environment . EUN Conference 2000. Workshop on Virtual Learning Environment (2000). Available in: http://tecfa.unige.ch/tecfa/publicat/dil-papers-2/Dil.7.5.18.pdf

[7] Fagundes, M.: An environment for development of BDI agents. Course Completion Work. Federal University of Pelotas (2004) Available in: http://www.inf.ufsc.br/~silveira/INE602200/Artigos/TCC_Moser.pdf. (in Portuguese)

[8] Frigo, L. B., Pozzebon, E., Bittencourt, G.. The Role of Intelligent Agents in Intelligent Tutoring Systems. In: Proceedings of the WCETE - World Congress on Engineering and Technology Education, São Paulo, Brasil, pp. 667-671. (2004). (in Portuguese)

[9] Giraffa, L.M.M., Viccari, R.M.: The Use of Agents Techniques on Intelligent Tutoring Systems. In: Computer Science SCCC '98. XVIII International Conference of the Chilean Society, pp. 76-83. IEEE Computer Society Washington, DC, USA (1998).

[10] Lima R.D., Rosatelli, M.C.: An intelligent tutoring system to a virtual environment for teaching and learning. In: IX Workshop de Informática na Escola, Campinas. Anais do XXIII Congresso da Sociedade Brasileira de Computação. (2003). (in Portuguese)

[11] Moran, J.M.: Computers in Education: Theory & Practice. Porto Alegre, vol. 3, n.1. UFRGS. Postgraduate Program in Computer oin Education, pp. 137–144 (2000). (in Portuguese)

[12] Mozzaquatro, P.M., Franciscato, F., Ribeiro, P.S., Medina, R.D.: Modeling a Framework for adaptation of Virtual Learning Environments Mobiles to different cognitive styles. vol 7, n 3, pp. 253-263. Renote - New Technologies in Education Journal. ISSN 1679-1916 (2009). (in Portuguese)

[13] Oliveira, C.L.V.: AutoExplC: intelligent tutoring system to aid the teaching of "C" languagebased on learning by self-explanation of examples. PUC, Campinas (2005). (in Portuguese)

[14] Pereira, A.T.C., Schmitt, V., Álvares, M.R.C.: Virtual Learing Environments. Culture Bookstore. (2007). Available in: http://www.livrariacultura.com.br/imagem/capitulo/2259532.pdf. (in Portuguese)

[15] Silveira, R. A. Ambientes inteligentes distribuídos de aprendizagem. Porto Alegre: CPGCC da UFRGS, (1998)

[16] Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice-Hall, Inc. New Jersey (2002)

[17] Santos, C.T., Frozza, R., Dahmer, A., Gaspary, L.P.: Doris - Pedagogical Agent in Intelligent Tutoring Systems. In: Cerri, S., Gouardères, G., Paraguaçu, F. Intelligent Tutoring Systems, pp. 91-104. Springer Berlin, Heidelberg (2002)

[18] Valente, J.A.: Computers in Education: conform or transform the school, pp. 42. (2009).

[19] Weinbrenner, S., Ulrich, H., Leal L., Montenegro, M., Vargas, W., Maldonado, L.: Supporting Cognitive Competence Development in Virtual Classroom In: 10th International Conference on Advanced Learning Technologies. IEEE Computer Society (2010)

[20] Wooldridge, M.: Reasoning about Rational Agents. The MIT Press. Cambridge, Massachussetts. London, England. (2000)

[21] Wooldridge, M.: An Introduction to Multiagent Systems, 2nd edition. John Wiley & Sons Ltd in Hoboken, NJ (2009)

[22] Zakrzewska, D.: Cluster Analysis in Personalized E-Learning Systems. Springer-Verlag, Berlin Heidelberg (2009)

Ms. Cecilia E. Giuffra P. Doctoral student of the postgraduate course in Engineering and Knowledge Management from the Federal University of Santa Catarina (UFSC). Master's (2013) and Bachelor's (2009) at Computer Science from the same University. Has experience in Computer Science, focusing on the following subjects: Artificial Intelligence, Learning Management Systems, Multiagent Systems, Computers in Education, Repositories, and Learning Objects. Has knowledge of the Moodle platform and had participation in teacher training and EaD staff.

Dr. Ricardo A. Silveira, PhD in Computer Science from Universidade Federal do Rio Grande do Sul (2001). Master's at Education from the Pontificial Catholic University of Rio Grande do Sul – PUC-RS (1992). Graduate at Electronic Engineering from PUC-RS from Departamento de Engenharia Elétrica (1980). Associate Professor of the Department of Informatic and Statistics (INE) on Federal University of Santa Catarina (UFSC). Has experience in Computer Science, focusing on Formal Languages and Automata, acting on the following subjects: Artificia Intelligence, Multiagent Systems, Computer in Education, Pedagogical Agents and Learning Objects.

GLOA: A New Job Scheduling Algorithm for Grid Computing

[1]Zahra Pooranian, [2]Mohammad Shojafar, [3]Jemal H. Abawajy, and [4]Mukesh Singhal
[1]Graduate School, *Dezful Islamic Azad University, Dezful, Iran*
[2]Dept.of Information Engineering, Electronic and Telecommunication (DIET), *''Sapienza'' University of Rome, Rome, Italy*
[3]School of Information Technology, *Deakin University, Geelong, Australia*
[4]Computer Science & Engineering, *University of California, Merced, USA*

***Abstract* — The purpose of grid computing is to produce a virtual supercomputer by using free resources available through widespread networks such as the Internet. This resource distribution, changes in resource availability, and an unreliable communication infrastructure pose a major challenge for efficient resource allocation. Because of the geographical spread of resources and their distributed management, grid scheduling is considered to be a NP-complete problem. It has been shown that evolutionary algorithms offer good performance for grid scheduling. This article uses a new evaluation (distributed) algorithm inspired by the effect of leaders in social groups, the group leaders' optimization algorithm (GLOA), to solve the problem of scheduling independent tasks in a grid computing system. Simulation results comparing GLOA with several other evaluation algorithms show that GLOA produces shorter makespans.**

***Keywords* — Artificial Intelligence, Distributed Computing, Grid Computing, Job Scheduling, Makespan.**

I. INTRODUCTION

NEW technology has taken communication to the field of grid computing. This allows personal computers (PCs) to participate in a global network when they are idle, and it allows large systems to utilize unused resources. Like the human brain, modern computers usually use only a small fraction of their potential and are often inactive while waiting for incoming data. When all the hardware resources of inactive computers are collected as an all-in-one computer, a powerful system emerges.

With the help of the Internet, grid computing has provided the ability to use hardware resources that belong to other systems. "Grid computing" may have different meanings for different people, but as a simple definition, grid computing is a system that allows us to connect to network resources and services and create a large powerful system that has the ability to perform very complex operations that a single computer cannot accomplish. That is, from the perspective of the users of grid systems, these operations can only be performed through these systems. As large-scale infrastructures for parallel and distributed computing systems, grid systems enable the virtualization of a wide range of resources, despite their significant heterogeneity [1].

Grid computing has many advantages for administrators and developers. For example, grid computing systems can run programs that require a large amount of memory and can make information easier to access. Grid computing can help large organizations and corporations that have made an enormous investment to take advantage of their systems. Thus, grid computing has attracted the attention of industrial managers and investors in companies that have become involved in grid computing, such as IBM, HP, Intel, and Sun [2].

By focusing on resource sharing and coordination, managing capabilities, and attaining high efficiency, grid computing has become an important component of the computer industry. However, it is still in the developmental stage, and several issues and challenges remain to be addressed [3].

Of these issues and challenges, resource scheduling in computational grids has an important role in improving the efficiency. The grid environment is very dynamic, with the number of resources, their availability, CPU loads, and the amount of unused memory constantly changing. In addition, different tasks have different characteristics that require different schedules. For instance, some tasks require high processing speeds and may require a great deal of coordination between their processes. Finally, one of the most important distinctive requirements of grid scheduling compared with other scheduling (such as scheduling clusters) is scalability.

With more applications looking for faster performance, makespan is the most important measurement that scheduling algorithms attempt to optimize. Makespan is the resource consumption time between the beginning of the first task and the completion of the last task in a job. The algorithm presented in this paper seeks to optimize makespan. Given the complexity and magnitude of the problem space, grid job scheduling is an NP-complete problem. Therefore, deterministic methods are not suitable for solving this problem. Although several deterministic algorithms such as min-min and max-min [4] have been proposed for grid job scheduling, it has been shown that heuristic algorithms provide better solutions. These algorithms include particle swarm

DOI: 10.9781/ijimai.2013.218

optimization (PSO)[5], genetic algorithms (GAs)[6], simulating annealing (SA)[7], tabu search (TS)[8], gravitational emulation local search(GELS)[9], ant colony optimization (ACO) [10], and recently Learning Automata (LA) [26]. Also, some researchers have proposed combinations of these algorithms, such as GA-SA[11], GA-TS[12], PSO-SA[13], GPSO[14], and GGA[15].

It is important that an optimization algorithm for optimization problems should converge to the optimal solution in a short period of time. The group leaders optimization algorithm (GLOA) [16] was inspired by the influence of leaders in social groups. The idea behind the algorithm is that the problem space is divided into several smaller parts (several groups), and each part is searched separately and in parallel to increase the optimization speed. Each separate space can be searched by its leader, who tries to find a solution by checking whether it is the closest member to the local and global minimum.

In this paper, we use GLOA for independent task/job scheduling in grid computing. In addition to the simplicity of its implementation, GLOA reduces optimization time. The remainder of this paper is organized as follows. Section II discusses related methods. Section III presents a general model for job/task scheduling. Section IV presents the GLOA method and modifies it based on our problem. Section V compares simulation results obtained with this algorithm and several other heuristic algorithms. Finally, the last section presents the conclusion of this study.

II. Related Work

In [17], the TS algorithm, which is a local search algorithm, is used for scheduling tasks in a grid system. In [18], the SA algorithm is used to solve the workflow scheduling problem in a computational grid. Simulation results show that this algorithm is highly efficient in a grid environment. The TS algorithm uses a perturbation scheme for pair changing.

In [19], the PSO algorithm is used for job scheduling with two heuristic algorithms, latest finish time (LFT) and best performance resource (BPR), used to decide task priorities in resource queues. In [20], the critical path genetic algorithm (CPGA) and task duplication genetic algorithm (TDGA) are proposed; they modify the standard GA to improve its efficiency. They add two greedy algorithms to the GA so that the wait times for tasks to start and ultimately the makespan can be reduced. The proposed algorithms consider dependent tasks, so that computation costs among resources are considered as well. Chromosomes are divided into two parts, and the graph under consideration is transformed into a chromosome that performs mapping and scheduling. The mapping part determines the processors on which tasks will execute, and the scheduling part determines the sequence of tasks for execution. In the representation of a chromosome, task priorities are considered by examining the graph.

The CPGA algorithm combines the modified critical path (MCP) algorithm [21] and a GA. The MCP algorithm first determines critical paths, and if the parent of tasks being executed on a processor is executing on another processor, these tasks are transported to the parent's processor to reduce the cost of transportation between processors.

The TDGA algorithm combines the duplication scheduling heuristic (DSH) algorithm [22] and a GA. This algorithm first sorts tasks in descending order and then repeats the parent task on all processors so that the children can execute earlier, because the transportation cost between processors becomes zero. By repeating the parent task, overload and communication delays are reduced and total execution time is minimized.

The resource fault occurrence history (RFOH) [23] algorithm is used for job scheduling fault-tolerant tasks in a computational grid. This method stores resource fault occurrence histories in a fault occurrence history table (FOHT) in the grid information server. Each row of the FOHT table represents are source and includes two columns. One column shows the failure occurrence history for the resource and the other shows the number of tasks executing on the resource. The broker uses information in this table in the GA when it schedules tasks. This reduces the possibility of selecting resources with more occurrences of failures.

The chaos-genetic algorithm [24] is a GA for solving the problem of dependent task/job scheduling. This algorithm uses two parameters, time and cost, to evaluate quality of service (QOS), and chaos variables are used rather than randomly producing the initial population. This combination of the advantages of GAs and chaos variables to search the search space inhibits premature convergence of the algorithm and produces solutions more quickly, with a faster convergence.

The integer genetic algorithm (IGA) [25] is a genetic algorithm for solving dependent task/job scheduling that simultaneously considers three QOS parameters: time, cost, and reliability. Since these parameters conflict with one another and cannot be simultaneously optimized—as improvement of one reduces the quality of another—weights are assigned to each parameter, either by the user or randomly. If the user provides the weighting, the parameter that is more important to the user is given more weight than the others.

III. Problem Description

The problem studied in this paper is independent task/job scheduling in grid computing. The proposed algorithm should be efficient in finding a solution that produces the minimum makespan. Thus, the problem is to assign a set of m input tasks ($T=T_1,T_2,...,T_m$) to n resources ($R=R_1,R_2,...,R_n$), with the minimum makespan.

IV. The GLOA algorithm

GLOA is an evolutionary algorithm that is inspired by the effect of leaders in social groups. The problem space is divided into different groups, and each group has its own leader. The members of each group don't necessarily have similar characteristics, and they have quite random values. The

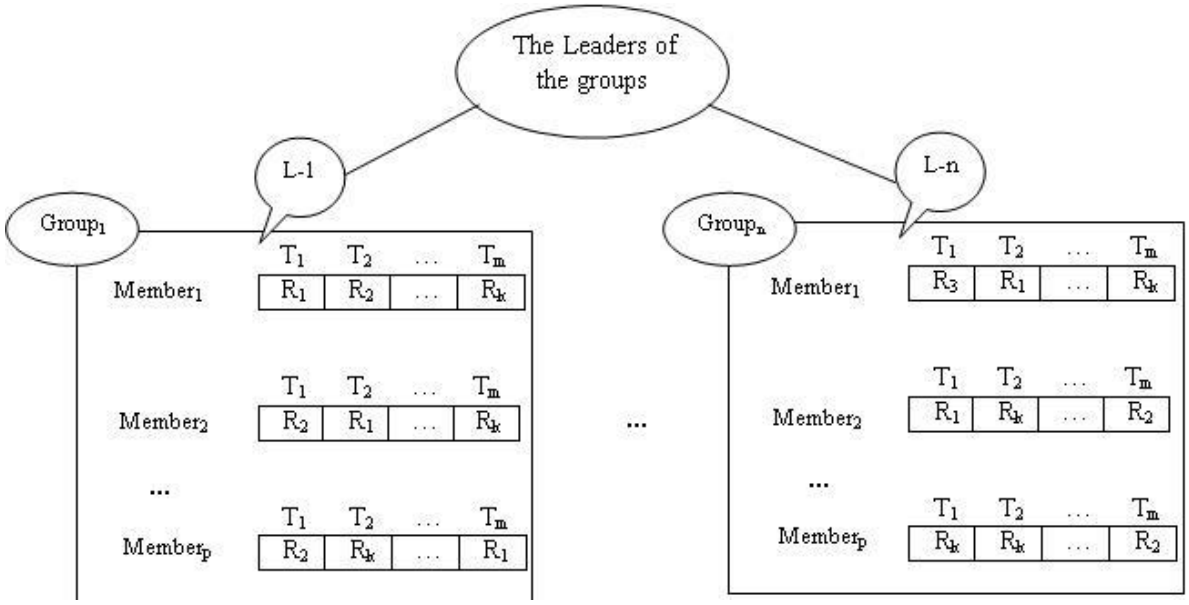


Fig. 1. Steps 1–3 of the algorithm: *n* groups consisting of *p* members are created, and their leaders are chosen based on their fitness values.

best member of each group is selected as the leader. The members of each group try to become similar to their leader in each iteration. In this way, the algorithm is able to search a solution space between a leader and its group members. It is obvious that after some iteration, members of a group may become similar to their leader. In order to introduce diversity within a group, one of its members is selected randomly and some of its variables are interchanged with a member of another group. In addition, a crossover operator helps a group come out of local minima, and the solution space can be searched again so to produce diversity. The algorithm steps are as follows:

A. Initial Population Production

A set of *p* members is produced for each group. The total population is therefore *n***p*, where *n* is the number of groups. Group and member values are produced randomly. Since the number of entering tasks is *m*, the members are represented as an *m*-dimensional array in which the stored values are resource numbers. For example, in Figure 1 we have *n* groups, each with *p* members.

B. Calculating Fitness Values of All Group Members

The fitness value is calculated for each member of each group. Since the purpose of task/job scheduling in a grid is to assign tasks to resources in a way that minimizes makespan, makespan has been chosen as the criterion for evaluating members. The less a member's makespan is, the greater is its fitness value, according to (1):

$$\text{fitness}(\text{member}_k) = \frac{1}{\text{makespan}(\text{member}_k)} \quad (1)$$

C. Determining Leader of Each Group

In each group, after the fitness value is computed for each member, the member with the best fitness value is selected as the group leader.

D. Mutation Operator

In this step, a new member is produced in each group from an older member, the leader of the group, and a random element, using (2). If the fitness value of the new member is better than the fitness value of the older member, it replaces the older member. Otherwise, the older member is retained.

$$\text{new} = r_1 * \text{old} + r_2 * \text{leader} + r_3 * \text{random} \quad (2)$$

where r_1, r_2, and r_3 are the rates determining the portion of the older member, the leader, and the random element that are used to generate the new population, such that $r_1 + r_2 + r_3 \leq 1$. Pseudocode for this step follows:

for i=1 **to***n* do {

for j=1 **to***p* do {

$\text{new}_{ij} = r_1 * \text{member}_{ij} + r_2 * Li + r_3 * \text{random}$

if fitness (new_{ij}) **better than** fitness (member_{ij})

then

$\text{member}_{ij} = \text{new}_{ij}$

end if

} end for

} end for

The value of r_1 determines the extent to which a member retains its original characteristics, and r_2 moves the member toward the leader of its group in different iterations, thus making the member similar to the leader. Careful selection of these two parameters plays an important role in the

optimization of the results. The main characteristic of this algorithm is that it searches the problem space surrounded by

TABLE I
PARAMETERS FOR THE ALGORITHMS

Algorithm	Parameter	Value
GLOA	Number of groups	3
	Population in each group	10
	r_1	0.8
	r_2	0.1
	r_3	0.1
GA	P-Crossover	0.85
	P-Mutation	0.02

the leaders. This leads to very rapid convergence to a global minimum. Note that eq. (2) is similar to the update equation for the PSO algorithm. The difference is that here, unlike PSO, the best position value of each member is not stored and so there is no information about the past positions of members.

TABLE II
THE ALGORITHMS' MAKESPAN AFTER 100 ITERATIONS (IN SECONDS)

(No. Tasks, No. Resources)	SA	GA	GSA	GGA	GLOA
(50,10)	136.742	99.198	95.562	90	89
(100,10)	307.738	183.49	190.353	181.028	167
(300,10)	973.728	638.082	626.66	597	581.842
(500,10)	1837.662	1105.56	1087	1087.216	1072.362

E. One-way Crossover Operator

In this step, a random number of members are selected from the first group and some of their parameter values are replaced with those of a member of another group that is selected randomly. It should be noted that in each iteration, only one parameter is replaced. If any new member is better it replaces the old one; otherwise the old member remains in the group. An important issue here is selecting the correct crossover rate, for otherwise all members will rapidly become similar to each other. The transfer rate t is a random number such that $1 \le t \le (\frac{m}{2}) + 1$ for each group. The purpose of the crossover operator is to escape local minima.

F. Repetition of Steps C to V according to the Determined Number of Iterations

This algorithm is repeated according to the determined number of iterations. At the end, from the different groups, the leader with the best fitness value is chosen as the problem solution.

V. SIMULATION

This section compares simulation results for our proposed algorithm with the results of several other algorithms. All algorithms were simulated in a Java environment on a system with a 2.66 GHZ CPU and 4GBRAM. Table I lists the parameters used in the performance study of our proposed algorithm and the other algorithms.

Table II shows the five algorithms' makespans for various numbers of independent tasks and 10 resources. As can be seen, SA has the worst makespans and GLOA has the best. We provide more details in Fig. 2.

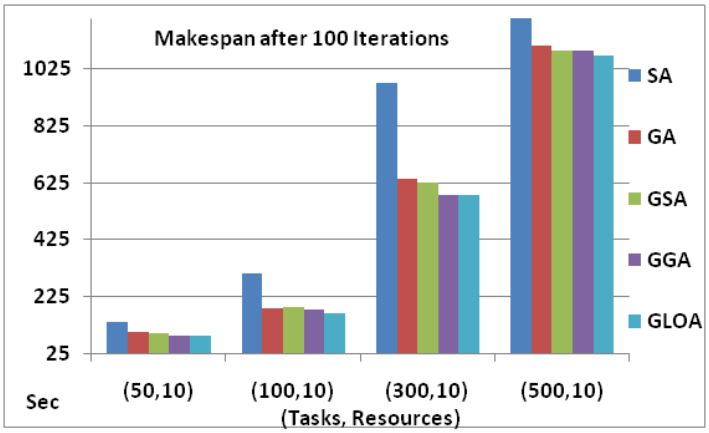


Fig. 2. Algorithms' makespan after100 iterations, with 10 resources

As we can see in Fig. 2, the SA algorithm's makespan increases rapidly as the number of tasks grows from 50 to 500.

TABLE III
ALGORITHMS' MAKESPAN AFTER 300 ITERATIONS (IN SECONDS)

(No. Tasks, No. Resources)	SA	GA	GSA	GGA	GLOA
(100,10)	233.2	172.628	179.062	175.598	166.14
(100,20)	173.116	111.946	105.314	103.092	94.55
(100,30)	120.452	90.716	87.846	80.086	77.75

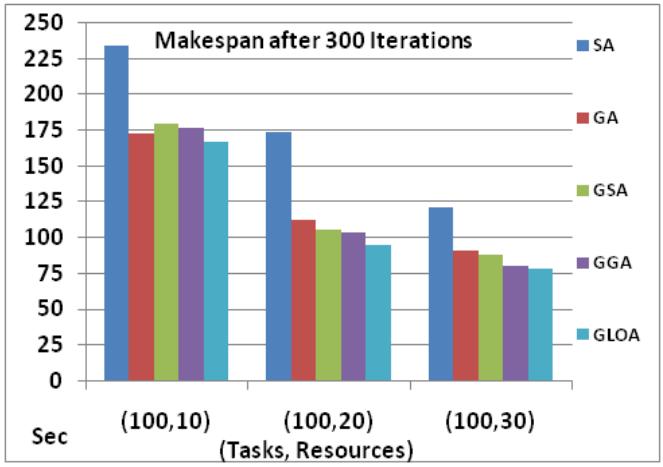


Fig. 3. Algorithms' makespan after 300 iterations, for various numbers of resources

Hence, SA is the worst algorithm for minimizing makespan and GLOA is the best in every case. In the 50-task case, the difference between SA and GLOA is approximately 48 seconds, which is less than half of the SA makespan. Here GLOA has the least makespan. When there are only a few tasks, the makespans for all of the algorithms are low, and GLOA produces the minimum. For the 300-task and 500-task cases, GGA has a similar makespan to the GLOA algorithm. For example, in the 300-task case, GGA's makespan is

approximately 597 seconds but GLOA's is approximately 582

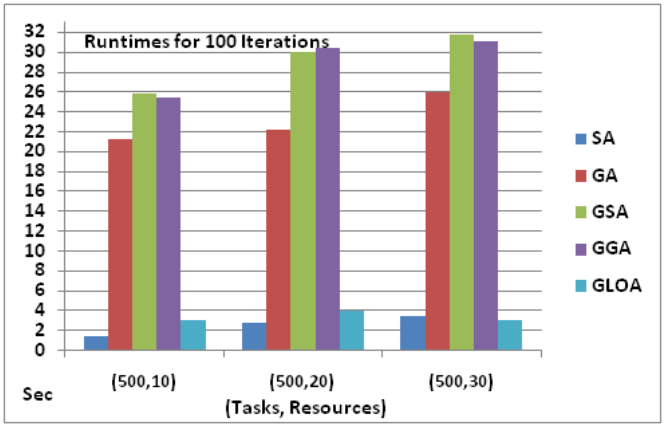


Fig. 4. Algorithm runtimes for 100 iterations with varying numbers of resources.

seconds.

Table III shows the makespans the algorithms produce for 100 fixed independent tasks for various numbers of resources. As can be seen, SA has the worst makespan in all of these cases and GLOA has the best. More details are shown in Fig. 3.

As can be seen in Fig. 3, as the number of resources increases, the makespan decreases for all algorithms, because when there are several ready resources with empty queues, tasks can be assigned to the new resources. The variation is the difference between the algorithms' structures. When the number of resources triples, the decrease for the makespan in SA is approximately 100 seconds and in GLOA it is approximately 90 seconds. As shown, GLOA has the minimum makespan in each case. Its structure provides it with the ability to be close to GGA, because like GGA, it can search the problem space both locally and globally. Hence, GLOA reaches the best solution more rapidly (e.g., in 95 seconds for 20 resources) than the other methods, particularly SA (which takes approximately 174 seconds). GLOA's makespan decreases up to 45% compared to SA, 15% compared to GA, 11% compared to GSA, and approximately 8% compared to GGA.

TABLE IV
ALGORITHMS' RUNTIME FOR 100 ITERATIONS (SECONDS)

(No. Tasks, No. Resources)	SA	GA	GSA	GGA	GLOA
(500,10)	1.4	21.2	25.8	25.4	3
(500,20)	2.8	22.2	30	30.4	4
(500,30)	3.4	26	31.8	31	3

Table IV shows the algorithms' runtime for job 500 independent tasks with varying numbers of resources. As shown, SA has the best runtime for 10 and 20 resources (because it considers only one solution, it can search more quickly than the other algorithms), and GLOA has the second best for 30 resources (because it divides the problem solutions into several groups that search in parallel, it reaches the optimum more quickly, but it takes some time to produce the several groups). Fig. 4 provides more details.

As can be seen in Fig. 4, when the number of resources increases, all algorithm runtimes increase, because when there are several new resources with empty queues, these resources must be searched and tasks assigned to them. SA is the least time-consuming algorithm (except for the 30-resource case) and GSA is the worst (except for the 500-task and 20-resource case). When the number of resources increases to 30, GLOA's runtime decreases less than SA's, because the resources have sufficiently many empty queues to be able to respond to 500 tasks more quickly, and SA considers the entire problem while GLOA divides the problem into several groups and considers the queue sizes and makespans for the tasks. When there are only a few resources (10), GA executes in just under 22 seconds, GSA and GGA have similar runtimes (just under 26 seconds), and GLOA requires just over 2 seconds, but SA requires less than 2 second. Although SA is the best algorithm in terms of runtime, it cannot produce better makespan results (as seen in Figure 2), and therefore we exempt this algorithm from consideration. When the number of resources triples (from 10 to 30), SA's runtime increases by 80%, GA's by 24%, GGA's and GSA's by 26%, but GLOA's increases by less than 10%. Therefore, while GLOA's runtime increases with the number of resources, it does so at a very low rate.

VI. CONCLUSION

Grid technology has made it possible to use idle resources as part of a single integrated system. The main purpose of grid computing is to make common resources such as computational power, bandwidth, and databases available to a central computer. The geographic spread and dynamic states of the grid space present challenges in resource management that necessitate an efficient scheduler. This scheduler should assign tasks to resources in such a way that they are executed in the shortest possible time.

This paper used a new evolutionary algorithm, GLOA, for scheduling tasks/jobs in a computational grid. Simulation results for GLOA were compared with results for four other intelligent algorithms: GA, SA, GGA, and GSA, and it was shown that in addition to wasting less computation time than the other algorithms, GLOA is able to produce shortest makespans. Also, GLOA could be applied in the real world because its runtime and makspan is less than other AI methods and produce less overhead on resources while responding the independent tasks.

In the future, we will change GLOA structure and apply it into dependent tasks in Grid Environment to cover the current gap into scheduling of dependent tasks.

REFERENCES

[1] J. Kołodziej, F. Xhafa, "Meeting security and user behavior requirements in Grid scheduling," *Simulation Modeling Practice and Theory, Elsevier*, pp. 213–226, 2011.

[2] S. Garg, R. Buyyaa, H. Siegel, "Time and cost trade-off management for scheduling parallel applications on Utility Grids," *Future Generation Computer Systems, Elsevier*, pp. 1344-1355, 2010.

[3] F. Xhafa, A. Abraham, "Computational models and heuristic methods for Grid scheduling problems," *Future Generation Computer Systems, Elsevier*, pp. 608-621, 2010.

[4] M. Shojafar, S. Barzegar, M. R, Meibody, "A new Method on Resource Scheduling in grid systems based on Hierarchical Stochastic Petri net," *3rd International Conference on Computer and Electrical Engineering (ICCEE 2010)*, pp. 175-180, 2010.

[5] Q. Tao, H. Chang, Y. Yi, Ch. Gu, W. Li, "A rotary chaotic PSO algorithm for trustworthy scheduling of a grid workflow, " *Computers & Operations Research, Elsevier*, Vol. 38, pp.824–836, 2011.

[6] R. P. Prado, S. García-Galán, A. J. Yuste and J. E. M. Expósito, "Genetic fuzzy rule-based scheduling system for grid computing in virtual organizations," *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, Vol. 15, No. 7, pp.1255-1271, 2011.

[7] S. Fidanova, "Simulated Annealing for Grid Scheduling Problem, International Symposium on Modern Computing," *IEEE John Vincent Atanasoff*, pp.41-45, 2006.

[8] B. Eksioglu, S. DuniEksioglu, P. Jain, "A tabu search algorithm for the flowshop scheduling problem with changing neighborhoods," *Computers & Industrial Engineering*, 54, pp.1–11, 2008.

[9] B. Barzegar, A. M. Rahmani, K. Zamani far, "Advanced Reservation and Scheduling in Grid Computing Systems by Gravitational Emulation Local Search Algorithm," *American Journal of Scientific Research*, No. 18, pp. 62-70, 2011.

[10] Y. Yang, G. Wua, J.Chen, W. Dai, "Multi-objective optimization based on ant colony optimization in grid over optical burst switching networks," *Expert Systems with Applications*, 37, pp.1769–1775, 2010.

[11] G. Guo-ning , "Genetic simulated annealing algorithm for task scheduling based on cloud computing environment," *International Conference of Intelligent Computing and Integrated Systems (ICISS)*, pp-60-63, 2010.

[12] J.M. Garibaldi, D. Ouelhadj, "Fuzzy Grid Scheduling Using Tabu Search," IEEE International Conference of Fuzzy Systems, pp.1-6, 2007.

[13] R. Chen, D. Shiau, Sh. Tang Lo, Combined Discrete Particle Swarm Optimization and Simulated Annealing for Grid Computing Scheduling Problem, Lecture Notes in Computer Science, Springer, Vol. 5755, 2009, pp.242-251.

[14] Z. Pooranian, A. Harounabadi, M. Shojafar and J. Mirabedini, "Hybrid PSO for Independent Task scheduling in Grid Computing to Decrease Makespan," *International Conference on Future Information Technology (ICFIT 2011)*, Singapore, pp.435-439, 2011.

[15] Z. Pooranian, A. Harounabadi, M. Shojafar, N. hedayat, " New Hybrid Algorithm for Task Scheduling in Grid Computing to Decrease missed Task," *World Academy of Science, Engineering and Technology 79*, pp. 262-268, 2011.

[16] A. Daskin, S. Kais, "Group leaders optimization algorithm, Molecular Physics," *An International Journal at the Interface Between Chemistry and Physics*, Vol. 109, No. 5, pp. 761–772, 2011.

[17] M. Yusof, K. Badak, M. Stapa, "Achieving of Tabu Search Algorithm for Scheduling Technique in Grid Computing Using GridSim Simulation Tool: Multiple Jobs on Limited Resource," *International Journal of Grid and Distributed Computing*, Vol. 3, No. 4, pp. 19-32., 2010.

[18] R. Joshua Samuel Raj, V. Vasudevan, Beyond Simulated Annealing in Grid Scheduling, *International Journal on Computer Science and Engineering (IJCSE)*, Vol. 3, No. 3, pp. 1312- 1318, Mar. 2011.

[19] Ruey-Maw Chen and Chuin-Mu Wang, Project Scheduling Heuristics-Based Standard PSO for Task-Resource Assignment in Heterogeneous Grid, *Abstract and Applied Analysis*, Vol. 2011, pp.1-20, 2011.

[20] F. A. Omaraa, M.M. Arafa, "Genetic algorithms for task scheduling problem," *Journal Parallel Distributed Computing, Elsevier*, Vol. 70, No. 1, pp. 13-22, 2010.

[21] M. Wu, D. D. Gajski, "Hyper tool: A programming aid for message-passing systems," *IEEE Transactions on Parallel and Distributed Systems*, Vol. 1, No. 3, pp. 330-343, 1990.

[22] H. El-Rewini, T. G. Lewis, H. H. Ali, *Task Scheduling in Parallel and Distributed Systems*, Prentice-Hall, 1994, ISBN:0-13-099235-6.

[23] L. Khanli, M. Etminan Far , A. Ghaffari, "Reliable Job Scheduler using RFOH in Grid Computing," *Journal of Emerging Trends in Computing and Information Sciences*, Vol. 1, No. 1, pp. 43- 47, 2010.

[24] G. Gharoonifard, F. Moeindarbari, H. Deldari, A. Morvaridi, "Scheduling of scientific workflows using a chaos- genetic algorithm," Procedia Computer Science, Elsevier, Vol. 1, No.1, pp. 1445- 1454, 2010.

[25] Q. Tao, H. Chang, Y. Yi, CH. Gu, "A Grid Workflow Scheduling Optimization approach for e-Business Application," *Proceedings of the 10th International Conference on E-Business and E-Government*, pp. 168-171, 2010.

[26] J. A. Torkestani, "A New Distributed Job Scheduling Algorithm for Grid Systems," An International Journal of Cybernetics and Systems, Vol. 44, Issue 1, pp.77-93, 2013.

Zahra pooranian received her Msc in Computer Architecture degree as honor student in Dezful Islamic Azad University since 2011. She is an instructor in Sama University in Dezful and Ahvaz since 2009. Her research interest in Grid computing specially in resource allocation and scheduling. She has worked on several papers in decreasing time and makespan in grid computing by using several AI methods such as GA, GELS, PSO, and ICA. She has published more than 5 papers especially in grid scheduling and resource allocation in various conferences, such as WASET 2010-11, ICCCIT 2011, and ICEEE 2011.Author's profile

.

Mohammad Shojafar is PhD Student in Information Communication Technology at Sapienza University of Roma from November 2012. He Received his Msc in Software Engineering in Qazvin Islamic Azad University, Qazvin, Iran in 2010.Also, he Received His Bsc in Computer Engineering-Software major in Iran University Science and Technology, Tehran, Iran in 2006. Mohammad is Specialist in Network Programming in Sensor field and Specialist in Distributed and cluster computing (Grid Computing and P2P Computing) and AI algorithms (PSO, LA, GA).

Jemal H. Abawajy is Professor in School of Information Technology of Deakin University in Australia. His interest is building secure, efficient and reliable infrastructures for large-scale distributed systems. He has published 10 books, more than 55 international well-known Journals, more than120 Conference papers and 58 Book chapters in his interest. He supervised more than 31 PhD students till now. Also, he is teaching Grid Computing Security, Distributed Computing, Information Technology Security Management and Information Security Management in Deakin University now.

Mukesh Singhal is the Professor and Gartner Group Endowed Chair in Network Engineering, Dept. of Computer Science, The University of Kentucky, Lexington, USA; and Chancellor's Professor School of Engineering of University of California, Meced, USA. He has authored and published over 200 research papers and four textbooks which 175 of them are indexed in DBLP. The books on emerging topics that Dr. Singhal edited include, Distributed computing, Cloud computing, Security and Networks. He is Journal Editorships of 5 outstanding journals such as IEEE Trans. on Dependable and Secure Computing and IEEE Trans. on Parallel and Distributed Systems. He is one of the highly cited authors in computer science and software engineering worldwide(h-index = 39, Citations = 5594).

An Analysis Architecture for Communications in Multi-agent Systems

Celia Gutiérrez
Complutense University in Madrid, Spain

***Abstract* — Evaluation tools are significant from the Agent Oriented Software Engineering (AOSE) point of view. Defective designs of communications in Multi-agent Systems (MAS) may overload one or several agents, causing a bullying effect on them. Bullying communications have avoidable consequences, as high response times and low quality of service (QoS). Architectures that perform evaluation functionality must include features to measure the bullying activity and QoS, but it is also recommendable that they have reusability and scalability features. Evaluation tools with these features can be applied to a wide range of MAS, while minimizing designer's effort. This work describes the design of an architecture for communication analysis, and its evolution to a modular version, that can be applied to different types of MAS. Experimentation of both versions shows differences between its executions.**

***Keywords* — Analysis, architecture, bullying, communications, multi-agent systems.**

I. Introduction

COMMUNICATIONS become complex to design in huge systems which interact frequently. In MAS, interactions among agents must be designed correctly to avoid behaviors that may collapse communications. The overall result of these behaviors is high response times, among other problems. Within this context, communication analysis techniques become relevant to evaluate the correct performance of the MAS. These techniques inspect the communications among agents in executions, to detect undesirable patterns of communications, like agents that are overloaded with the reception of too many messages. Once the undesirable situation is detected, the re-design the MAS communications is a straightforward task [1]. Other non desirable situations appear when there are not expected sequences of agents that interact in a conversation [2].

The effect of overloading can be compared to bullying, as explained in [3]. There are agents that play the bully role, when they send too many messages; other agents play the mistreated role when they received too many messages; other agents that play both roles, mistreated and bully; other ones that are considered as isolated because they neither send nor receive messages; and there are regular agents that behave correctly because they send and receive messages in a balanced way. There are metrics to measure the proportion of sent and received messages; these metrics are the values to classify agents into the mentioned patterns. The detection of non desired patterns in certain conversations can help the designer to modify the interactions, obtaining better response times and higher QoS results, [1], [3].

Previous frameworks for the analysis of these behaviors have been designed embedding the evaluation and debug tools within the execution of the MAS. Results can be inspected after the execution, and in consequence a straightforward re-design can be made.

Despite the satisfactory results obtained with this approach, reusability for other types of MAS becomes a difficult task, that involves re-codification of the evaluation and debug functionality. An efficient architecture is basic for the designer/tester, not only to obtain satisfactory results, but also to reuse the analysis tool in other type of MAS.

This work represents one step forward in architectures for MAS analysis. We provide a new framework for the MAS execution and evaluation in order to reach complete independence of both tasks. The result is a new architecture with two modules: one for the execution and another for the evaluation and debugging.

This research is presented in the following order: Section 2 describes the related work. The description of the new architecture is within Section 3. The results of the execution of the new architecture are included in Section 4. Finally, conclusions and future work can be found in Section 5.

II. Related Work

Literature regarding load balancing in MAS is relevant and plentiful. This problem has been focused using different strategies. [4] apply learning techniques in MAS load balancing. The task of the agent is to choose the correct resource using local information. Its objective is to optimize the resource usage. Unlike our work objective, they are not concerned in the scalability and adaptability of their solution to other problems or platforms.

The work in [5] resembles ours because they also use classification techniques and metrics to analyze the organization of MAS. They also relate their metrics and the response time, which is used as indicator of QoS. But it differs our work in the use of their metrics, which are used just to evaluate architectures; instead, we present an architecture to evaluate the communications in MAS.

AntNet [6], Challenger [7], and DIET (Decentralised Information Ecosystem Technologies) [8] use mobile agents to use their respective resources equitably, but they do not identify the cause of the overloading/bullying problem. DIET

DOI: 10.9781/ijimai.2013.219

overcomes multi-agent platforms limitations in terms of adaptability and scalability, providing a foundation for an open, adaptive and scalable agent organization. In this way, they share the same interests as we do, but they are focused on supporting basic mobile agent capabilities.

Messor [9] uses adaptive system approach. It uses an algorithm that emulates the ant behavior to distribute workload among distributed nodes. In this case, they are specially focused in peer-to-peer systems.

Other work, the Anticipate Agent Assistance (AAA) [10] also uses an agent-based metric for testing and managing the resource information of the wireless points, choosing the less overloaded access points. They are also concerned in achieving high QoS indicators of communications. However, they have confined their solution to the wireless networks.

Finally, [11] perform debugging process on recorded data of the MAS execution, like in the current work. Their analysis helps understand the behavior of the system and can reveal undesirable social behaviors. So their testing and debugging of complex MAS remains just at social level.

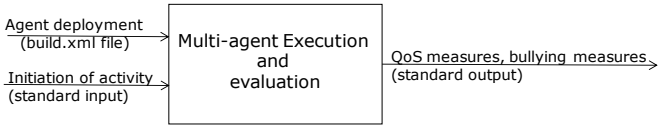
In summary, there are works that are concerned in achieving equitable behaviors of agents in MAS executions. All of them differ in the way they make the analysis, design, or evaluation, and their purpose: ones are focused on load balancing in general, others on load balancing in communications, and others in social behaviors. But neither of them has the purpose of building a scalable architecture of MAS to evaluate its communications. This architecture can integrate the elements which are present in MAS communications, as the following section describes.

III. DESIGN OF THE NEW ARCHITECTURE

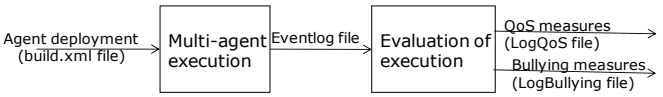
The new architecture, called IDKAnalysis 2.0, is based on a previous version, IDKAnalysis 1.0.

Both architectures follow the Ingenias methodology [12] and have been executed on Ingenias Development Kit (IDK) case studies, although they use different versions of IDK (IDKAnalysis 1.0 uses IDK 2.7, whereas IDKAnalysis 2.0 uses IDK 2.8). IDK versions use a template (build.xml) to detail the agent deployment of the case study one wants to run. At the same time, user inputs can be necessary to start the case study activity, although these inputs vary on each case study. Further details on this framework can be seen at [13]. Both versions of IDK are available at http://sourceforge.net/projects/ingenias/files/INGENIAS%20Development%20Kit/Aranjuez/, on their corresponding option.

Fig. 1 shows the differences of both versions of the architectures:



(a)



(b)

Fig. 1. Block diagram of the IDKAnalysis, version 1.0 (a) and 2.0 (b).

The first version performs the MAS execution and evaluation at the same time. The outputs only refer to the analysis, and extract the analysis measures and QoS measures. The second version is based on an architecture with a front-end that executes the main functionality of the MAS, and a back-end that analyzes the communications generated by the front-end.

There are also differences in the inputs and outputs of both architectures:

- In the first one, apart from the agent configuration, it is necessary human interaction to start the activity, whereas in the second one, the execution starts automatically (without the user input).
- The outputs of the first version are shown at the same time. In the second version, the front-end outputs a log file with the events recorded; the back-end receives as input the event log file, and produces the two outputs physically separated in two files.

Inputs and outputs of the back-end are described and analyzed in the following subsections.

A. Event log file

The event log file registers the main events of the MAS execution with certain format that corresponds to the main features of these events. The generation of this file is a characteristic functionality of IDK2.8.

The standard format of a line is as follows:

```
Timestap(hours:minutes:seconds:milesecond)
;Name of the event;Additional fields
```

Additional fields depend on the type of event it represents. Below there is an example of the event that represents *A new entity is added to the agent mental state*:

```
23:53:47:187;MEAddedToMentalState;BuyerAge
nt_0multipleBuyers@viriato:60000/JADE!Curr
entAssistedAgent!ME0
```

where the content of the additional fields are:

involved agent –> `BuyerAgent_0multipleBuyers@viriato:60000/JADE`
kind of entity –> `CurrentAssistedAgent`
entity id –> `ME0`

To register all communication information, the types of

events of this version include message shipping and reception events, and others that are necessary to measure response times. Even more, with the intention of using the event log file for other purposes than communication analysis, a wide range of types of events is included:

1) A new entity is added to the mental state.
2) An agent was initialized completely.
3) A task was scheduled within the agent.
4) A task was executed.
5) An agent is starting collaboration as initiator.
6) An agent has accepted to participate in an interaction as collaborator.
7) An agent has received a request to participate in collaboration.
8) A mental entity has been removed from the conversation.
9) An agent received a message.
10) An agent sent a message.
11) An entity was added to a conversation.

In IDK 2.8 the name of the event log file is generated in such a way that it contains the day, month, year, hour and minute of its creation.

An excerpt of an event log file can be found in the Appendix section.

B. Outputs of the Evaluation Module

As Fig. 1 (b) shows, the outputs of the second module are the QoS measures (in this case response times) and the bullying measures. This module is coded in Java, JDK1.7.0_04. For this purpose, there are two types of events selected from the event log file.

The first output depends on each case study and basically is the time elapsed since a service is requested until an offer of that service is proposed. For MAS with a lot of service responses (as a consequence of having many agents offering services), it is may be useful to establish a number of iterations or responses until a response time is recorded.

It is necessary to choose the task when the time measuring process initiates and the task when it finishes. The response time is the elapsed time between them. This depends on each case. In the experimentation of Section 4, the initiating task is ChooseMovie, and the finishing task is ChooseCinema. The type of event that records the executed task is *TaskExecuted*. In the example below, the log refers to the starting time of execution of ChooseMovie task.

```
18:22:02:355;TaskExecuted;InterfaceAg
ent_3expInterfaceAgentwithprofile!ChooseMo
vie!ME103705
```

where the additional fields mean:

involved agent –>
`InterfaceAgent_3expInterfaceAgentwithprofi le`
task type –> `TaskExecuted`
task name –> `ChooseMovie`
task id –> `ME103705`

The second output is the bullying measures, which are described in detail in a previous work [3]. In this case, *MessageReceived* event is used each time a message is received by an agent, as in the following example:

```
23:53:48:885;MessageReceived;BuyerAssignme
nt!0.InterfaceAgent_9multipleInterfaceAgen
tsvir1225148028355!RejectBecomingAssistant
!BuyerAgent_4multipleBuyers!InterfaceAgent
_9multipleInterfaceAgents
```

where the additional fields are:

protocol –> `BuyerAssignment`
conversation id –>
`0.InterfaceAgent_9multipleInterfaceAgentsv ir1225148028355`
protocol state from which the message is sent –>
`RejectBecomingAssistant`
sender –> `BuyerAgent_4multipleBuyers`
receiver –>
`InterfaceAgent_9multipleInterfaceAgents`

In this way, information about senders and receivers is enough to compute the measures of [3] and start the evaluation process.

Although the measures are standard for any type of MAS with agents playing different roles, the designer must also specify which the role is going to be analyzed as bully, and which one as the mistreated. Besides, he must tune a threshold. As explained in [3], the computed measures are compared with the indicated values for each pattern, although a margin between both values is established as threshold.

Considering that all these features must be customizable for executions of other types of MAS, this module contains the following parameters:

1) Path of the Eventlog file, LogBullying file, QoS file.
2) Name of the LogBullying file
3) Name of the Qos file.
4) Role that is suspected to be the Bully in the conversations.
5) Role of that is suspected to be the Mistreated in the conversations.
6) Threshold for the bullying metrics.
7) Number of iterations that a task must be executed to calculate the response time.
8) The initial task that must be executed to start the response time counting.
9) The final task that must be executed to end the response time counting.

C. Advantages of IDKAnalysis 2.0 over IDKAnalysis 1.0

Case studies built under IDKAnalysis 2.0 offer several aspects of the executions that make it applicable to other case studies. These features appear on each module:

1) The event log file generated by the first module does not only record communication related to events, but also other events that can be analyzed for different purposes.
2) The second module produces two different files, so bullying measures and response times can be analyzed separately. Besides, this module contains some parameters that can be tuned, so it can be adapted to other methodology case studies.

Figs. 4 and 5 (in the Appendix section) show the running architectures of both versions using the experimentation described in Section 4. Fig. 4 (a) shows the architecture of the first version, where the distinction between the front-end and the back-end does not exist. The second version in Fig. 4 (b) contains the *srceclipse* package, which is the back-end, whereas the rest of the packages compose the front-end. The *srceclipse* package, which does not appear in Fig. 4 (a), is also composed of the *bullying* package and the *logs* package, as Fig. 5 shows. The first one contains the source and binary files for the evaluation process, and the second one is the directory where the log files (inputs and outputs) are placed. As explained in the previous subsection, this directory is the first parameter the designer/tester can customize.

IV. Experimentation Results

Executions of both versions have been carried out using the Cinema case study, pursuing the objective of acquiring cinema tickets according to certain user's preferences. The participant roles are the following:

- *Interface agent*, which represents the customer.
- *Seller agent*, which represents the cinema.
- *Buyer agent*, which represents the intermediary between the Seller and the Interface.

The hardware of the experimentation has been a machine with 2 GHz and 2GB RAM, using 32-bit Windows 7 Professional.

The Cinema case study uses Java Agent DEvelopment (JADE) platform. JADE framework uses the Foundation for Intelligent Physical Agents (FIPA) standard for communications among agents.

As table 1 shows, configurations with different numbers of agents for each role have been run:

TABLE I
CONFIGURATIONS FOR CINEMA CASE STUDY

Configuration	Number of Interface Agents	Number of Seller Agents	Number of Buyer Agents
Serious	10	5	10
Simple	20	4	20
FullSystem	100	8	100

The following subsections include examples of executions on both versions of the tool.

A. Execution using IDKAnalysis 1.0

The Cinema case study begins with two possible options for the use, as Fig. 2 shows.

Fig. 2. Initial GUI of the Cinema case study built with IDKAnalysis 1.0

It is necessary to start running by selecting *Start monthly activity*. This will produce the conversations between the agents, in order to get the proposed tickets. As this is not the relevant part of this work, no output has been extracted. Then, *Bullying Measures* can be selected, to obtain the values for the bullying metrics and response times from the generated communications.

A snapshot of this execution on console can be seen in Fig. 3, where the metrics and classification for IntergaceAgent_16, IntergaceAgent_19, IntergaceAgent_18 agents, and the corresponding values for the roles and the system, can be seen alongside the extraction of a response time.

B. Execution using IDKAnalysis 2.0

Mentioned parameters in subsection 3.B, numbered from 4 to 9, have been tuned as follows:

- Role that is suspected to be the Bully in the conversations: Interface
- Role of that is suspected to be the Mistreated in the conversations: Buyer
- Threshold for the bullying metrics: 1.0
- Number of iterations that a task must be executed to calculate the response time: 10
- The initial task that must be executed to start the response time counting: ChooseMovie
- The final task that must be executed to end the response time counting: ChooseCinema

In this way, the response time which is recorded, is the elapsed time between the ChooseMovie task and the tenth occurrence of the ChooseCinema task.

In the Appendix section, there are examples of the two outputs generated by the IDKAnalysis 2.0 using the FullSystem configuration. They are generated in two separate files, to facilitate the designer analysis.

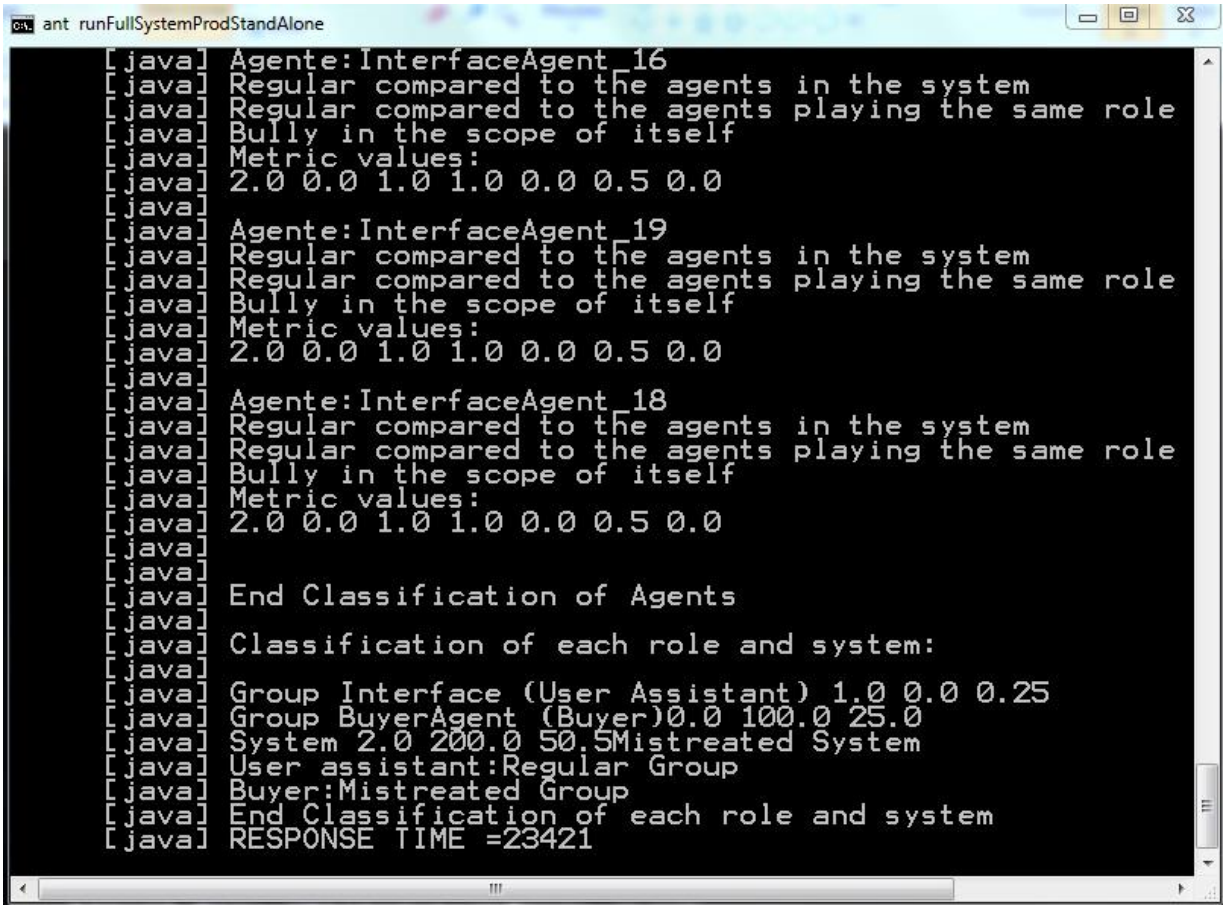


Fig. 3 Output of the IDKAnalysis 1.0 for a FullSystem configuration.

V. Future Work and Conclusions

In this work we have presented a new framework which separates the multi-agent system execution and the evaluation of the communication among agents.

This perspective provides several advantages from the Software Engineering point of view:

- To work on the functionality or the evaluation process directly, by introducing changes in the front-end (for the first purpose), or the back-end (for the second purpose).
- To inspect bullying behaviors and QoS measures separately, by the analysis of the LogBullying file (in the first case), or the QoS file (in the second case).
- To reuse the evaluation module in other case studies, by tuning some parameters accordingly to each multi-agent
- system circumstances. The range of events generated by IDK8.0 (and IDKAnalysis2.0 in consequence) offers different possibilities to record QoS, which does not necessarily use the TaskExecuted event, but other ones.

This architecture offer several possibilities of future work. It is thought to use the evaluation module in MAS with different purposes and frameworks:

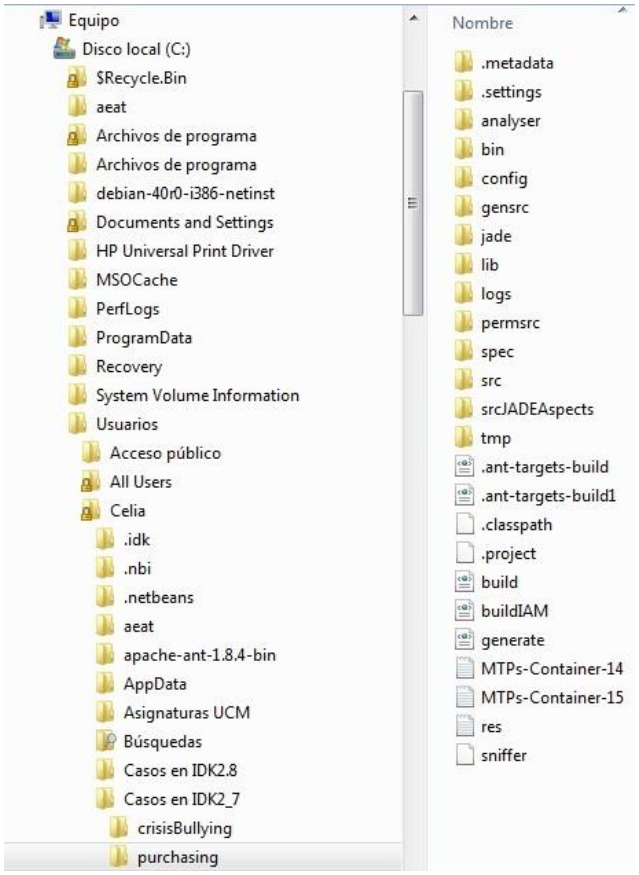
- ADELFE methodology [14] for Adaptive MAS.
- ICARO-T framework [15] for agent organizations. Available at http://icaro.morfeo-project.org/
- Agent Based Social Simulation frameworks.

The combination of IDKAnalysis 2.0 with the above methodologies will provide experimentation outputs with two purposes:

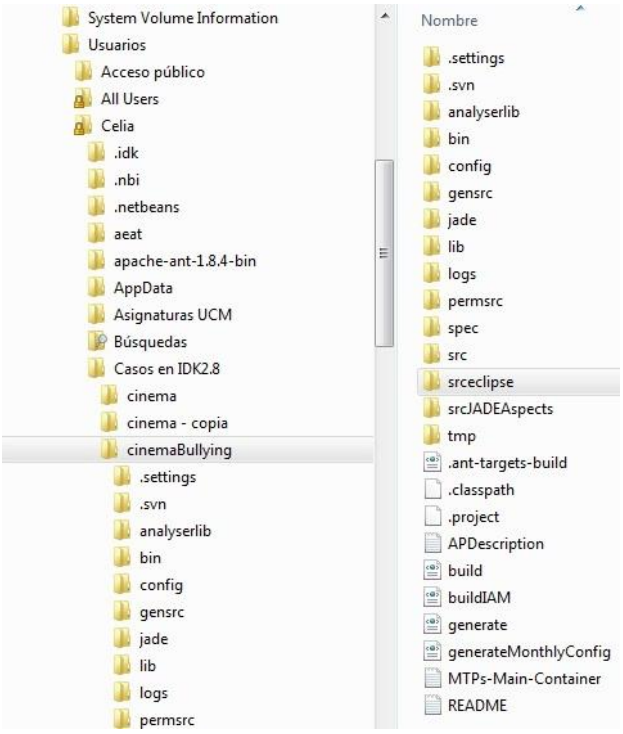
1) Validate and enlarge the evaluation framework with the experimentation results. In particular, it is necessary a previous extraction of the event logs. These logs must accomplish the basic format of the log file mentioned in subsection 3.A. Even more, as log extraction is used for other purposes, an ontology may be parsed to get the correct parameters for each purpose. This new component and other ones will be incorporated in a new version of the tool, IDKAnalysis 3.0.
2) Enlarge the mentioned methodologies and frameworks from the AOSE point of view, with a complete module that provides testing and debugging tools.

Acknowledgment

C. Gutiérrez thanks Jorge Gómez-Sanz for his continuous help on the use of IDK releases, and also for providing the infrastructure for the Cinema case study.



(a)



(b)

Fig. 4 A snapshot of the running architecture top level in IDKAnalysis1.0 (a) and IDKAnalysis2.0 (b).

APPENDIX

This section contains two types of information:

1) Snapshots of the running architecture of the Cinema case study using IDKAnalysis1.0 and IDKAnalysis2.0. In the first snapshot, belonging to IDKAnalysis1.0, the package deployment does not show the distinction between the front-end and the back-end. This fact is reflected in the second snapshot, belonging to IDKAnalysis2.0. The third snapshot shows the content of the back-end. Further explanations are provided in subsection 3.C.
2) Samples of the input and outputs of the evaluation module for the execution of the Cinema case study, using the parameter configuration in subsection 4.B.

- This is an excerpt of an event log file. Each line contains the information of an event, according to the syntax described in subsection 3.A:

```
18:21:26:770;TaskExecuted;InterfaceAgent_6
5expInterfaceAgentwithprofile!Look_for_an_
assistant!ME1044
18:21:26:770;MessageSent;BuyerAssignment!0
.InterfaceAgent_67expInterfaceAgentwithpro
filePC-
1227028885694!enable!InterfaceAgent_67expI
nterfaceAgentwithprofile!BuyerAgent_6expBu
yerAgentWithProfile@PC-
sheilacg:60000/JADE,
18:21:26:770;TaskScheduled;InterfaceAgent_
7expInterfaceAgentwithprofile!Look_for_an_
assistant!ME1167![ME19:GetAssignments]
18:21:26:770;MessageSent;BuyerAssignment!0
.InterfaceAgent_67expInterfaceAgentwithpro
filePC-
1227028885694!RequestBeingAssistant!Interf
aceAgent_67expInterfaceAgentwithprofile!Bu
yerAgent_6expBuyerAgentWithProfile@PC-
sheilacg:60000/JADE,
18:21:26:786;MessageSent;BuyerAssignment!0
.InterfaceAgent_1expInterfaceAgentwithprof
ilePC-
1227028886568!enable!InterfaceAgent_1expIn
terfaceAgentwithprofile!BuyerAgent_6expBuy
erAgentWithProfile@PC-sheilacg:60000/JADE,
```

- This is an excerpt of the LogBullying file. It reflects the classification values and measures for one of the Interface agents, both roles and the whole system:

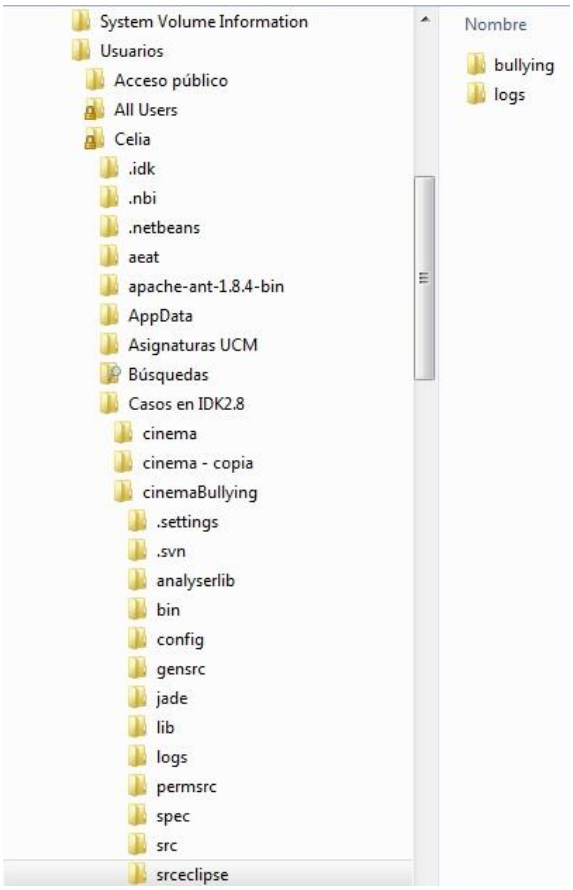


Fig. 5 A snapshot of the running architecture second level (back-end) in IDKAnalysis2.0.

```
Agente:InterfaceAgent_7expInterfaceAgentwi
thprofile
numOutputAgent =26.0 NumAgent =40
numOutput = 514.0 Bully proportionally to
the bully agents in the system
Regular compared to the agents playing the
same role
Bully in the scope of itself
Metric values:
2.0233462 0.0 1.0116731 0.50583655 0.0
0.25291827 0.0

End Classification of Agents

Classification of each role and system:

Group CoordA 0.0 0.0 0.0
Group NetworkA 0.28266892 4.5903044
1.2182432
System 2.2613513 36.722435
9.745946Mistreated System
Bully System
CoordA:Regular Group
NetworkA:Mistreated Group
End Classification of each role and system
```

- This is an exerpt of the LogQoS file. Each line contains the response times (in milliseconds) obtained with a frequency of 10 iterations:

```
10 iterations 6755
20 iterations 8106
30 iterations 8994
40 iterations 9511
```

REFERENCES

[1] C. Gutiérrez, I. García-Magariño, and R. Fuentes-Fernández, "Detection of undesirable communication patterns in multi-agent systems," *Engineering Applications of Artificial Intelligence*, vol. 24, no. 1, pp. 103-116, 2011.

[2] C. Gutierrez and I. García-Magariño, "Extraction of execution patterns in multi-agent systems," *IEEE Latin America Transactions*, vol. 8, no. 3, pp. 311-317, 2010.

[3] C. Gutiérrez and I. García-Magariño, "Revealing bullying patterns in multi-agent systems," *Journal of Systems and Software*, vol. 84, no. 9, pp. 1563-1575, 2011.

[4] S. Andrea, Y. Shoham, and M. Tennenholtz, "Adaptive Load Balancing: A Study in Multi-Agent Learning," *Journal of Artificial Intelligence Research*, vol. 2, no. 1, pp. 475-500, 1994.

[5] S. K. Lee and C. S. Hwang, "Architecture modeling and evaluation for design of agent-based system," *Journal of Systems and Software*, vol. 72, no. 2, pp. 195-208, 2004.

[6] R. Schoonderwoerd, O. Holl, J. Bruten, and L. Rothkrantz, "Ant-based load balancing in telecommunications networks," *Adaptive Behavior*, vol. 5, no. 2, pp. 169-207, 1996.

[7] A. Chavez, R. Moukas, and P. Maes, "Challenger: A Multiagent System for Distributed Resource Allocation," *Autonomous Agents*, vol. 97, pp. 323-331, 1997.

[8] C. Hoile, F. Wang, E. Bonsma, and P. Marrow, "Core Specification and Experiments in DIET: A Decentralised Ecosystem-inspired Mobile Agent System," in *Proc. 1st Int. Conf. Autonomous Agents and Multi-Agent Systems*, Bologna, 2002, pp. 623-630

[9] A. Montresor and H. Meling, "Messor: Load-Balancing through a Swarm of Autonomous Agents," in *Proc. 1st Workshop on Agent and Peer-to-Peer Syst.*, Bologna, 2002, pp. 125–137.

[10] Y. C. Chen and W. Y. Chen, "An agent-based metric for quality of services over wireless networks," *Journal of Systems and Software*, vol. 81, no. 10, pp. 1625-1639, 2008.

[11] E. Serrano, A. Quirin, J. Botia, and O. Cordón, "Debugging complex software systems by means of path finder networks," *Information Sciences*, vol. 180, no. 5, pp. 561-583, 2010.

[12] J. Pavón and J.J. Gomez-Sanz, "Agent Oriented Software Engineering with INGENIAS," in *Proc. 3rd Int. Central and Eastern European Conf. on Multi-Agent Systems*, vol. 2691, Prague, pp. 394-403, 2003.

[13] J. J. Gomez-Sanz, R. Fuentes-Fernández, J. Pavón, and I. García-Magariño, "INGENIAS Development Kit: a visual multi-agent system development environment," in *Proc. 7th Int. Conf. Autonomous Agents and Multiagent Systems*, 2008, pp. 1675-1676.

[14] G. Picard and M-P. Gleizes, "The ADELFE Methodology," *Designing Adaptive Cooperative Multi-Agent Systems*, chapter 8, pp. 157-176, Kluwer Publishing, 2004.

[15] J. M. Gascueña, A. Fernández-Caballero, and F. J. Garijo, " Programming Reactive Agent-based Mobile Robots using ICARO-T Framework," in *2010 Proc. ICAART Conf*., vol. 2, pp. 287-291.

C. Gutierrez was born in Bilbao, Spain, in 1969. She received the B.Eng. and Ph.D. degrees in Computer Science from the University of Deusto in 1992, and from the University of the Basque Country in 2000, respectively. She has worked for Labein technological research center, in Bilbao. She has also worked for the Basque Government and Indra, as a software Engineer. After doing teaching and research at some private universities, her current position is Assistant Professor at the Complutense University in Madrid, Spain, with the Grasia research group. She has published in relevant journals and has also direct financially supported projects. Her current interests focus on Data Mining, Multiagent Systems and Accessibility.

A Case-based Reasoning Approach to Validate Grammatical Gender and Number Agreement in Spanish language

Jorge Bacca, Silvia Baldiris, Ramon Fabregat, Cecilia Avila
University of Girona, Spain

***Abstract* — Across Latin America 420 indigenous languages are spoken. Spanish is considered a second language in indigenous communities and is progressively introduced in education. However, most of the tools to support teaching processes of a second language have been developed for the most common languages such as English, French, German, Italian, etc. As a result, only a small amount of learning objects and authoring tools have been developed for indigenous people considering the specific needs of their population. This paper introduces Multilingual–Tiny as a web authoring tool to support the virtual experience of indigenous students and teachers when they are creating learning objects in indigenous languages or in Spanish language, in particular, when they have to deal with the grammatical structures of Spanish. Multilingual–Tiny has a module based on the Case-based Reasoning technique to provide recommendations in real time when teachers and students write texts in Spanish. An experiment was performed in order to compare some local similarity functions to retrieve cases from the case library taking into account the grammatical structures. As a result we found the similarity function with the best performance.**

***Keywords* — Authoring tool, Second language acquisition, indigenous people, Case-based reasoning, local similarity functions.**

I. INTRODUCTION

IN bilingual virtual training programs for teachers that have an indigenous language as mother tongue [1] [2], there are some difficulties when teachers design and create learning objects to teach Spanish as a second language for indigenous population. Some of those difficulties were reported in [3] and are mainly related to the process of writing texts, in particular the use of grammatical gender and number in the Spanish language. The main cause of this situation is that some indigenous do not have masculine or feminine distinction, or there are particular ways to express grammatical number that differs significantly from Spanish language.

In consequence, teachers have to be aware of some rules in order to properly apply the grammatical rules of Spanish and take care of teaching them correctly to their students. Nevertheless, in some cases, indigenous teachers of Spanish language use some didactic strategies, as reading from textbooks and the language class [4], designed to teach indigenous languages but they apply them to teach Spanish language. This situation can create some problems in students, because they do not reach a good Spanish level, so it will affect the learning process of other subjects in the future. As a result of these issues, some learning objects that are written by indigenous teachers in Spanish may contain some grammatical errors in the texts.

As a solution, in this paper we introduce Multilingual-Tiny, a web authoring tool based on the TinyMCE [5] web content editor which consist of a complete set of plug-ins and online services for teachers to support them in the learning objects design and development. Multilingual-Tiny also has a module that applies Case-based reasoning (CBR), in order to provide recommendations (based on the grammatical structure of sentences) and taking into account the previous experience of skilled teachers from writing Spanish texts and well-formed texts obtained from the Internet. All of this process support teachers of Spanish language when they are creating their learning objects, mainly when they are writing texts in Spanish language.

This document is organized as follows: In section 2, some concerns about teaching Spanish as a second language are presented. In section 3 the architecture design of Multilingual-Tiny is described, including the applied CBR cycle. Section 4 describes an illustrative scenario which present the complete process performed by Multilingual-Tiny and also how the CBR technique was applied. Section 5 describes the followed validations process as well as the obtained results. Finally conclusions are presented in section 6.

II. TEACHING SPANISH AS A SECOND LANGUAGE

Teaching Spanish as a second language to indigenous communities is not a trivial task. It supposes a challenge to governments and universities in which is important to promote effective Bilingual Intercultural Programs (BIE) and at the same time, training teachers effectively in order to introduce Spanish in a coordinated bilingualism method [4], in which both, mother tongue or L1 and second language or L2 are developed at the same time. In this context, the mother tongue (which is an indigenous language), is acquired by a natural

DOI: 10.9781/ijimai.2013.2110

process [6]. The second language – L2, in this case, the Spanish language, is taught for facilitating indigenous people communication with Spanish speakers and also to receive instruction in some knowledge areas which are taught in Spanish.

Despite the efforts and advances obtained by applying the Bilingual Intercultural Programs in some countries such as Mexico and Peru, teachers of Spanish language may face some difficulties when they have to teach indigenous people how to read and write in Spanish and in the indigenous language [7] at the same time. Some of those difficulties are due to the fact that teachers of Spanish have an indigenous language as mother tongue and they learnt Spanish in a non-systematic way. The consequence is that those teachers use the same strategies for teaching both languages, so it could be counterproductive in student's learning process [8].

When teaching Spanish, teachers usually can follow two complementary strategies: reading from textbooks and the language class [5]. The former is a strategy in which teachers introduce and explains the topic in the indigenous language and then students read the book in Spanish language so that students identify vocabulary and pronunciation. Finally, the teacher explains vocabulary or concepts that students may have lost in the reading. The latter strategy is the language class, in which teachers of Spanish compare the indigenous language with the Spanish language in terms of grammar, vocabulary and structure in order to promote reflection and develop the meta-linguistic awareness [5].

In this context, in teacher's training, when universities are preparing indigenous students that will be future teachers of Spanish language for teaching in their indigenous communities, students have to develop competencies and skills in order to effectively apply the teaching strategies mentioned above and other didactic and pedagogic methods. Multilingual-Tiny, the web authoring tool developed, takes a relevant role in this task; giving recommendations to teachers to avoid grammatical errors. As a result, teachers can create quality educational content to teach Spanish and create learning objects in their mother tongue.

III. Multilingual-Tiny Approach

A. Overview Architecture

Multilingual-Tiny is a web authoring tool developed in order to support indigenous students which will be future teachers of Spanish language in indigenous communities and teachers of this population, when they are creating learning objects, in particular, when they have to deal with some grammatical structures of sentences in Spanish. Multilingual-Tiny consist of plug-ins and online services to provide a virtual environment to design and develop learning objects in Spanish and indigenous languages and has a module based on the case-based reasoning technique, to provide recommendations in order to avoid grammatical errors and develop quality educational content.

The architecture of Multilingual-Tiny is depicted in Fig. 1. The architecture has 4 layers, from top to the bottom: The users layer, represent indigenous teachers and students that interact with Multilingual-tiny. The interface layer includes the authoring tool and shows the recommendations that come from the CBR module. The services layer provides a group of services for text processing and includes the CBR based module to provide the recommendations. Finally the data access layer includes services for data storing, such as the case library.

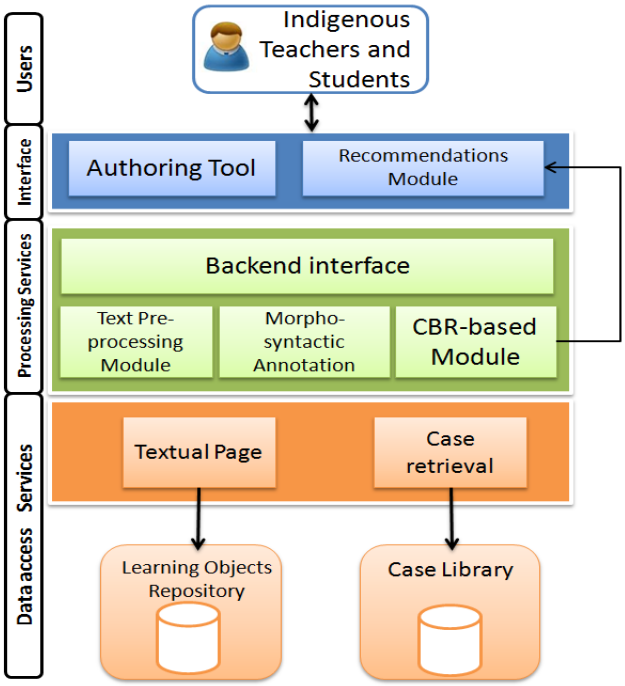


Fig. 1. Multilingual-Tiny Architecture.

B. Layer Description

The following paragraphs provide a detailed description about each of the layers mentioned before.

1) Users Layer

This layer represents the users that interact with Multilingual-Tiny, for instance, indigenous students that will be future teachers of Spanish language in indigenous communities and indigenous teachers. These users interact with the interface layer to use the service in order to create the learning objects.

2) Interface Layer and Authoring tool

Interface layer includes the authoring tool and the recommendations. The authoring tool is based on the TinyMCE [5] web content editor, which is an open source JavaScript based web editor that provides a group of services in order to create web pages without worrying about HTML code, because HTML is generated by it. The authoring tool can be integrated in the ATutor [9] e-learning platform or in other platforms.

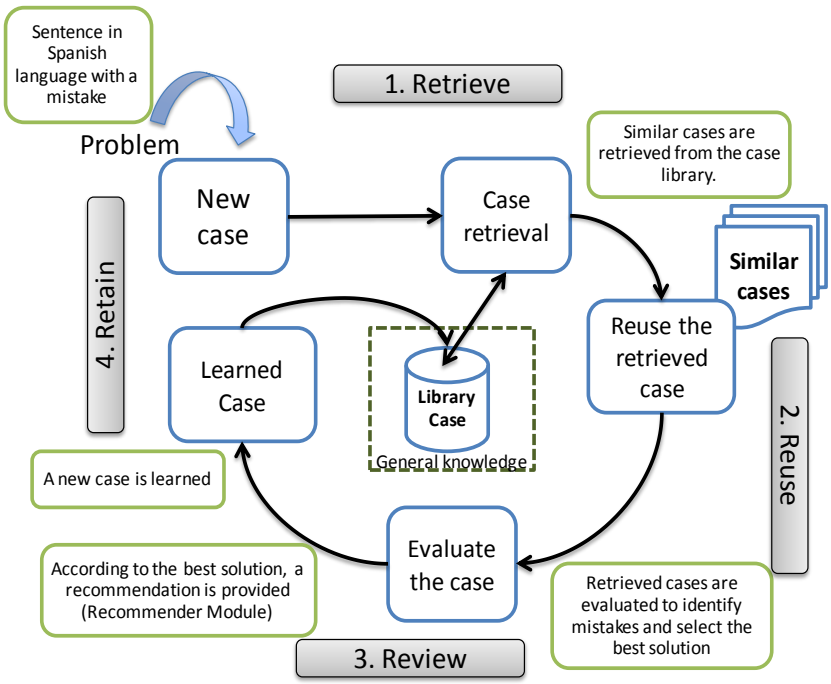


Fig. 2. CBR Cycle applied to generate grammatical recommendations for indigenous population.

As a result teachers can easily create web pages which will be part of a course in the ATutor e-learning platform as learning objects.

The authoring tool establishes communication with the Processing services layer when a learning object is being created. The text written in Spanish by indigenous teachers or students in the authoring tool is then sent to the Processing services layer to be analyzed.

The recommender module in the interface layer shows the recommendations that come from the CBR based module. These recommendations include suggestions on how to correct grammatical errors. The recommendation process is described in detail in next sections.

3) Processing Services Layer

This layer includes the services for text processing, the morpho-syntactic annotation module and the CBR based module. Those services are combined in order to provide recommendations to teachers when they are creating the learning objects to teach Spanish. The input of this layer is the text of the learning object that is being created in the authoring tool. The components of this layer are:

1) **Text Pre-processing Module:** The text pre-processing module is based on the open source FreeLing [10] library for Natural Language Processing. The input of this module is a text which has been written by the teacher as part of a learning object. This text is automatically split into sentences and the resultant sentences are split into words. This process is based on dictionaries and rules of the FreeLing library. The result of this process will be the input of the morpho-syntactic annotation module.

2) **Morpho-syntactic Annotation Module:** This module provides the morpho-syntactic annotation, which is a process of assigning tags for every word in the text, depending on the grammatical category. This process is based on the PoS (Part of Speech) tagging of FreeLing library. The input of this module is the output of the pre-processing module (which is a group of words). The PoS tagging is based on the EAGLES [11] recommendations. EAGLES define a group of standard tags for every grammatical category. As a result, each word of the text is assigned a tag depending on the context and grammatical structure of each sentence. The outputs of this module are groups of part-of-speech tags which represent a sentence. These tags will be an important component of the case representation in the case based reasoning module.

3) **Cased-based Reasoning Module:** This module is based on the Case-based reasoning technique. It takes the output of the morpho-syntactic annotation module, and executes the CBR cycle. As a result it provides the recommendations to indigenous teachers and students in order to correct grammatical errors when they write texts in Spanish language during learning objects creation. The module was built with jCOLIBRI framework [12], each case from the case library consists of a group of tags (part of speech tags) which represent a well-formed sentence. The CBR cycle, which includes 4 steps (Retrieve, Reuse, Review and Retain), is applied to grammatical sentence analysis in Spanish language and the process is depicted in fig. 2.

The main steps of this process are:

1) The Retrieve step: In this step a new case that comes from the morpho-syntactic annotation module, which is a new sentence, is compared with the cases stored in the case library by means of the similarity algorithm. As a result the most similar cases are retrieved. Both components are used:
   - Case library: Composed by a group of cases which are well-formed sentences in Spanish language obtained from a wide variety of texts from Internet. The case library is updated and new cases are stored when teachers add a new sentence structure. The case library is part of the Data Access layer which establishes communication with the services layer in order to store and retrieve cases.
   - Similarity Algorithm and Retrieve Component: Based on the JCollibri framework, the nearest neighborhood algorithm K-NN [13] is applied in order to retrieve the most similar cases when a new sentence is being analyzed. This process uses a global similarity function and a local similarity function for each attribute from the case.

2) The Reuse Step: In this step the K most similar cases obtained, by computing similarity, as described above are selected and the CBR Module organizes the cases according with the weights defined by the Morpho-syntactic Annotation Module.

3) In the next step, Review, the cases are evaluated in order to identify if the sentence is correct or if the sentence has a grammatical error. Besides, the case could be adapted or transformed to provide a recommendation about how to properly write the sentence. Further details about the overall process are depicted in section 4.

4) In the next step, which is called Retain, a new case obtained from the adaptation of the retrieved case is converted into a new case. Which is part of the recommendations provided by the recommender module and on the other hand it is stored in the case library as a new case. As a result from the process, grammatical errors in terms of using gender and number could be identified and a recommendation on how to correct it is provided to students.

4) Recommender Module

This module takes the cases retrieved from the case library as an input for providing recommendations to teachers or students on how to correct the sentence if a grammatical error in gender and number is identified. These recommendations take into account the indigenous language of teachers and students in order to explain why the sentence was incorrect from the indigenous language grammatical perspective.

IV. An Illustrative Scenario

As well known, the CBR cycle includes 4 steps (Retrieve, Reuse, Review and Retain) as shown in Fig. 2. In this section a step-by-step illustrative case based on CBR cycle is applied in order to show how the grammatical sentence analysis in Spanish language is performed in Multilingual-Tiny to provide recommendations to students and teachers.

Step 1 - Writing the text:

Indigenous students which are preparing to be future teachers of Spanish language write a text in the web content editor when they are creating learning objects. In this step it is probably that students make mistakes in terms of grammatical issues when they write a text in Spanish but they are frequently thinking in their mother tongue which is an indigenous language. For instance:

― *Me gustan el gatos blancos* (sentence with a mistake in Spanish).

― *I like white cats* (English translation only for illustrative purposes).

The above sentence in Spanish has a mistake in the definite article ("el") because it is in singular form but it must be in plural form ("los").

Step 2 – Text Pre-Processing (Morpho-syntactic annotation of the initial text):

In this step the system takes the initial text and applies the morpho-syntactic part-of-speech annotation of the text according to EAGLES recommendations [11]. Taking the example mentioned above the morpho-syntactic annotation is depicted in table 1. It is important to remark that in table 1 for English language the sentence seems to be grammatically correct, but in Spanish language there is a mistake when using the definite article ("el") (which in English is "the") in singular form with a noun "gatos" (in English "cats") in plural form.

Step 3 – Case retrieval

Based on the morpho-syntactic annotation from step 2, in which each word has a specific tag (as depicted in table 1), a new case is created; this case is composed by the group of EAGLES tags. The new case could be: Case[PP1CS000, VMIP1P0, DA0MS0, NCMP000, AQAMP0]. This case is equivalent to the sentence: "*Me gustan el gatos blancos*" (in English: I like white cats). The new case is compared by means of the nearest-neighbor algorithm [13] with cases previously stored in the case library. The most similar cases are retrieved. For instance, if the following case is retrieved: Case=[PP1CS000,VMIP1P0,DA0MP0,NCMP000,AQAMP0], with a computed similarity of 96% from the global similarity function. It is important to remark that cases stored in the case library have been obtained from texts without grammatical errors.

TABLE I
MORPHO-SYNTACTIC ANNOTATION OF THE EXAMPLE.

Word – Token		Part-of-Speech tagging (EAGLES)	Meaning of the tag assigned to each word
Words in Spanish	Translation to English		
Me	I	PP1CS000	Personal pronoun, first person, common gender in singular form.
Gustan	Like	VMIP1P0	Main verb, indicative, present form, first person and plural.
El	the	DA0MS0	Define article, masculine, in singular form.
Gatos	cats	NCMP000	Common noun, masculine, in plural form.
Blancos	white	AQAMP0	Qualified adjective, masculine in plural.

Step 4 – Comparison of cases and recommendations

In this step a comparison between the new case and the most similar case retrieved is performed in order to find differences in terms of the sentence grammatical structure. By means of this comparison and the analysis performed is possible to identify for example if there are mistakes of grammatical gender or grammatical number which are common when indigenous people is learning Spanish. For instance the comparison of the example proposed ("Me gustan el gatos blancos" in English "I like white cats") with the case retrieved from the cases library is depicted in Fig. 3.

As a result of the comparison in this example, the system identifies a difference in the third element of the new case (DA0MS0) and the corresponding element of the re-trieved case (DA0MP0). Those tags are described as follows:

- DA0MS0 = Definite article (DA), Neutral (0), Masculine (M), in singular form (S), is not a possessive article (0).
- DA0MP0 = Definite article (DA), Neutral (0), Masculine (M), in plural form (P), is not a possessive article (0).

The difference was identified around the use of the grammatical number: In the new case the article is in singular form, but in the retrieved case (which has been extracted from a text correctly spelled) the article is in plural form. When the mistake has been identified, a recommendation is provided in order to correct the sentence; this recommendation takes information from the case retrieved in the CBR cycle in order to suggest the correct form that the sentence should have. As a result indigenous students and teachers can also learn by interacting with the authoring tool. Fig. 4 shows the graphical user interface of the CBR module. In this case the interface shows the sentence that will be analyzed to identify possible mistakes in grammatical number and gender agreement.

Fig. 3. Comparing the example proposed with a retrieved case from case library.

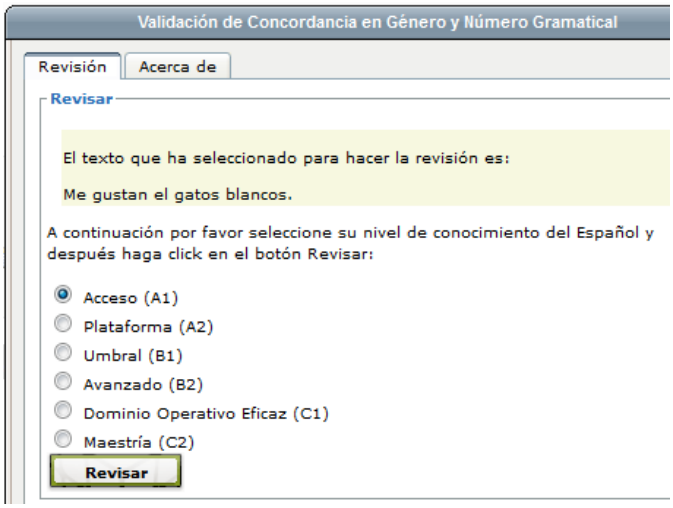


Fig. 4. CBR Module graphical interface in TinyMCE.

V. EVALUATION

A. Description

The purpose of the evaluation process in to validate the main of our approach which is to support indigenous teachers and students when they write a text in a web content editor for creating learning objects. As mentioned before, the support we offer to indigenous teachers and students refers to automatically generate recommendation in order to avoid grammatical errors and develop quality educational content.

In particular, we validate the case retrieval process, because this is the process that ensures that the offered recommendation is the best one that the user could receive.

We applied the K-NN algorithm [13] in order to retrieve the most similar cases from the case library to check the grammatical number and gender agreement.

The K-NN algorithm in the jCOLIBRI framework uses a local similarity function and a global similarity function. The former is used to compute the similarity in every attribute of the cases; the latter is used to compute de global similarity considering the results of the local similarities from all the attributes of the case. We design an experiment to compare and choose the best local similarity functions that allows retrieving the most similar sentence to check the grammatical number and gender agreement. In this section we describe the methodology and the main results of the comparison.

TABLE II
TEST PERFORMED IN THE EXPERIMENT WITH VALIDATION METHOD AND VOTING METHOD APPLIED

Similarity Function	Validation method					
	Leave One Out			N-Fold Random Crossvalidation		
	Weighted Voting Method	Majority Voting Method	Unanimous Voting Method	Weighted Voting Method	Majority Voting Method	Unanimous Voting Method
Levensh.	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Overlap	Test 7	Test 8	Test 9	Test 10	Test 11	Test 12
Smith-Waterm.	Test 13	Test 14	Test 15	Test 16	Test 17	Test 18
Jaccard	Test 19	Test 20	Test 21	Test 22	Test 23	Test 24
Dice	Test 25	Test 26	Test 27	Test 28	Test 29	Test 30
TokensC	Test 31	Test 32	Test 33	Test 34	Test 35	Test 36

B. Methodology

In each case stored in the case library, the attribute with the highest weight is the morpho-syntactic annotation, which is basically a group of tags where each tag has been assigned to each word in the sentence according to the context and the grammatical structure. Since this group of tags is represented by means of a string data type, the local similarity function applied to this attribute should be able to compute the similarity between strings. There are many similarity functions for strings in literature some of them are described in [14], [15]. There are similarity functions based on fuzzy sets [16], and set-based string similarity [17] and [18].

For this experiment we chose four similarity functions commonly used in textual case-based reasoning. In addition we improved two similarity functions to consider the word order of the sentences during the analysis and deal with disambiguation by means of the FreeLing library. These are some important drawbacks described in [19] to be tackled in information retrieval and textual CBR. As a result 6 similarity functions were applied in the experiment. These are listed in table 3.

The validation methods used in this experiment were:

- Leave One Out
- N-Fold Random Cross-validation with 10 folds.

The voting methods selected for the K-NN algorithm were:

- Weighted Voting Method
- Majority Voting Method
- Unanimous Voting Method

The experiment was performed by means of 36 tests that combine the validation methods and the voting methods. Table 2 summarizes the group of tests in the experiment. For each test we obtain the following data:

- Precision
- Recall
- True Negatives
- False Positives
- False Negatives
- True Positives

TABLE III
SIMILARITY FUNCTIONS APPLIED IN THE EXPERIMENT

Simmilarity Function	**Description**	**Improvement**
Levenshtein	Also known as the edit distance	-
OverlapOrdered	Overlap Coefficient	Improved to consider the order of the tokens in each sentence analyzed.
Smith-Waterman	Based on the algorithm of dynamic programming for local alignment.	-
Jaccard	Jaccard coefficient for strings of characters	-
Dice	Dice coefficient for strings of characters	-
Tokens Contained Weighted First	Based on the TokensContained similarity function of jCOLIBRI framework.	Improved to consider the order of the tokens in each sentence analyzed.

C. Results

The following paragraphs summarize the main results we obtained in the experiment.

The results of tests 1,7,13,19,25 and 31 are show in fig. 5. In this case the validation method was Leave One Out and the voting method was Weighted Voting Method. On the other hand Fig. 6 shows the results of tests 4, 10, 16, 22, 28 and 34

where we used the same voting method but using the N-Fold random cross-validation method.

The F-Measure graphic in fig. 5 shows that the OverlapOrdered and TokensContained functions outperform the other functions compared, and the F-measure graphic in fig. 6 shows that Smith-Waterman, OverlapOrdered and TokensContained are functions with a better performance than the others.

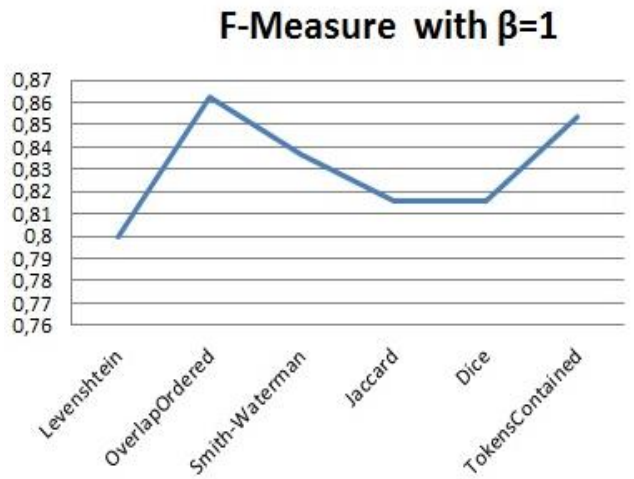


Fig. 5. Results of tests 1, 7, 13, 19, 25 and 31.

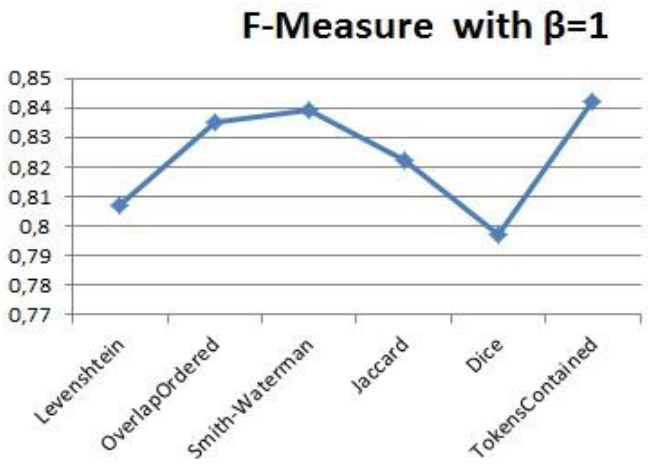


Fig. 6. Results of tests 4, 10, 16, 22, 28 and 34.

The results of tests 2, 8, 14, 20, 26 and 32 are shown in fig. 7 using the Leave One Out validation method as mentioned before, but in this case using the Majority Voting Method for the K-NN classifier.

In contrast fig. 8 shows the results of tests 5, 11, 17, 23, 29 and 35 using the same voting method but using the N-Fold random cross-validation method. In this case the results also shown that OverlapOrdered offers better performance than the other functions evaluated and TokensContained has a good performance when the voting method is the Majority Voting Method.

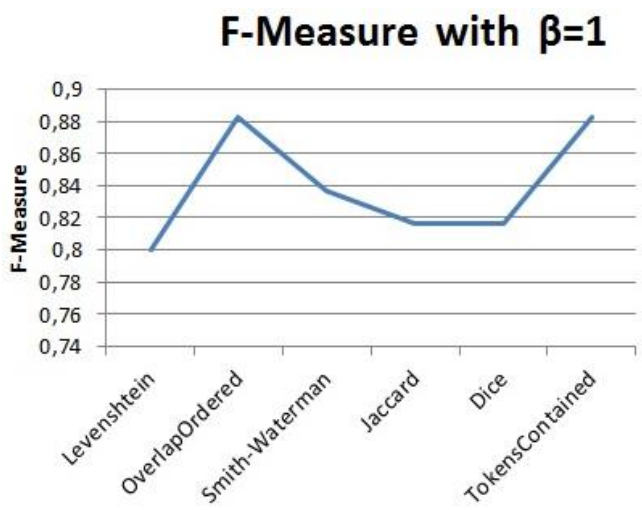


Fig. 7. Results of tests 2, 8, 14, 20, 26 and 32.

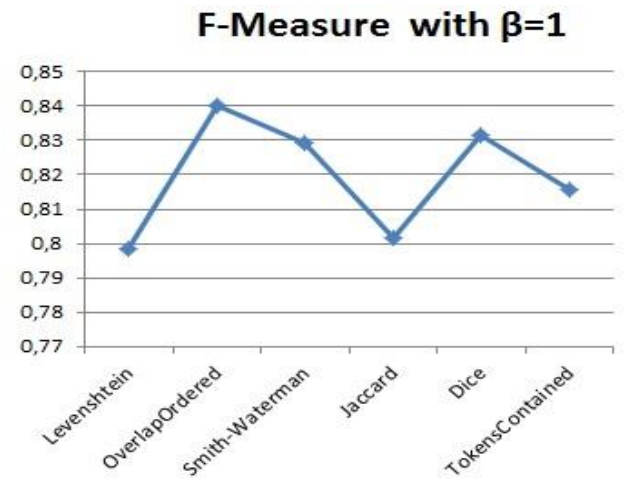


Fig. 8. Results of tests 5, 11, 17, 23, 29 and 35.

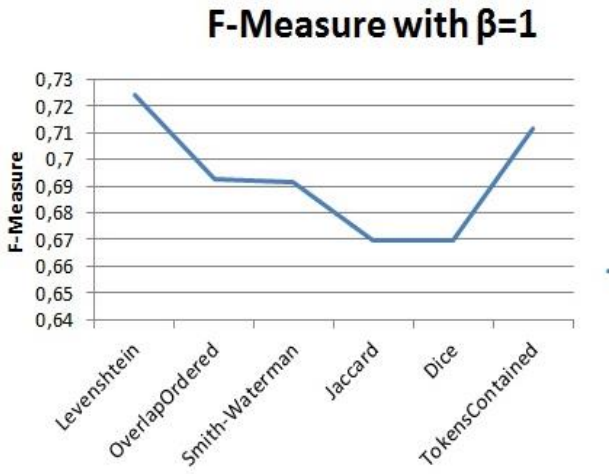


Fig. 9. Results of tests 3, 9, 15, 21, 27 and 33.

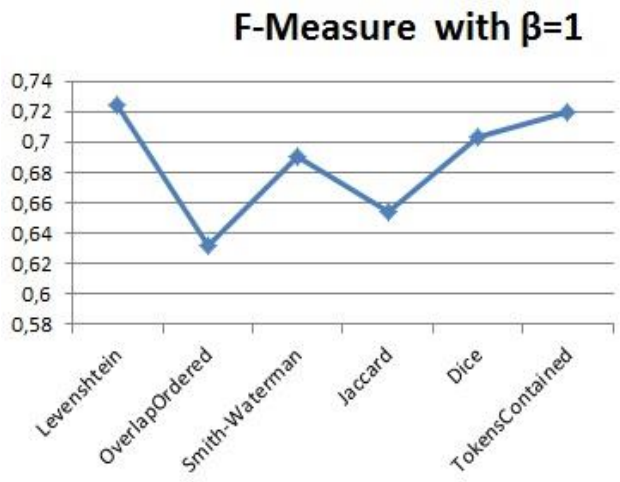


Fig. 10. Results of tests 6, 12, 18, 24, 30 and 36.

Finally the results of tests 3, 9, 15, 21, 27 and 33 are shown in fig. 9, in this case we use the same validation method (Leave one out), but we use the Unanimous Voting Method in the K-NN classifier.

In contrast, fig. 10 shows the results of tests 6, 12, 18, 24, 30 and 36 using the same voting method but the N-Fold random cross-validation method. In this case the Levenshtein distance has a better performance than OverlapOrdered in both methods of validation but Tokens Contained has almost the same performance than Levenshtein.

VI. Conclusion

Multilingual-Tiny as an authoring tool to support indigenous students that will be future teachers of Spanish language when writing texts in Spanish, takes a relevant role in order to help students to improve their writing skills at grammatical level so that they will be proficient teachers of Spanish. Multilingual-Tiny also provides a group of services that allow creating learning objects and design activities in the context of learning Spanish as a second language. This tool can be considered an advance in information and communication technologies to support the training process of indigenous students in the context of bilingual intercultural programs.

The case-based reasoning technique applied to the process of sentence analysis in order to identify grammatical errors mainly in terms of grammatical number and gender, is an efficient technique due to the use of the past user experience. Besides, the similarity algorithm, including the local similarity functions, the global similarity function and the retrieval process based on the K-NN algorithm in JColibri applied in the retrieval step works as it was expected in order to retrieve the most similar cases compared with a new case provided.

Evaluation developed shows that with respect to the algorithms used to retrieval cases that OverlapOrdered and TokensContained are functions with better performance to retrieve cases from the case library, so we can confirm that they are useful when we are dealing with grammatical structures of sentences in Spanish in form of part-of-speech tags. As a result of this experiment we decided to combine both functions when performing the retrieval phase of the case-based reasoning cycle.

This strategy allows improving the system's performance in order to identify possible grammatical errors in gender and number agreement in Spanish language.

Acknowledgment

This work is supported in part by the European Commission through the funding of the ALTERNATIVA Project (DCI-ALA/19.09.01/10/21526/245- 575/ALFA III (2010)88) and for the Spanish Science and Education Ministry in the ARrELS project (TIN2011-23930). The BCDS group (ref. GRCT40) is part of the DURSI consolidated research group COMUNICACIONS I SISTEMES INTELLIGENTS (CSI) (ref. SGR-1202).

References

[1] UNICEF & PROEIB Andes, *Atlas Sociolingüistico de pueblos indígenas en América Latina - Tomo I.* FUNPROEIB Andes, Bolivia, 2009.

[2] UNICEF – *Oficina Regional para America Latina y el Caribe*, [Online], Available: http://www.unicef.org/lac/lenguas_indigenas.pdf

[3] Trillos, M, "Bilingüismo Desigual en las escuelas de la sierra nevada de Santa Marta". THESAURUS. Tomo LI. Núm. 3, pp. 401-486. 1996.

[4] Hamel, R. E., Francis, N, "The Teaching of Spanish as a Second Language in an Indige-nous Bilingual Intercultural Curriculum". Language, Culture and Curriculum. Vol 19:2, pp. 171-188. 2006

[5] Moxiecode Systems AB - TinyMCE. [Online]. Available at: http://www.tinymce.com/

[6] Lier, L. "The Ecology and Semiotics of Language Learning – A Sociocultural Perspective". Springer Netherlands – Kluwer Academic Publishers, Boston. 2004.

[7] Valdiviezo, L. A., "Indigenous worldviews in intercultural education: teachers' construction of interculturalism in a bilingual Quechua–Spanish program. Intercultural Education.". Vol. 21,1, pp. 27-39. 2010.

[8] Arévalo, I., Pardo, K., Vigil, N, "Enseñanza de castellano como segunda lengua en las escuelas EBI del Peru". Dirección Nacional de Educación Bilingüe Intercultural.

[9] University of Toronto, ATutor Learning Management System, [Online], Available: http://atutor.ca/

[10] Padró, L., "Analizadores Multilingües en FreeLing", *Linguamatica*. 3 (2), pp. 13-20. 2011.

[11] Expert Advisory Group on Language Engineering Standards – EAGLES. "EAGLES Guidelines" [Online]. Available: http://www.ilc.cnr.it/EAGLES96/home.html

[12] J.A. Recio-García, B. Díaz-Agudo, M.A Gómez-Martín, and N. Wiratunga, "Extending jCOLIBRI for Textual CBR". In: Muñoz-Ávila, H., Ricci, F. (eds.) Case-Based Reasoning Research and Development. LNCS, vol. 3620, pp. 421-435. Springer, Heidelberg 2011.

[13] K. Yu, L. Ji, and X. Zhang, "Kernel Nearest-Neighbor algorithm", Neural Processing Letters 15, pp. 147-156, 2002.

[14] H. Núñez, M. Sànchez-Marrè, U. Cortés, J. Comas, M. Martínez, I. Rodríguez-Roda and M. Poch," A comparative study on the use of similarity measures in case-based reasoning to improve the classification of environmental system situations, Environmental Modelling & Software", vol. 19, 9, pp. 809-819, Sep. 2004.

[15] N. Elita, M. Gavrila, and C. Vertan, "Experiments with string similarity measures in the ebmt framework". In Proc. RANLP 2007 Conf., Bulgaria, Sep. 2007.

[16] J.J. Astrain, J.R. González de Mendívil and J.R. Garitagoitia, "Fuzzy automata with -moves compute fuzzy measures between strings, Fuzzy Sets and Systems", vol. 157, 11, pp. 1550-1559, Jun. 2006.

[17] M. Hadjieleftheriou and D. Srivastava, "Weighted set-based string similarity", IEEE Data Eng. Bull, 33(1), pp. 25-36, 2010.

[18] P. Achananuparp, X. Hu, and X. Shen, "The Evaluation of Sentence Similarity Measures", In: Song, I., Eder, J., Nguyen, T. (eds.) Data Warehousing and Knowledge Discovery 2008. LNCS, 5182, pp. 305-316. Springer, Heidelberg, 2008.

[19] D. Patterson, N. Rooney, M. Galushka, V. Dobrynin and E. Smirnova, "SOPHIA-TCBR: A knowledge discovery framework for textual case-based reasoning", Knowledge-Based Systems, vol. 21, 5, pp. 404-414, Jul. 2008

Jorge Luis Bacca Acosta is Systems Technologist and Telematics Engineer from the Universidad Distrital Francisco José de Caldas in Colombia, MSc. on Industrial Informatics from University of Girona in Spain, as well as Ph.D Student at University of Girona. He belongs to the BCDS (Broadband Communications and Distributed systems) group from the University of Girona. He is also a developer at the ALTER-NATIVA European Project. His research interest are technology enhanced learning, case-based reasoning, intelligent agents, semantic web and augmented reality.

Silvia Margarita Baldiris Navarro Systems and Industrial Engineer from Industrial University of Santander (UIS), Colombia Master in Industrial Informatics and Automatic from University of Girona and PhD in Technologies from University of Girona. The topics of User Modelling and Adaptations applied to learning management systems, Knowledge management and standardizations in elearning are her research interest. She has participated in several projects: Adaptation based on machine learning, user modelling and planning for complex user-oriented tasks (ADAPTAPlan), EIE-Surveyor: Reference Point for Electrical and Information Engineering in Europe, A2UN@: Accessibility and Adaptation for ALL in Higher Education, and in the European project ALTER-NATIVA.

Ramon Fabregat Gesa Ramón Fabregat is professor at the Computer Architecture and Technology Department at the University of Girona, Spain. He received his degree in computer engineering from the Universitat Autónoma de Barcelona and his doctorate degree in technology from the University of Girona. His research interests are adaptive hypermedia systems, user modeling and e-learning. He has participated to the technical program committees of several conferences. He is director of the master of computer science and automation and leads a research group on BCDS.

Cecilia Avila Garzón is a researcher at BCDS Group from Universitat de Girona (Spain). She graduated in Telematics Engineering from the Universidad Distrital Francisco José de Caldas (Colombia) and has a MSc in Industrial Informatics from the Universitat de Girona. She is a PhD student in the area of technology. His work addresses several topics as web accessibility, collaborative learning, learning objects, and work in communities. She has been researching integrated in the ALTER-NATIVA European Project as a software developer for the last 2 years.

Más de 40.000 alumnos en toda España y Latinoamérica ya confían en nosotros

Másteres ■ Programas Experto ■ Itinerarios Formativos ■ Cursos Universitarios

¿Por qué estudiar en BUREAU VERITAS Centro Universitario?

- Flexibilidad de Horarios
- Infórmate de cómo conseguir tu Máster a través de los Itinerarios Formativos
- Master Class
- Prácticas garantizadas
- Facilidades de pago para todos nuestros alumnos

Tel. : 912 702 178
informacion@bvbs.es
www.bvcu.es

Centro adscrito
UNIVERSIDAD CAMILO JOSÉ CELA

BUREAU VERITAS
Centro Universitario
eLearning

Move Forward with Confidence