Intelligent Learning Environments

Research in Computing Science

Series Editorial Board

Research in Computing Science es una publicación trimestral, de circulación internacional, editada por el Centro de Investigación en Computación del IPN, para dar a conocer los avances de investigación científica y desarrollo tecnológico de la comunidad científica internacional. **Volumen 148, No. 5,** mayo de 2019. *Certificado de Reserva de Derechos al Uso Exclusivo del Título* No.: 04-2005-121611550100-102, expedido por el Instituto Nacional de Derecho de Autor. *Certificado de Licitud de Título* No. 12897, *Certificado de licitud de Contenido* No. 10470, expedidos por la Comisión Calificadora de Publicaciones y Revistas Ilustradas. El contenido de los artículos es responsabilidad exclusiva de sus respectivos autores. Queda prohibida la reproducción total o parcial, por cualquier medio, sin el permiso expreso del editor, excepto para uso personal o de estudio haciendo cita explícita en la primera página de cada documento. Distribuida por el Centro de Investigación en Computación, Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othón de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738, México, D.F. Tel. 57 29 60 00, ext. 56571.

Editor responsable: *Grigori Sidorov, RFC SIGR651028L69*

Volume 148(5)

Intelligent Learning Environments

María Lucia Barrón Estrada
Ramón Zatarain Cabada
Yasmín Hernández
Carlos A. Reyes García (eds.)

Instituto Politécnico Nacional, Centro de Investigación en Computación
México 2019

ISSN: 1870-4069

Instituto Politécnico Nacional (IPN)
Centro de Investigación en Computación (CIC)
Av. Juan de Dios Bátiz s/n esq. M. Othón de Mendizábal
Unidad Profesional "Adolfo López Mateos", Zacatenco
07738, México D.F., México

http://www.rcs.cic.ipn.mx
http://www.ipn.mx
http://www.cic.ipn.mx

Indexed in LATINDEX, DBLP and Periodica

Electronic edition.

ISSN 1870-4069

Editorial

Artificial Intelligence (AI) is present in many aspects of our lives and Education is not an exception. The educational field is changing by continuously developing new theories and methodologies in the learning process, as well as adopting technology to provide personalized learning environments for each student.

Artificial intelligence is part of many computer systems used to teach several topics and the learning environments are moving on providing personalized support to help each student in the learning process, considering previous knowledge, learning styles, emotions and affective states, among other characteristics of students.

Researchers have shown that applying artificial intelligence techniques to educational programs, makes possible to provide software systems that students can use to gain knowledge and learn at their own pace in different virtual or digital environments.

In this volume, we present fifteen research works that apply AI techniques to education in several fields of intelligent learning systems.

At least three members of the Technical Committee reviewed each work to select the chapters presented in this volume. The reviewers took into account the originality, scientific contribution to the field, soundness and technical quality of the papers.

We appreciate the funding provided by RedICA (Conacyt Thematic Network in Applied Computational Intelligence) and we thank its members that were part of the Technical Committee as well as members of Mexican Society of Artificial Intelligence (SMIA Sociedad Mexicana de Inteligencia Artificial). Last, but not least, we thank Centro de Investigación en Computación-Instituto Politécnico Nacional (CIC-IPN) for their support in preparation of this volume.

María Lucía Barrón Estrada
Ramón Zatarain Cabada
María Yasmín Hernández Pérez
Carlos Alberto Reyes García
May 2019

ISSN 1870-4069

Table of Contents

Page

ISSN 1870-4069

A Content Model based on LOM Specification Integrating Learning Disabilities: Toward an Adaptive Framework

Jaime Muñoz-Arteaga[1], Julien Broisin[2], Miguel A. Ortiz-Esparza[1]

[1] Universidad Autónoma de Aguascalientes, Mexico
[2] Institut de Recherche en Informatique, France
jaime.munoz@edu.uaa.mx, broisin@irit.fr,
miguel.ortiz@edu.uaa.mx

Abstract. Learning disabilities can be defined as neurologically based processing problems that interfere with basic learning skills such as understanding, reading or counting, but also with higher-level skills such as space and time coordination. Adaptive learning systems are usually not designed to take into account learning disabilities, even if the need for enhancing support of learners with disabilities is more and more important within nowadays society. We first identify the set of learning disabilities to consider in an accessible adaptive learning system; then a content model is proposed integrating the matching abilities and extending a standard to ensure interoperability with existing solutions. Finally, a case study is presented to apply the proposed model in order to identify a learning object designed to support math skills at elementary school.

Keywords: technology enhanced learning, LOM specification, learning disabilities.

1 Introduction

Technology Enhanced Learning (TEL) aims at supporting learning, and includes both educational and assistive technologies to improve access to educational tools and learning materials, to encourage involvement in learning activities, and to overcome barriers limiting the learning process [7]. Some of these barriers are the Learning Disabilities (LD) that can be defined as neurologically-based processing problems that interfere with basic learning skills such as understanding, reading or counting, but also with higher-level skills such as space and time coordination.

Types of learning disabilities are considered as dyslexia, dysgraphia, dyscalculia, dyspraxia and non-verbal learning disabilities [5]. Learning disabilities can also impact one's social relationships with family, friends or colleagues. They are all of different forms, but they are all learning disorders.

Unfortunately, learning systems are usually not designed to take into account learning disabilities. The approach adopted to address this problem often consists in designing systems specifically intended to learners with LD.

One important risk of this approach is to increase social exclusion of LD learners, as they do not use the systems used by learners without LD. Yet other environments could be designed, such as adaptive systems, to adapt existing learning tools to LD learners and thus to enhance their feeling of belonging to the "regular" learning community.

Adaptive systems are usually composed of four main models [2]: (1) the learner model comprises both domain-dependent and independent characteristics of the learner, (2) the content model describes the learning resources to ensure their mapping with the previous model, (3) the tutoring model includes the adaptive techniques while de- fining what can be adapted, and (4) the user interface model specify the interaction and feedback according to user needs.

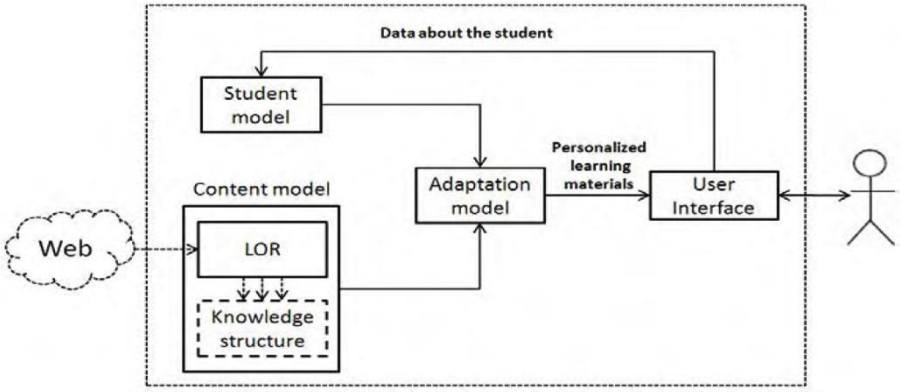


Fig. 1. Architectural model for a recommender system, inspired by [2].

Most adaptive system designed to support learning difficulties don't give information about an effective use in a teaching session [4]. Then teachers spend a lot of time in the selection of educational application according to the learning needs of their students. Learning disabilities are necessary to take into account in particular at elementary school since they reflecting a low academic performance in the child with the possibility of academic desertion [3].

Accessibility aspects can be considered throughout different models of architectural of figure 1, this work focuses only in the content model in terms of learning objects considering learning disabilities, then the content model requires to offer an explicit accessibility specification; This helps a better match with the learner model so that adaptive techniques and algorithms can be further developed [7].

Current work is structured in six sections, section two describes the theoretical background concerning the accessibility and learning object concepts. A related work is presented in section three. The proposal of current work is presented in the fourth section; it is based on standardized initiatives and extensions dedicated to the abilities previously identified.

Next section presents a case study in order to give an overview of existing adaptive approaches regarding learners with disabilities. Finally, in section six, we sum up the proposal and expose the future experimentation designed to evaluate our proposals.

2 Theoretical Background

This sections presents a short description of standard LOM (Learning Object Metadata) is made, starting from the definitions of learning object and its metadata. The term learning disability is also described in order to present some research well known in TEL literature.

2.1 Learning Object

In the LOM standard [6] a learning object is defined as: "any entity, digital or non-digital, that can be used to learn, teaching or training". On the other hand, one of the most common and simple metadata descriptions is that they are considered as "data of the data", this information is mainly used to facilitate the identification, organization and interoperability of learning objects. In fact, the metadata is a kind of interface specifying main characteristics of a learning object [10].

Based on this definition, the metadata of learning objects can be considered as a document that contains structured information by categories and subcategories about an entity that can be used for learning and teaching purposes [6].

The general structure of this standard is composed of nine sections: General, Life Cycle, Metadata, Technical, Educational, Rights, Relation, Annotation and Classification, whose purpose is to allow mainly interoperability between various operating systems, as well as share and reuse the information in different systems or components, through the creation of LOM instances through XML (eXtensible Markup Language). Barker [1] indicates that the design and use of LOM helps us obtain a description that facilitates discovering, locating and acquiring learning resources for teachers, students and automated software processes. It also allows to produce and share the description of resources to adapt to the special needs of a community, thus controlling the vocabulary by classification and reducing the number of elements that are described or added from another resource description scheme.

2.2 Learning Disabilities

The term of accessibility is related to the characteristics of the environments, services and products, which include adaptations necessary to be available to all. The WWW Consortium [13] responds to this need by establishing a web accessibility initiative that provides extensive guidelines and recommendations on the characteristics that certain content must meet to make them accessible and available on various devices. In addition to this initiative, in the document COGA (Cognitive Accessibility User Research) [13], it provides information of users with learning disabilities versus cognitive disabilities, the objective of the description of this information is to provide the characteristics and difficulties in each of these problems in order to consider the accessibility characteristics that must be considered to offer services and technology for this user.

This table is an important guide to take into account in the instructional design of accessible content in terms of learning object, for example a user with any learning disability; it is necessary to consider strategies into the content to reinforce the

memory disorder or visual recognition. In fact, this tables are the starting point to justify the accessibility concept preconized by current work.

Table 1**.** Learning disabilities versus cognitive function according to COGA [13].

Learning Disabilities/ Cognitive Function	Dys-lexia	Aph-asia	Non Verbal	Down Synd-rom	Au-tism	Dys-calculi a	Aging-Related CognitiveD ecline	ADD
Memory	*	*	*	*	*	*	*	*
Executive Functions	*	*		*	*	*	*	*
Reasoning	TNA	*	*	*	*	*	*	TNA
Attention	*	*	*	*	*		*	*
Language		*	*		*			
Speech Perception	*	*	*	*			*	
Understanding figural language	NA	*	*				*	
Literacy	*	*	*	*	*		*	
Visual Perception	*	*	*	*	*	*	*	*
Other Perception	*	*	*	*	*	*	*	*
Knowledge	*	*	*	*	*	*	*	
Behavioral	*	*	*	*	*	NA	*	
Consciousness	TNA	*	*			NA	*	

*TNA= typically not affected, NA= not affected, ADD = Attention Deficit Disorder.

3 Related Work

We have conducted a literature review focusing on content models that could be used by adaptive learning systems, to study in which extent they take into account learning disabilities [7].

Nowadays, many knowledge pools have been built to enable sharing and reuse of learning materials. Most of them implement the LOM [3] standard to describe the content they offer, which provides a controlled structure and vocabulary to expose details about the properties of a learning object. Since this standard has not been designed to support learning disabilities, several proposals emerged to integrate some of the facets of LD: Karampiperis and Sampson [11] proposed an application profile focusing on accessibility. Other approaches have been built, such as TASS [3] and PBAE [8], to facilitate the identification of accessible resources by considering the properties defined by the IMS standard, but previous works have not taken into account in an explicit manner the specification of learning disabilities into the learning object; current work represent a basis for designing a content model supporting learners with LD.

4 Content Model

In order to allow systems to self-adapt according to the above learner model while preserving reuse of content stored into existing repositories, we propose a content model based on the LOM standard [6] and comprising a new category dedicated to accessibility, this expresses the level of ability required by a learner to use efficiently the learning resource. We propose in the figure 2 an extension for LOM specification for metadata of learning objects is inserting a new category called Accessibility with the COGA criteria described in the table 1. Then, the possible values for the field titled "learning disability" such as: *Dyslexia, Aphasia, Non Verbal, Down Syndrome, Autism, Dyscalculia, Aging related Cognitive Decline, ADD (Attention Deficit Disorder)*. For the field "Cognitive function" one of values are considered: *Memory, Executive Functions, Reasoning, Attention, Language, Speech Perception, Understanding figural language, Literacy, Visual Perception, Other Perception and Knowledge.*

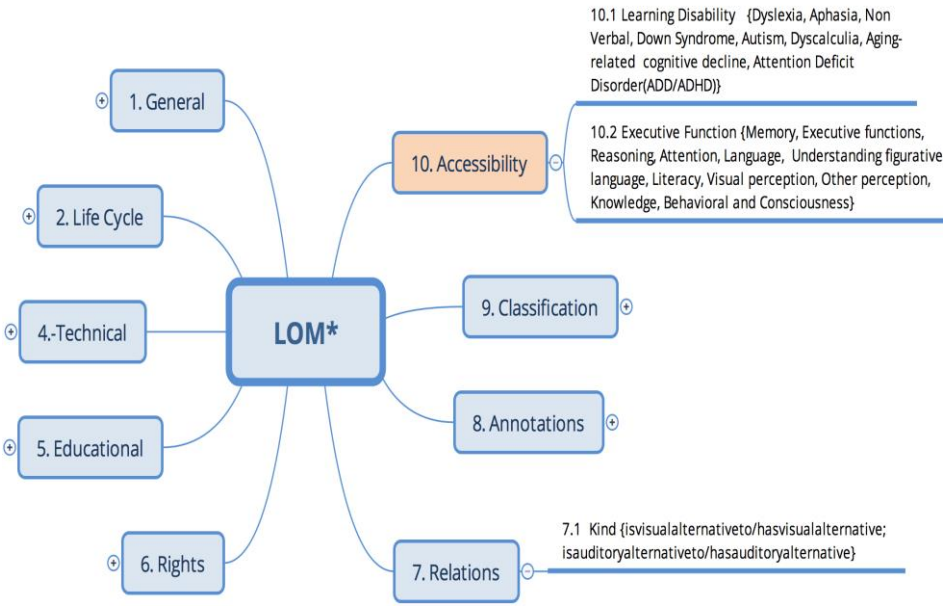


Fig. 2. Extension of IEEE LOM as a content model considering the accessibility aspect.

Also, to allow an adaptive learning systems to deliver alternative resources to learners according to their preferences, we extended the Relation category of the LOM standard. Indeed, this primary objective of this category is to depict the different kinds of relations learning resources might have between them. Therefore, to express the fact that a given learning material represents an alternative to a visual or auditory resource in a similar way of Sampson [11], we defined the new *isvisualalternativeto/hasvisualalternative* and *isauditoryalternativeto /hasauditoryalternative kind of Relations category*. Table 2 presents this information.

Table 2. Specification for a visual or auditory content to consider in the *Relations* category.

#	Name	Explanation	Size	Order	Value space	Data Type	Example
7	*Relation*	This category defines the relationship between this learning object and other learning objects, if any.	LangString (Smallest permitted maximum 100 items)	Unordered			
7.1	*Kind*	Nature of the relationship between this learning object and the target learning object.	1	Unspecified	*isviualalternativ to/hasvisual alternative; isadtoryalterntivto/hasauditoryalternative*	Vocabulary (*state*)	*"isvisualalternativeto/hasvisualalternative"*

5 Case Study

Current proposal has applied in the following case study: a special education teacher needs to teach a student of eight years old under third grade of primary school. Dyscalculia has been the result of diagnosis conducted by the teacher, since this student has some difficulties for the identification and manipulation of numbers, in particular the child has a limited ability to grasp the concept of value of money. One of proposed solution by the teacher is one activity where the student gets skills to select, compare and buy items in a corner store.

Fig. 3. User interface of learning object "La Tiendita".

The teacher has used an educational recommender system using the proposed content model to look for a learner object to support this activity, an application called "La Tiendita" is the answer given by the system (see figure 3). This is because people with dyscalculia often buying far too much or not-nearly enough because it is difficult for them to work out exactly how much they need. In order to design

personalized functionalities such as recommendations for different user types in this case study, the recommender system should first establish a user model to represent users' profile, as well as current content model to organize and classify the resources accessed by users. It is possible to enrich their semantics in a similar approach of work [12], such models can be expressed in terms of ontology concepts.

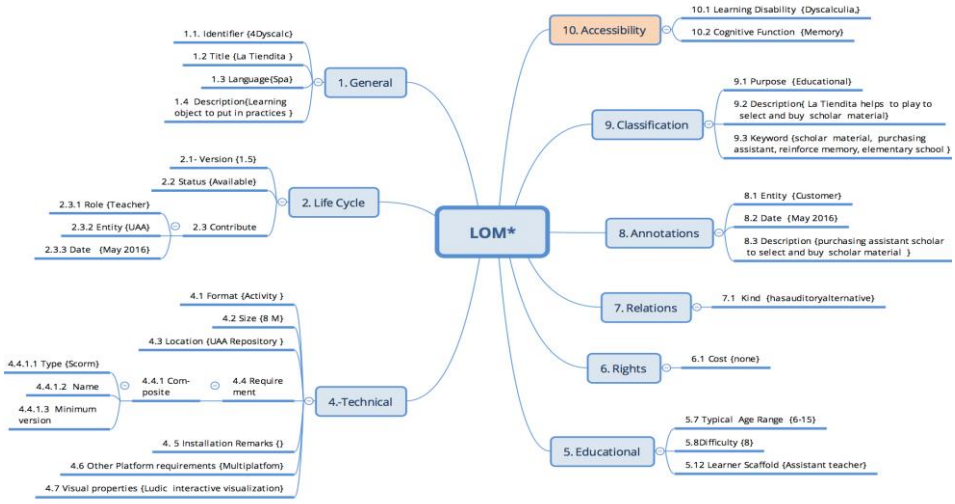


Fig. 4. Content model for the learning object *"La Tiendita"*.

The proposed content model of figure 4 describes the characteristics of learning object "La Tiendita", note the accessibility category is described in terms of a learning disability and the cognitive function where dyscalculia is the value of learning disability and for the field cognitive function is the reasoning and memory.

The student has used the learning object called "La Tiendita" to select and play to buy scholar material such as notebooks, pencil, eraser, scissor, backpack, etc… In addition, the application helps to get and confirm an accurately payment in cash, in case of error a message is sent to user in order to give a better solution and get a successful result (see figure 3).

With all this information, the teacher can use the recommender system and begin with the search of learning object that allow, for example, to approach visual perception or another cognitive function. In this sense, it is possible to illustrate the need to use this extension of LOM, since the keywords that the teacher uses are related to learning problems and cognitive functions, and it is possible to find only the resources related to that category or subcategories.

6 Conclusions and Future Work

This paper has proposed a content model in order to design adaptive learning systems able to take into account learners with disabilities. The proposed content model extends existing standardized initiatives to ensure interoperability with existing

solutions, and stands on the concept of learnability to consider abilities that affect the learning process [7]. There are several initiatives in progress as future work to develop the tutoring model. The adaptive rules exploit the usability features of both the learner and the content models to apply strategies for the identification of appropriate resources that can be recommended and delivered to learners according to their LD profile. For this, several user evaluations can be conducted in order to evaluate the capacity of our adaptive system to deliver pertinent resources to children in order to increase their reading skills; tutors of the children will be responsible for filling the initial profiles of learners as well as ensuring the evaluation of our adaptation process, whereas a repository [9,10] comprising a set of learning objects will be used as the source of learning content.

References

1. Barker, P.: What is Ieee learning object metadata/ims learning resource metadata. Cetis standards briefings series (2005)
2. Butz, C., Hua, S., Maguire, R.: A web-based bayesian intelligent tutoring system for computer programming. Web Intelligence and Agent Systems, Vol. 4, pp. 77–97 (2006)
3. Green, S., Jones, R., Pearson, E., Gkatzidou ,S.: Accessibility and adaptability of learning objects: responding to metadata, learning patterns and profiles of needs and preferences. ALT- J, Research in Learning Technology 14(1), pp. 117–129 (2006)
4. Bustos-López, M., Vásquez-Ramírez, R., Alor-Hernández, G.: An Architecture for Developing Educational Recommender Systems. Research in Computing Science 106, pp. 17-26 (2015)
5. Hammill, D.: On Defining Learning Disabilities: An Emerging Consensus. Journal of Learning Disabilities 3(2) (1990)
6. IEEE Learning Technology Standards Committee (LTSC). Draft Standard LOM (2002)
7. Erdt, M., Fernández, A., Rensing, C.: Evaluating Recommender Systems for Technology Enhanced Learning. IEEE Transactions on Learning Technologies 8(4) (2015)
8. Jemni, M., Laabidi, M., Ben Ayed, L.: Accessible E-learning for Students with Disabilities: From the Design to the Implementation. In: Huang, R., Kinshuk, Chen, NS. (eds) The New Development of Technology Enhanced Learning. Lecture Notes in Educational Technology. Springer, Berlin, Heidelberg (2014)
9. Margain Fuentes, L., Muñoz Arteaga, J., Santaolaya Salgado, R., Alvarez Rodriguez, F., Vanderdonckt, J., Orey, M.: MIRROS: Intermediary Model to Recovery Learning Objects. Revista Computación y Sistemas 13(4) (2010)
10. Muñoz Arteaga, J., Alvarez-Rodriguez, F., Chan-Nuñez, M.E.: Tecnología de objetos de Aprendizaje. Editorial Universidad Autónoma de Aguascalientes & UdG Virtual (2007)
11. Sampson, D., Fytros, D., Zervas, P.: Supporting Lifelong Learning Programmes: Defining an Accessibility and Competence Based Application Profile for Educational Metadata. In Proceedings of the 11th IASTED International Conference, Vol. 614, No. 125, pp. 343–348 (2008)
12. Triomphe, M., Vidal, P., Broisin, J.: Opening Learner Profiles across Heterogeneous Applications. In: ICALT 2009, IEEE Computer Society, pp. 504–508 (2009)
13. W3C, Cognitive Accessibility User Research. In: Web Accessibility Initiative (WAI), http://www.w3.org/TR/coga-user-research/ (2018)

ISSN 1870-4069

A Model for Identifying Steps in Undergraduate Thesis Methodology

Samuel González-López[1], Aurelio López-López[2], Steven Bethard[3], Jesús Miguel García-Gorrostieta[2]

[1] Technological Institute of Nogales, Sonora, Mexico
[2] National Institute of Astrophysics, Optics and Electronics, Tonantzintla, Puebla, Mexico
[3] University of Arizona School of Information, Tucson, Arizona, USA
samuelgonzalezlopez@gmail.com, allopez@inaoep.mx, bethard@email.arizona.edu, jesusmiguelgarcia@inaoep.mx

Abstract. Knowledge generation is an important asset of great economic powers, and knowledge societies are a fundamental part in the development of countries. Mexico is a country that is in the process of development and improvement of its education system, according to the Educational Reform promoted since 2012 by the Federal Government. We identified an area of opportunity at the undergraduate level to help improve the writing of students, specifically in draft theses and research proposals. This work focuses its efforts on analyzing with natural language processing techniques the "Methodology" section, an important element for the development of a thesis, that helps the reader to understand if the techniques and data used are appropriate in an investigation. This paper proposes a Model to identify a series of steps in such a section. In addition, preliminary results of a basic exploration of a collected corpus are presented, pre-processing the text to generate a representation according to Language Models. The corpus contains documents of graduate and undergraduate levels in the computer science and information technologies domain. The preliminary results showed that the information extracted from the corpus serves to adequately differentiate the methodologies of both levels.

Keywords: automated text evaluation, natural language processing, corpus creation, language models, methodology analysis.

1 Introduction

Knowledge generation is an important asset of great economic powers, while knowledge societies are a fundamental part in the development of countries. Mexico is a country that is in the process of development and improvement of its education system, as confirmed by the Educational Reform approved in December 2012 by the Federal Government. This reform establishes that the

State will provide the educational materials and methods, the school organization, and the educational infrastructure for the continuous improvement and the maximum educational achievement of the students.

One of the relevant principles in this reform is that all public and private sectors must collaborate so that education improves and achieves a high quality[4]. Under the dynamics of this principle of improvement and in accordance with the educational reform, an area of opportunity has been identified at the undergraduate level, which includes the support in the writing of documents of students who are finishing their educational program, specifically in documents such as theses and research proposals.

In this paper, we present a component, part of a wider project, which seeks to provide a series of tools to the student. Using these tools, the students can analyze and evaluate their texts, and obtain feedback to improve their writings [5]. In this study, we seek to analyze the element of "Methodology". The methodology has steps and procedures used to develop the research which should provide a step-by-step explanation of the aspects necessary to understand and possibly repeat the research [8]. The methodology should include: the techniques and procedures employed, type of research, population studied, the sample, collection instruments and description of the selection of data, description of the validation instrument, and a description of the statistical analysis process.

In this study, first we analyze one of the elements, specifically the sequence of steps. This feature is implicitly related to the first point since, when showing a procedure, it is presumed that there are a series of steps. However, in this work, the proposed method does not analyze the content. In the following list, we show the elements to be identified to indicate a series of steps. In addition, we identify the actions expressed by the verbs used in the methodology, as well as if there is a logical sequence in the use of verbs. In particular, we focus on:

- **Series of steps:** ordered activities to be carried out in the methodology by applying some technique. They are not necessarily expressed by numbering.
- **Verbs:** words that represent the actions to be performed in the series of steps.
- **Logical sequence:** The student is expected to use verbs with a hierarchical order. Example of expected logical sequence: it would appear first the verb "to explore" (in Spanish "explorar") and then the verb "to implement" (in Spanish "implementar").

The paper is structured as follows. In section 2, we present related work in automatic writing assessment and machine learning. The collection collected and used in the first experiments, and the proposed model to analyze a series of steps in student theses is detailed in section 3. The results of the experimentation are shown in section 4. In section 5, we conclude with some final remarks.

[4] http://www.presidencia.gob.mx/reformaeducativa/#sobre-la-reforma

2 Related Work

The Automated Writing Evaluation (AWE), also called Automated Essay Scoring (AES), refers to the process of evaluating and scoring written text using a computer system. This kind of system builds a scoring model by extracting linguistic features (lexical, syntactic or semantic) on a specific corpus that has been annotated by humans. For this task, researchers have used Artificial Intelligence techniques such as natural language processing and machine learning methods. The system can be used to directly assign the score or the quality level of a student's text [4]. The use of AWE systems offers students a way to improve their writing during the document review process.

AWE systems help to reduce the review time dedicated by academic instructors, i.e., they are complementary tools to the reviewer's work. Currently, advances in AWE systems include the use of Natural Language Processing technologies to perform the evaluation of student texts and provide feedback to students. Under this context, the Writing Pal (WPal) system offers instructions and practice based on game theory. The WPal system evaluates the quality of the essay using a combination of linguistic computation and statistical models. The authors of this system selected different linguistic properties that were used as predictors [2]. In a similar way, our work evaluates the quality of the text but focusing on the methodology section.

The Machine Learning approach has been used to assess student essays, with the aim of finding the main topic and conclusion in essays [1]. The authors used two annotators to identify the main topic and the conclusions section. Among the features used to train the algorithm are words and phrases. For example, the phrase "in conclusion" is associated with the conclusions section. Another feature considered is the position of the text. In contrast, in our corpus, we assume that the methodology is clearly delimited by a subtitle. Our proposed method includes sending recommendations to students to improve the methodological approach of their project. In previous work, we have presented a system called TURET (Tutor for the Writing of Thesis -in Spanish "Tutor para la Redacción de Tesis"), which analyzes the lexical richness of seven sections of a thesis, applying natural language processing techniques [7].

In the phrase extraction approach, scientific articles similar to a thesis have been studied. In the work of [9], a sentence extraction method was developed to identify the most relevant phrases, which define the document and differentiate it from other types of documents. In a similar way, our work identifies elements that represent the methodology and capture it through a language model.

3 Method

A collection was created using the ColTyPi site, which stores theses and research projects from the area of information technology. In Table 5, we present the corpus extracted from ColTyPi [6].

The graduate level is composed of Doctoral and Master theses. The Undergraduate level is composed of Bachelor and Advanced College-level Technician

Table 1. Collection of methodology texts.

Sets	Graduate Level	Undergraduate Level
Train	90	56
Test	18	14
Total	108	70

(TSU) theses. The theses and research proposals of the collection have been reviewed at some point by a review committee. ColTyPi includes texts in Spanish in the area of Information Technologies. However, the proposed method does not limit to thesis of this area, but it does have a qualitative cut. In Table 2, we show an example of an item in the collection labeled by an annotator (person with experience in reviewing thesis).

Table 2. Annotated (translated) example of Methodology Collection [Graduate Level].

Element	Methodology	Logical Sequence
1	To develop the proposed work, a set of steps was followed to ensure each of the objectives presented. The following are the needs surpassed for the development of the research: *1.* ***To compile*** *the Bibliographic and detailed analysis of existing disambiguation approaches.* *2.* ***To characterize*** *the language families and their relationship with the Spanish language.*	Yes

Table 2 shows three features that a methodology must include. The "series of steps" in italics, the "verbs" used (in bold), and "YES" in the logical sequence column. Observe that the methodology contains the three elements, however, sometimes the methodology does not show a series of steps and even without a numbering. The collection used in this study has been tagged by two annotators with experience in reviewing theses, reaching F-scores of 0.9 and 0.89 in Set of Steps and Verbs respectively, and Kappa of 0.46 (Moderate) in Logical Sequence.

Figure 1 shows the main components that are proposed to identify a series of steps in the Methodology section. The central part of the Analyzer is the language model, which captures the main characteristics of the different methodology sections in the collection. Below, we detail the components used in the proposed method.

Transitional Devices (TD): During the corpus analysis, the use of specific terms have been observed in the methodologies. These devices work as a connection bridge between sentences. The online writing lab at Purdue University [5] identifies the following categories: add, compare, test, show an exception, show

[5] https://owl.english.purdue.edu/owl/resource/574/02/

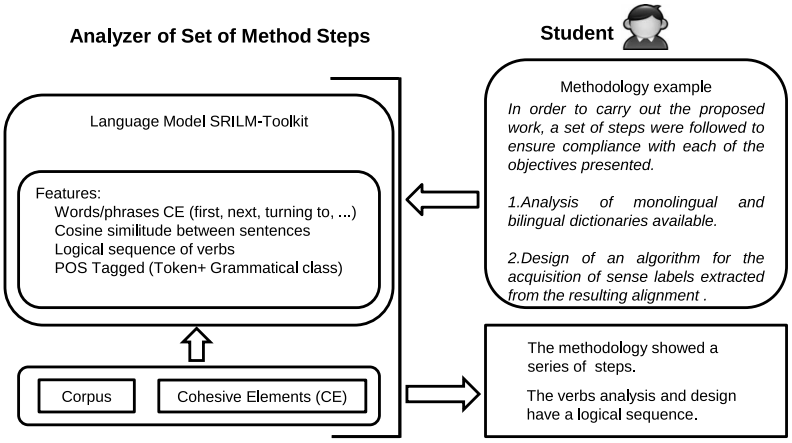


Fig. 1. Analyzer Model "Series of steps in Methodology".

time, repeat, emphasize, show sequence, give an example, summarize. We focus on the "Sequence" category, presented in Table 3.

Table 3. Transitional Devices in Methodology.

Category	Transitional Devices
Sequence	First, second, third, next, then, following, at this point, now, before, later at this moment, subsequently, finally, consequently, previously simultaneously, therefore.

For example, the cosine similarity between the two weighted term vectors of sentences in (1) is 18.26%. In a series of steps, it is expected to find this type of result, thus giving a series of transitions of similarity between the steps.

(1) 1. To compile the Bibliographic and detailed analysis of existing disambiguation approaches.
2. To characterize the language families and their relationship with the Spanish language.

Logical Sequence of Verbs: This feature captures the order of the verbs used in the series of steps. We used the Bloom Taxonomy , consisting of three hierarchical models (cognitive, affective and psychomotor). We employed the cognitive model. In the work of [3], they proposed an adapted taxonomy for the computer tools implementation.

Part of Speech (PoS) Tagging (Term+Grammatical Category): The grammatical position of each term within sentences captures some patterns in the use of the terms, specifically in Methodology. For example, the use of a verb

(possibly appearing at the beginning of sentence) that indicates the action to be performed in the series of steps.

Language Model "Series of Steps in Methodology": The Statistical Language Model extracts models that estimate the probability of word sequences. In our approach, each training element of the language model contains the features: transition devices, cosine similarity, logical sequence of verbs and the grammatical categories (PoS tags). The objective of the model is to capture information that identifies a series of steps. To build the model, we employed the SRILM [6] tool. With the trained model, we expect to evaluate new methodology formulations of students. It is worth mentioning that each feature listed provides a value, which is part of the vector of features that the model takes for training.

4 Results

In the first stage of development of the proposed model to identify a series of steps, the corpus validation was carried out. This validation was with the objective of identifying if the corpus of graduate and undergraduate levels differed. Otherwise, the corpus would not be useful to construct the language model with the features described in Section 3. In this first experiment, we used the corpus detailed above and the SRILM tool for the model construction.

The first step was the construction of three language models, taking into consideration only one of the characteristics of PoS Tagging. The first model was built using the graduate level training group consisting of 90 elements, the second language model was built with 56 elements of the undergraduate level, and the third group was built with 70 conclusions. The third group was created to test the behavior of models 1 and 2 with different texts, out of the corpus of methodologies.

Each methodology element was lemmatized, to obtain the root of the word using the FreeLing tool. The purpose of lemmatizing is to obtain more coverage when training the language model and minimize the burden of processing. We employed sequences of 2, 4 and 8 terms (n-grams) for the training with a smoothing value of 0.01 (to avoid the effect of null probabilities). The bigrams model (two terms) achieves the best performance, having the lowest values of perplexity. A low value of perplexity indicates that the language model is capturing the analyzed sequences.

The model comparison was developed by evaluating the test sets in the models, with the hypothesis that each model would better identify the test set corresponding to its level. Figure 2 shows the results achieved when evaluating the test groups in each model.

In Figure 2, we observe that the graduate model and the undergraduate model have a similar performance (25 and 28 of perplexity, respectively). Between these two models, the graduate model has a lower average perplexity in tests performed with graduate texts, while the undergraduate bigram model has

[6] www.speech.sri.com

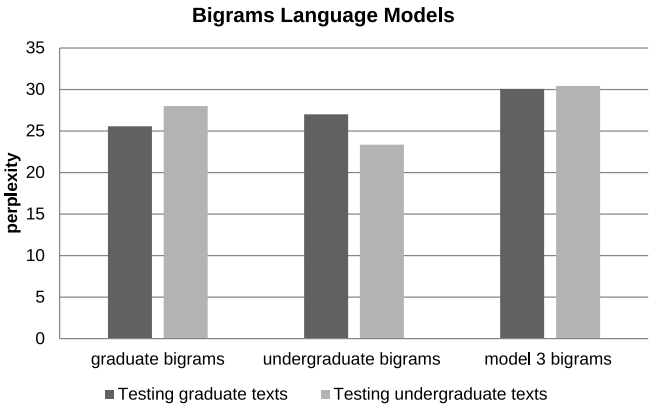


Fig. 2. Comparison chart between bigram models.

a lower perplexity in tests with undergraduate texts. This result is expected since each model obtained low perplexity with the sets of the same level, that is, the models are recognizing their respective level. The third model shows the highest average perplexity (30 and 31 of perplexity) with both graduate and undergraduate texts, which indicates that the models can be differentiated between the texts of each level of different texts.

In Figure 3, which shows perplexity for each undergraduate text in the test set, we can notice that the "bigrams 3 model" (in green), obtains high values which indicates that is not capturing the sequences as well as the "bigrams model of undergraduate", which achieves lower values of perplexity. Therefore, bigrams model of undergraduate responds better to its corresponding set, that is, of undergraduate level.

5 Conclusion

In this paper, we presented a method to evaluate the methodology section of students writings. The results of these first experiments give us a guideline for the development of further methods. We observed that although the distances between the perplexity results of the models were small, they were however consistent in the test set. Therefore, the differences found between graduate and undergraduate groups with the n-gram language model gives evidence that in a future stage of experimentation, the proposed model will be feasible.

During the qualitative analysis of the methodology statements of both levels, we observed that some of them did not present a list of numbered steps, rather, a writing in separate paragraphs. Our model expects to identify this type of methodology statements since transitional devices are expected to be found.

Finally, in future work, we expect to develop the proposed method using the annotated corpus to perform agreement tests between the annotators and

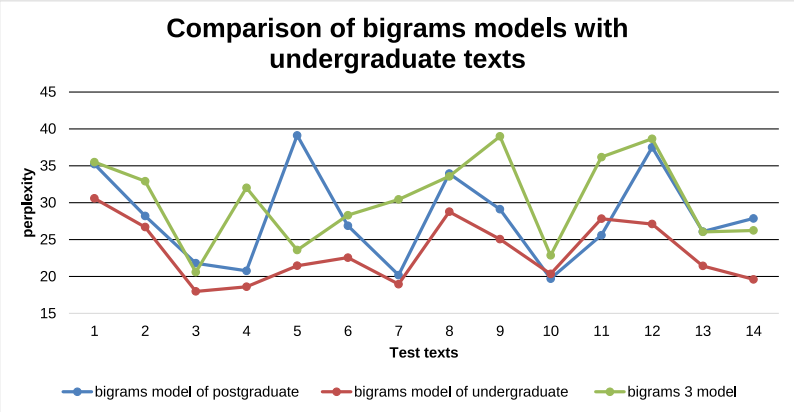


Fig. 3. Comparison between models with the 14 undergraduate texts in the test set.

the analyzer. We also plan to integrate the proposed method into a TURET2.0 system to analyze and evaluate student drafts online.

References

1. Burstein, J., Marcu, D.: A machine learning approach for identification thesis and conclusion statements in student essays. Computers and the Humanities 37(4), 455–467 (2003)
2. Crossley, S.A., Varner, L.K., Roscoe, R.D., McNamara, D.S.: Using automated indices of cohesion to evaluate an intelligent tutoring system and an automated writing evaluation system. In: International Conference on Artificial Intelligence in Education. pp. 269–278. Springer (2013)
3. Fernández-Sánchez, P., Salaverría, A., Perez, E.M., Valdes, V.: Taxonomía de los niveles del aprendizaje de la ingeniería y su implementación mediante herramientas informáticas pp. 522–527 (2012)
4. Gierl, M.J., Latifi, S., Lai, H., Boulais, A.P., De Champlain, A.: Automated essay scoring and the future of educational assessment in medical education. Medical education 48(10), 950–962 (2014)
5. González López, S., López-López, A.: Supporting the review of student proposal drafts in information technologies. In: Proceedings of the 13th annual conference on Information technology education. pp. 215–220. ACM (2012)
6. González-López, S., López-López, A.: Colección de tesis y propuesta de investigación en tics: un recurso para su análisis y estudio. In: XIII Congreso Nacional de Investigación Educativa. pp. 1–15 (2015)
7. González-López, S., López-López, A., García-Gorrostieta, J.M., Espinoza, I.R.: Turet2. 0: Thesis writing tutor aimed on lexical richness in students' texts. Intelligent Learning Environments 129, 9–17 (2016)
8. Muños Razo, C.: Como elaborar y asesorar una investigación de tesis. Libertad: Pearson México (2011)
9. You, W., Fontaine, D., Barthès, J.P.: An automatic keyphrase extraction system for scientific documents. Knowledge and information systems 34(3), 691–724 (2013)

ISSN 1870-4069

A Web-based Didactic Tool for Teaching of Distributed Consensus

Abdiel González-Ortega, Francisco de Asís López-Fuentes

Universidad Autónoma Metropolitana-Cuajimalpa, Mexico City, Mexico
abditiger123@gmail.com, flopez@correo.cua.uam.mx

Abstract. Currently, educational models face the challenge of using information and communication technologies (ICT) to provide their students with the necessary tools and required knowledge in the 21st Century. The study of distributed systems is increasingly its important due to the popularity of web applications. For this reason, we address in this paper a topic known as the Byzantine Generals problem, which is a topic that has boomed in recent years, with the emergence of Bitcoin and Blockchain. This paper presents a web tool to support the teaching-learning process, which is given between the teacher and the students. The objective of this tool is to familiarize students with notions of consensus algorithms. Using our didactic tool users can visualize the operation of the Byzantine Generals algorithm and improve their knowledge about this consensus case.

Keywords: distributed systems, Web 2.0, byzantine generals, fault tolerance.

1 Introduction

Computer systems are already immersed in virtually all human activities. In particular, real-time systems are present in increasingly complex tasks and where an error can lead to catastrophic situations. Therefore, the fault tolerance capabilities of this type of systems are important for its success throughout its life cycle. Different characteristics of real-time systems must be converted to make them tolerant to failures, which is why there is a large field of research and development.

At present, the use of large distributed systems is becoming more common. For example, online services on the internet, cloud computing, internet of things and artificial intelligence are booming. Many of these systems are distributed systems that need to be scalable, because they constantly add and remove nodes (servers and clients). Distributed systems require being tolerant to failures to guarantee the service both to the current nodes and to those new nodes that want to join the system [1]. In this context, the concept of "distributed consensus" is very important to guarantee reliability in the system.

This paper presents a didactic tool to support the study of distributed consensus in the distributed systems. Our didactic tool simulates the case of Byzantine generals for different numbers of nodes with different numbers of traitor generals in order to know if the system can reach a consensus (agreement) [8].

The rest of this paper has the following organization. Section 2 gives an overview about related work, while an introduction to Byzantine general algorithm is given in Section 3. Design and implementation is given in Section 4. Operation and evaluation of our didactic tool is presented in Section 5. Paper concludes in Section 6.

2 Related Work

The main topics addressed in this work are two. First issue is about importance of information technologies for teaching-learning process and second issue is about the Byzantine Generals Problem. Several interactive tools can be found in the literature. Interactive Tools for Learning Sensor Network Basics [3] is a tool used to show the nodes in a network as intelligent autonomous sensors, which can measure certain characteristics of their environment, such as temperature, pressure, humidity, acceleration, etc. Wireless sensor networks [4] have a wide range of applications: climate monitoring, flood prevention, seismic monitoring, early detection of forest fires, etc. This tool is useful for master's students interested in wireless sensor networks as a research topic for their thesis, but who do not know enough to make a rational choice. This led the authors to start implementing an easy-to-use multimedia e-learning course, explaining the essential aspects of sensor networks and their routing protocols.

It has been observed that students, who have the opportunity to freely experiment with the systems learn faster and with more depth than those who do not have such opportunities. For that reason, a very simple interactive simulator that shows the principles of routing is included in the e-learning course.

One of the systems in which the fault tolerance of the Byzantine generals is being used is the Boeing 777 aircraft (through its ARINC 659 SAFEbus network) [5]. The flight control system of the Boeing 777 and the Boeing 787 flight control systems, use tolerance to Byzantine faults. Because these are real-time systems, Byzantine fault tolerance solutions must have very low latency. For example, SAFEbus can achieve tolerance to Byzantine faults with an order of one microsecond of added latency. The Boeing 777 is the first com-mercial aircraft manufactured by Boeing that employs the fault tolerance of Byzantine generals as the primary flight control system.

3 Byzantine Generals Algorithm

The problem [6] is that some of the generals are traitors and have as their objective that the armies be defeated. Therefore, our objective is to define an algorithm that allows loyal generals reach a consensus on the action plan. The final decision will be by majority vote on the initial choices. If they tie, the decision is withdrawn. The problem assumes that the generals are computers and the messengers are communication channels. Generals may fail, but not messengers. The requirements to obtain consensus are:

- Termination: all non-defective processes must decide [2].
- Agreement: the final decision of each non-defective process must be identical.

– Validity: if every non-defective process starts with the same value (V), its final decision must be (V).

The problem of the Byzantine generals is frequently referred in the failure tolerance, and it is also known as a Byzantine problem. Lamport et al. [2,7] demonstrated that in a system with *k* defective processes, consensus can be achieved only if *2k + 1* processes are functioning correctly, for a total of *3k + 1* processes.

3.1 Case with Three Generals

Figure 1 shows the basic example of Byzantine Generals problem. This scenario consists of three generals, which are represented in a hierarchical way. There is a commander (there can only be one) who sends the orders to the general lieutenants that in this first case we have two lieutenants. We note that the generals with the rank of lieutenant cannot send messages to the commander or return those received, because the final agreement is carried only with the generals of the same rank as lieutenants and only if they are loyal generals. In this scenario, three nodes are shown as loyal generals. We see that the general who has the commander hierarchy is who sends the message to the other lieutenants in this case Lieutenant1 and Lieutenant2. The messages that the commander sends are "Attack" given that the commander is a loyal general.

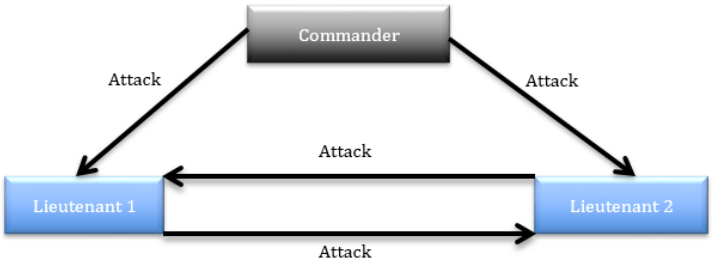


Fig. 1. Basic model of the Byzantine generals.

Consequently, Lieutenant generals send messages to each other because, as previously explained, the Lieutenant generals cannot communicate with a superior general. Lieutenant1 sends Lieutenant2 the message received by the Commander and consequently Lieutenant2 sends Lieutenant1 the message he received. In this scenario it is assumed that all generals are loyal.

As shown in Figure 1 we have the base case which consists in that all generals are loyal. The general who will be responsible for deciding if consensus is reached will be one of the loyal lieutenants, never a traitor general will have that responsibility to make the decision final in this case can be both Lieutenant1 and Lieutenant2 since both are loyal. After receiving the message from the commander and the lieutenants have finished sending their messages, an evaluation is made of how many are in favor of attacking and how many in favor of retreat and depending on the results of those messages will be how an agreement is reached. In this case, we have that the Lieutenant1 will receive the message from the commander who has sent a message to

"Attack" and the other from the Lieutenant2 who in turn also sends the "Attack" message, resulting in the Lieutenant1 reaching an agreement to attack.

Now, we analyze what happen when we have a traitor general in the group. Figure 2 shows this scenario. Communication between generals is same and the hierarchy of them is also maintained, the only variant in the scheme is the traitor node. The Commander sends the message "Attack" to the two Lieutenants, but here we see that one of them is a traitor in this case the Lieutenant2. Both lieutenants have already received the message. Then, both nodes exchange its messages together. Lieutenant1 as a loyal general sends the message "Attack" received from the Commander. However, Lieutenant2 is a traitor, then this general changes the message received from the Commander and sends a false message ("Retreat") to Lieutenant1. In this scenario, each Lieutenant receive two messages, one from Commander and the other from another Lieutenant. We can see that Lieutenant 1 receives from Commander the order to "Attack", while from the traitor Lieutenant receives the order "Retreat", resulting in a tie. In this case is not possible to reach an agreement between the nodes. Then, we will analyze a case with four nodes.

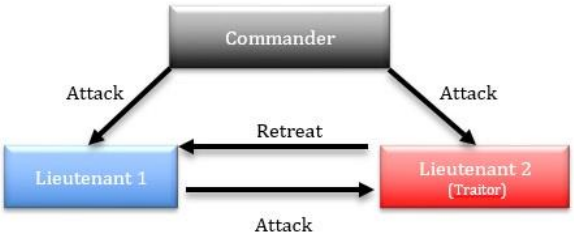


Fig. 2. Case three generals with a traitor (consensus is not possible).

3.2 Case with Four Generals

In this scenario, we have four nodes, the hierarchy always remains. Now, there are three Lieutenant generals and a Commander. Commander sends the message (attack) correctly to each of the lieutenants, and after this each lieutenant communicates this message to each other. Thus, Lieutenant1 sends to Lieutenant2 and Lieutenant3 the message "attack" received from the commander. Lieutenant2 and Lieutenant3 performs similar operation. In this case, there is no problem in reaching a consensus. Now, we analyze what happen when we have a traitor general in the group. The analysis is carried out in parts in order to understand well if consensus is reached or not. In the first part, the Commander sends the "Attack" message correctly to the three Lieutenants, as we can see in Figure 3. However, Lieutenant3 is traitor. Thus, Lieutenant1 and Lieutenant2 forward the message to the other two lieutenants as it was received from the commander, but Lieutenant3 changes the "Attack" message received form the Commander and sends a "Retreat" message to Lieutenant1 and Lieutenant2. Each Lieutenant must take a decision based on the number of messages received. For example, Lieutenant1 receives three messages, one from the Commander and two from the Lieutenants, but one of them is incorrect. Lieutenant1 has two "attack" messages and a "Retreat" message (from traitor general), so as the majority of messages are "Attack", then the consensus is carried out to attack. A similar way is for Lieutenant2. We can see that Commander, Lieutenant1 and Lientenant2 can reach a consensus.

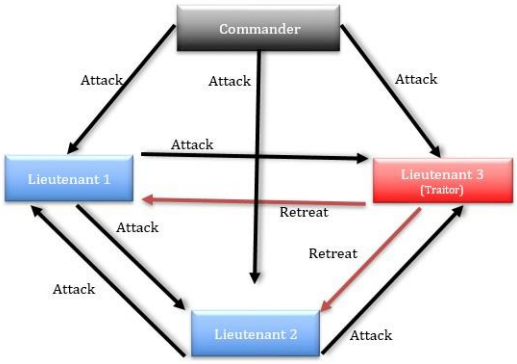


Fig. 3. Communication of the four generals and evaluation (consensus is possible).

4 Design and Implementation

This section gives a briefly description about design and implementation of our web tool. Figure 4 shows a sequence diagram which describes our system for different types of users: a student, a teacher and a general user. Users can see the options shown in the menu and they can choose one of those. In start section a brief description of the origin of this problem is shown. In problem section an explanation about algorithm operation is presented. Finally, in algorithm part we find the Byzantine general model. There each user can perform tests, and analyze the results.

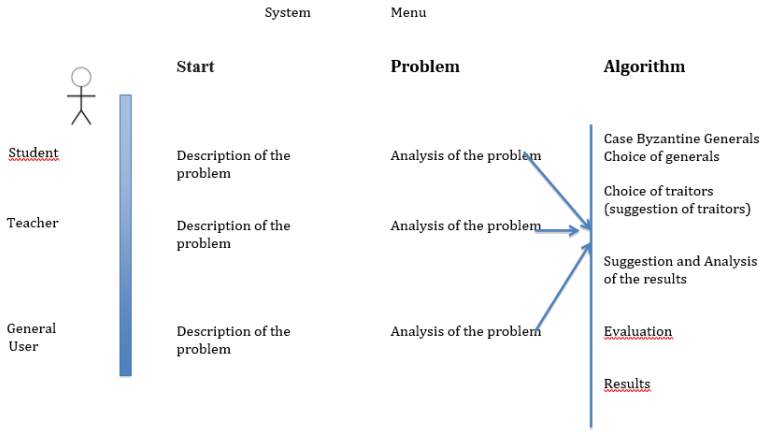


Fig. 4. Sequence diagram.

To implement our design, we are used several programming tools such as JavaScript (Vis.js library), HTML and CSS. Our implementation considers the following tasks

– **Draw function.** This function creates the array that will contain the nodes, and these are assigned values to each of the variables, which will be as these are identified in the program.

- **Commander and Lieutenant Node**. After having created the node to which we are going to connect our other generals, we create our Lieutenants. We use a "FOR" cycle to perform the evaluation according to the number of Lieutenants requested by the user. Once nodes have been created, the connection is made with each one of them. We need to avoid overlapping or some inconvenience in the program.
- **Restrictions.** In order to program works according with the base formula, it was necessary to make the restriction indicated from design, which is the maximum number of traitors that can be in each of the scenarios.
- **Check function** helps to the user to know the maximum of traitor generals that can be introduced to reach an agreement to attack. This message informs to a user the maximum number of traitor generals to have a consensus of "Attack", but if the user decides to enter another amount the program does not restrict doing this.

5 Operation and Evaluation

To evaluate our web tool, we organized its operation in three parts:

1. **Start.** In this section a detailed description of the algorithm is made as well as the analysis of the equation. The base cases that were analyzed in the project are proposed too. Figure 5 shows this scenario.

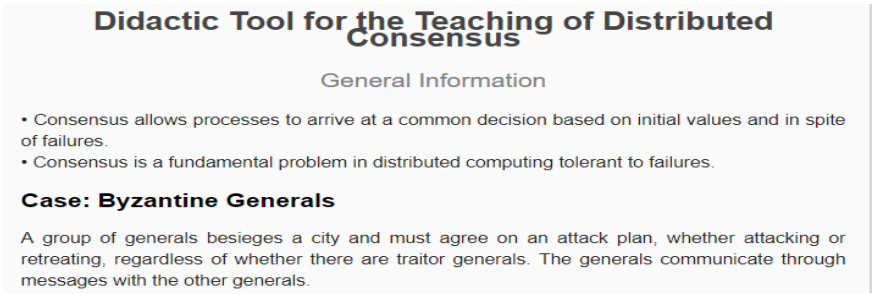


Fig. 5. A part of home section of our web page.

2. **Issue.** Figure 6 shows the part where the user analyzes the algorithm of the problem, similar to the analysis made in Section 3.
3. **Algorithm** is the graphic and interactive part that the user can see. Figure 7 gives an example about the program that indicates the maximum amount of traitor generals that the user can enter. It should be noted that this does not limit the user to enter more, if not, which is a suggestion for the user to obtain a consensus to attack.

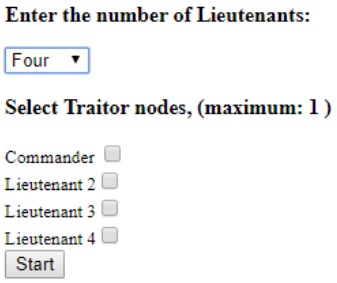


Fig. 6. Description of the problem to be retreated.

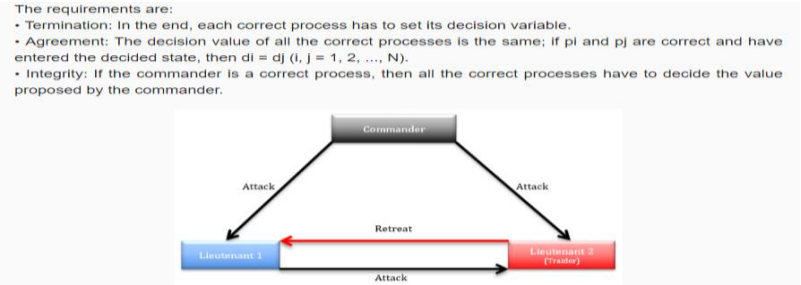


Fig. 7. Selection of traitorous nodes and suggestion.

The user has the option to choose the number of generals to be shown on the platform. A user can also choose how many traitor nodes must appear and at the end of the evaluation a table with the final result is shown. We can see in Figure 8 that result is "CONSENSUS ATTACK". When many nodes are evaluated a graphical representation becomes complex, and our tool uses a table with status of all participant nodes to deal with this problem.

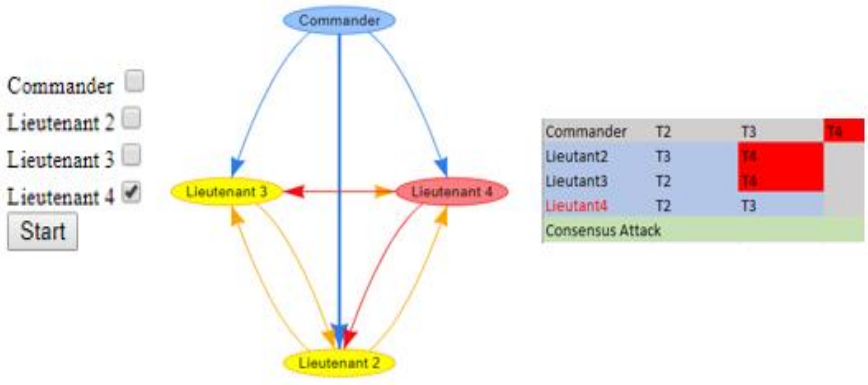


Fig. 8. An example for four generals using our didactic tool.

6 Conclusions

This paper presents a didactic tool to help understand the distributed consensus, in particular, the Byzantine Generals problem. In addition, we tested the use ICT as facilitator in the teaching process for the new generations of students who are now more familiar with new devices to access information in Internet. We hope that our didactic tool also helps to disperse the knowledge of a simpler and cheaper way to those places where is difficult to acquire a large number of books. Our didactic tool has had a good acceptance in our academic community because it solves the Byzantine Generals problem of a graphical an interactive way. We think that with the growth of the Internet this type of didactic tools will be very important in the teaching/learning process. As future work, we plan to add a database to our didactic tool in order to storage information about user experience (UX). In such a way that using artificial intelligence techniques we could increase emphasis on knowledge representation as an activity within a perceptual space and organized by social interactions.

References

1. Kindberg, T., Dollimore, J., Coulouris, G., Blair, G.: Distributed Systems. Pearson Education Limited, pp. 5–25 (2013)
2. Lamport, Shostak, R., Pease, M.: The Byzantine Generals Problem. ACM Transaction on Programming Languajes and Systems 4(3), pp. 382-401 (1982)
3. Uwase, M.P., Tiberghien, J., Steenhaut, K.: Interactive Tools for Learning Sensor Network Basics. In: Advances in Engineering and technology. MacMillan UGANDA, pp. 148–154 (2011)
4. Wein, J., Kourtchikov, K., Cheng, Y., Gutierrez, R., Khmelichek, R., Topol, M.: Virtualized Games for Teaching about Distributed Systems. In: Proceeding of the 40th ACM Technical Symposium on Computer Science Education, Chattanooga, TN, USA, pp. 246–250 (2009)
5. Yeh, Y.C.: Design Considerations in Boeing 777 Fly by Wire computers. Boeing Commercial Airplane. In: Proceedings Third IEEE International High-Assurance Systems Engineering Symposium, Washington, DC, USA, pp. 64–72 (1998)
6. Arévalo-Viñuales, S.: Tolerancia a fallas en sistemas distribuidos mediante replicación de procesos. Tesis Doctoral, Universidad Politécnica de Madrid (1988)
7. Pérez-Porto, J., Merino, M.: Definición de consenso. http://definicion.de/consenso, Last accessed 2018/08/20
8. Jiménez-Peris, R., Patiño, M., Jiménez-Merino, E.: Consenso. Obtained from: http://lsd.ls.fi.upm.es/lsd.html. Last accessed 2018/08/20.

ISSN 1870-4069

Analysis of Speech Acts for the Design of a Corpus of Phrases used in an Intelligent Learning Environment

Xochitl Samantha Delgado-Hernández, María Lucila Morales-Rodríguez, Nelson Rangel-Valdez*, Laura Cruz-Reyes, Claudia Gómez-Santillán, Juan Javier González-Barbosa

TecNM/Instituto Tecnológico de Ciudad Madero, Tamaulipas, Mexico
* CONACyT Research Fellow
xsam.delgado@gmail.com, {lucila.morales, nelson.rangel, claudia.gomez, lauracruzreyes, jjgonzalezbarbosa}@itcm.edu.mx

Abstract. Artificial intelligence is a fundamental part of our life today. An example is the simulators integrated by intelligent virtual characters that serve as instructors in interactive environments where no matter the place and time, people can take a class and even train in a video game. These virtual agents require a corpus of phrases to have a dialogue similar to that of a human being. This article describes an investigation of the types of speech acts immersed in an interactive dialogue. They describe the most relevant characteristics for an educational environment, such as the type of phrase, the function it performs, its meaning, etc., thus discarding those that do not belong to this environment. The main objective is to develop a specific corpus of phrases for an intelligent virtual agent belonging to a serious game for learning purposes used as an interactive learning environment.

Keywords: speech acts, intelligent learning environments, intelligent virtual agent, corpus, BDI architecture.

1 Introduction

Currently we can find very advanced technology based on artificial intelligence, which tries to emulate human behavior through machines, achieving satisfactory objectives in various fields, such as industrial, health, military, education, among others. In these fields, the so-called intelligent learning environments are developed, which are simulated environments with a problem to be solved, which are able to interact or communicate with various systems through the Internet. Technologies through the web give way to interact with other users and in the same way with instructors or agents, who assist and monitor learning.

Smart agents are a fundamental part of this technology, often serve to guide users within the system. Its application is increasing, since as time goes by, new improvements are made in its development and operation, getting closer and closer to its purpose, which is to achieve an entity as intelligent as possible. For this, it is necessary to improve the answers that the agent issues in his dialogue with the user, which leads us to analyze each phrase belonging to a corpus of available phrases to establish an interactive communication. This is where the theory of speech acts

intervenes, since they can be used to classify sentences and select the most appropriate to generate a coherent and natural response.

2 State of the Art

Below are some works related to the development of a corpus, the main theme of this article, which helped to prepare it, contributing ideas or comparing its development.

Úbeda [1] presents a collection of recordings to analyze and generate a corpus suitable to the objective that is, from an oral corpus, to describe the manifestation of the use of language in a situation, through its constant and frequent cognitive premises. Although the objective is different from that of this work, the idea that record a real dialogue in the study environment is retaken to analyze each phrase used.

In Casacuberta et al. [2] three different corpus are designed and constructed, the first with phonetically balanced sentences, the second with sentences corresponding to a task of consulting a geographic database and the third of speech produced in adverse conditions. For the research carried out in this work, the corresponding corpus classification would be the second, since it is intended to generate a specific corpus for an educational environment, which will help the selection of phrases.

The proposed methodology is based in the analysis of speech acts, similarly as it is developed in Torralvo [3] which focuses on finding the conversational, syntactic and pragmatic characteristics that are repeated in several dialogue recordings. With this, it is established a set of models that can serve as a guide in the construction of an automatic system capable of interacting with a human speaker.

3 Background

This section defines the following concepts: 1) Intelligent Learning Environment; 2) Virtual Agents; and 3) Theory of Speech Acts. Knowing these concepts will guide the reader through the understanding of the fundamental parts that indicates what type of agent is used, and why, in a dialogue inside a learning environment that be similar to that of a human being.

3.1 Virtual Agents in Intelligent Learning Environments

According to Hwang [4], Intelligent Learning Environments (ILE) are learning environments supported by technology. They make adaptations and provide adequate support with intelligent detection and recommendation technology in the right places at the right time. According to the individual needs of each student, the environments are determined through the analysis of their learning behaviors, their trajectory of performance and the online and real-world contexts, and composed by the resources that are provided or accessed, downloaded, read or elaborated.

According to Ros [5], ILE systems aim to help students learn even when doing other activities. They are the teacher friends, which look for opportunities to advise the student in their daily life taking into account their needs and preferences.

Within the ILE you can find instructors that serve as a guide to the student, which can be represented by virtual agents with a specific architecture and function (e.g. [6,7]).

An intelligent agent can be physical, like a robot, or virtual, like a computer program with which you can interact. The study in this article is related exclusively with intelligent virtual agents (IVA). This research specifically pursues the provision of an infrastructure to generate a coherent and natural interactive dialogue for intelligent virtual agents.

There are several types of intelligent virtual agents, classified mainly depending on their architecture, which is the internal structure that determines the mechanisms that an agent uses to react to stimuli.

According to Iglesias [8], the architectures of the agents can be classified into deliberative architectures and reactive architectures. The deliberative architectures follow the current of the symbolic artificial intelligent, which is based on the hypothesis of the physical-symbol systems enunciated by Newell y Simons [9], according to which a system of physical symbols capable of manipulating symbolic structures can exhibit intelligent behavior. In order to work at Newell's level of knowledge, our problem will be how to describe the objectives and means of satisfying them, and how to translate the level of knowledge to the symbolic level.

In this case we have selected an agent with BDI architecture (with Beliefs, Desires and Intentions) aka deliberative agents, that besides being very intelligent agents (mainly in their interaction with the user), they also contain variables for the selection of the dialogue that will help with the characterization of the sentences.

A BDI agent architecture must offer responses according to the program in function. To achieve this goal, the agent must be able to express itself through specific speech acts.

3.2 Theory of Speech Acts for the Generation of Interactive Dialogue

According to Lozano [10], the first theory of speech acts was proposed by J. L. Austin in 1962. This theory basically states that, in saying one thing, another was also done besides the simple act of saying it, such as asking or promising, among others. Later, in 1990, Searle deepened the subject by perfecting Austin's theory of speech acts, and made an extension of the analysis of it.

Austin [11] identified three different acts when making a sentence. Such acts are discussed as follows:

a. Locutionary act: the act of utterance a sentence with a certain meaning or reference.
b. Illocutionary act: the communicative force that accompanies a sentence, such as asking, questioning and promising, among others.
c. Perlocutionary act: the effect on the receiver, either on their feelings, thoughts or actions.

Also, speech acts can be divided into two types according to Austin [11]: direct acts, and indirect acts. The direct acts are those statements in which the sequence and illocutive aspect coincide, that is, the intention is expressed directly. The indirect acts

are those phrases in which the sequence and illocutive aspect do not coincide, therefore the purpose of the sentence is different from what is expressed directly.

Searle [12] proposes as classification of the illocutionary acts the assertions, directions, commissions, expressions, and declarations.

In the acts of assertion, the intention of the speaker is to express to the receiver how things are, the correspondence address is of words to the world; the condition of sincerity, that is, the one that expresses the psychological state of the speaker in carrying out the speech act, is to believe that the proposition is true.

In the acts of direction, the intention of the speaker is to make the receiver do something. The correspondence address is from the world to the words; the condition of sincerity is desire.

In the commission acts, the speaker's intention is to commit to a future act. The correspondence address is from the world to the words and the condition of sincerity is the intention.

In the acts of expression, the intention of the speaker is to express their feelings and attitudes. There is no correspondence address, since, when performing an act of this type, the speaker will try to make the correspondence address from the world to the words or vice versa.

In the declaration acts, the speaker's intention is to cause a change in the world through the statements that she/he utters. The correspondence address can be from the world to the words or the other way around, there is no condition of sincerity. For a better understanding of the classification provided by Searle, Table 1 presents examples of each type of illocutionary act that has been taken from Alanis [13].

Table 1. Types of illocutionary acts.

Type	Example
Assertive	statements, claims, hypothesis, descriptions, explanations, and reports
Commissive	promises, vows, oaths, offers, or threats
Directive	suggestions, challenges, requests, questions, orders, and mandates
Declarative	baptize, bids, blessings, name, inaugurate, and dismiss
Expressive	apologies, complaints, congratulations, thanks, deplores, and condolences

Based on the analysis of definitions of types of speech acts, it was identified that it is necessary to have a database or corpus of phrases, where the possible answers are stored, which is a fundamental part of this project for the good functioning of the system.

In the field of linguistics, according to Torruella [14], the word corpus is a somewhat ambiguous word, and it is currently used in a general sense to refer to any type of compilation of texts. Then the appropriate corpus will be made to classify according to the speech acts previously seen. The following section presents the proposed methodology that should be followed to accomplish such task.

4 Methodology

The methodology proposed for the creation of a corpus consists of four stages: 1) Obtaining the dialogue; 2) Obtaining the sentences; 3) Analysis and classification of sentences; and, 4) Generation of the specific corpus.

In stage 1, a dialogue was collected as an example of a real conversation in an educational environment, for its use in the analysis. At this point, Google's ListNote application [15] was used to record the phrases of the conversation, from a questionnaire given by a teacher to his student.

In stage 2, with the written dialogue, we proceeded to pass it to a file where each of the sentences was broken down, eliminating errors when recording and completing some words, which many times, due to the low tone of voice, the application does not reach to distinguish.

In stage 3, once the necessary sentences were collected, each one was analyzed and classified through speech acts previously investigated, and with this, selecting the most appropriate for an intelligent virtual agent to serve as an instructor in the educational environment.

In stage 4, the selected phrases are finally collected for the creation of the corpus, which will serve as the speech acts of a conversational virtual agent in an ILE.

5 Results

In this section,we present results obtained for the application of a methodology for the creation of a specific corpus, proposed in this work.

The first stage was the obtaining of dialogue, in this case, so that the dialogue was real, some mentioned phrases were recorded in an educational environment, where the instructor is a teacher and the user is a student of the basic math material. The dialogue consists of 29 sentences with 435 words. Below is a part of the dialogue obtained:

> *"Hello, good afternoon. How can I serve you seriously extra? Good afternoon? I have the doubt in this problem, you are in love, and you do not pay attention to the class. Hehe, not the teacher as I believe I want to know what it is What operation is the one that I will solve first to be able to continue passions is a equation ma of fractions with multiplication of fractions with square roots okay To begin what is your first doubt for the next thing you need law of order you know what is the law of order roots multiplication with exponents and root then here I am going to draw what is the root of 9 and under the same the whole problem and now where do I have the root I put in what is the 3 What is the square root 9 "*

The second stage proposed in the methodology was obtaining the sentences, where the sentences of the dialogue are separated to give them order and clarity, so that we can work with each one of them. The dialogue is not very clear in some sections, or the words appear incomplete, since the application cannot clearly understand them when the dialogue is performed with low voice.

Once the sentences were separated, it proceeds with the identification of the actor involved with each of them; the results from such task are shown in Table 2. There, the actors identified were the student and the Professor involve in the dialogue derived from the case of study.

The main idea under the case of study is to represent through an IVA, the speech act of instructor in an educational environment. The IVA must act as a teacher in such a way that it incorporates a direct, serious and firm character in its speech such that it maintains the order within a class room. To achieve this purpose, the third stage of the methodology analyze each sentence shown in Table 2 in order to describe and classify it according to the revised speech acts theory.

Table 2. Dialog phrases.

Actor	Phrases
Professor	Hello good afternoon, How can I serve you?
Student	Good afternoon, I have a doubt in a problem
Professor	Surely you are in love, and do not pay attention in class.
Student	Hehe, not master as he believes
Professor	What kind of operation do you want to solve?
Student	It is an equation of fractions with multiplication of fractions and square roots
Professor	Okay To begin what is your first doubt
Student	Where to start if the problem is 5 (1/4) + (2/3) -√9
Professor	For the following you need law of order, do you know what the law of order is?
Student	Yes, first the roots and exponent are solved, then multiplication and divisions and at the end, addition and subtraction. So here I'm going to take root first of 9?

Table 3 summarizes the results obtained from this step. In column 2 are shown the sentences from the dialog that corresponds to its locutionary act. Column 3 shows the description obtained from analyzing the intentions in the locutionary act, which were obtained by interpreting the real intentions of the subject of study (i.e. the teacher). Column 4 shows the description obtained from analyzing the real effect of the locutionary act, which was obtained through the examination of the sequences of statements and the impact among the interlocutors. Columns 5 and 6 presents the classification according to the type of speech acts (TSA) and the type of illocutionary act (TIA); this classification is derived from the comparison among the concordance among the illocutionary, and perlocutionary acts.

Let us point out that those that do not comply with the overall purpose of the dialogue and the context are highlighted in bold.

Finally, in the stage four it is generated the specific corpus. After having classified the sentences, it was determined that the speech acts need the direct type, since the direct phrases are very clear and do not hide another meaning in the sentence, so the agent would be understood easily which is one of the purposes. Also, an agent must use the directive illocutionary act type, since, like any instructor, the user must be questioned, guided and given instructions.

In the case of the study the other interlocutor besides the instructor is a student; this situation obligates an instructor to be firm and consistent in his/her dialogue in order that the student achieve his/her goals.

The phrases that were eliminated are of the indirect type. They do not comply with the formal behavior of an instructor. In order that an IVA succeed it cannot be assumed on its design that there is trust from the students at the beginning, or that the agent has prior knowledge of their life, without that information an indirect speech act is impossible.

Once the sentences that do not correspond have been eliminated, the final corpus will have the selected phrases shown in Table 4. It should be mentioned, as in the previous examples, that only a small part of the entire dialogue is being integrated.

Let us observe it is desirable to organize the phrases on the corpus according to an expected characterization that are suitable for the specific application of an IVA in a ILE. This way it is possible to progress in the automation of the dialogue generation used in interactions with human beings.

Table 3. Analysis of sentences.

#	Locutionary	Illocutionary	Perlocutionary	TSA	TIA
1	Hello good afternoon, How can I serve you?	Say hello and ask what you can do	He is greeted and receives service	Direct	Directive
2	Good afternoon, I have a doubt in a problem	Say hello and declare a doubt a problem	Corresponding greeting and declares the doubt	Direct	Directive
3	**Surely you are in love, and do not pay attention in class.**	**It means that he is distracted and does not pay attention**	**He's joking to break the ice**	**Indirect**	**Declarative**
4	**Hehe, not master as he believes**	**Laughs to be nice and denies being distracted**	**The joke corresponds**	**Indirect**	**Expressive**
5	What kind of operation do you want to solve?	Ask about the type of operation	Ask about the type of operation	Direct	Directive
6	It is an equation of fractions with multiplication of fractions and square roots	Describe the type of equation	Describe the type of equation	Direct	Directive
7	Okay To begin what is your first doubt	Ask for the first doubt	Ask for the first doubt	Direct	Directive
8	Where to start if the problem is 5 (1/4) + (2/3) -√9	Do not know where to start and declare the problem	Do not know where to start and declare the problem	Direct	Directive
9	For the following you need law of order, do you know what the law of order is?	Explain what you need and ask if you know	Explain what you need and want to know if you know	Direct	Directive
10	Yes, first the roots and exponent are solved, then multiplication and divisions and at the end, addition and subtraction. So here I'm going to take root first of 9?	Explain the law of order and ask if it's okay as it starts	Explain to check what you know and ask if it's okay	Direct	Directive

Table 4. Corpus of selected phrases used by professors or students in ILE.

Actor	Corpus of phrases
Professor	Hello good afternoon, How can I serve you?
Professor	What kind of operation do you want to solve?
Professor	Okay To begin what is your first doubt
Professor	For the following you need law of order, do you know what the law of order is?
Student	Good afternoon, I have a doubt in a problem
Student	It is an equation of fractions with multiplication of fractions and square roots
Student	Where to start if the problem is 5 (1/4) + (2/3) -√9
Student	Yes, first the roots and exponent are solved, then multiplication and divisions and at the end, addition and subtraction. So here I'm going to take root first of 9?

6 Conclusion and Future Works

This research work proposes a speech acts approach to analyze the dialogue from a specific context to generate corpus of phrases. For the elaboration of the specific

corpus, a methodological application was made to analyze sentences of a real dialogue, through speech acts, which led to classifying the sentences and selecting only those that fulfilled the function of a teacher and a student. The case of study to test the proposed methodology was a learning environment.

Specific phrases were selected for an intelligent learning environment, focused on education, through the analysis of speech acts, creating a guide for the conformation of a specific corpus. The results shown that it is possible to eliminate redundant or unnecessary phrases from a dialogue in a specific context. Also, it was possible to characterize the sentences involved in the dialogue according to speech act theory.

From here a broader study can be made with more dialogues to analyze and shape the corpus suitable for the instructor agent. It also leaves a previous work for the characterization of the phrases using the BDI architecture, with this the agent will be able to select the best answer for the user.

Acknowledgments. This work has been partially supported by the following projects: (a) CONACYT/Fronteras de la Ciencia Project 1340; (b) Projects Tecnológico Nacional de México – Instituto Tecnológico de Ciudad Madero: "Análisis de Modelos para incorporar Preferencias en Problemas de Optimización Multicriterio con Dependencias Temporales"; (c) Project 280081 Red Temática para el Apoyo a la Decisión y Optimización Inteligente de Sistemas Complejos y de Gran Escala (OPTISAD), Universidad Autónoma de Nuevo León; and (d) Project 3058 from the program Catedras CONACyT Project 3058 from the program Cátedras CONACyT. The authors also acknowledge the support from the "Red Temática en Inteligencia Computacional Aplicada del CONACyT" (Red ICA).

References

1. Úbeda-Mansilla, P.: Estudio de un corpus de textos conversacionales en inglés realizado en estudios de arquitectura: Su aplicación al diseño de un programa de inglés para arquitectos (2001)
2. Casacuberta, F., García, R., Llisterri, J., Nadeu, C., Pardo, J.M., Rubio, A.: Desarrollo de corpus para investigación en tecnologías del habla (1992)
3. Torralvo, I.: Análisis de modelos de diálogos en contextos específicos para su tratamiento automático (2007).
4. Hwang, G.J., Chin-Chung, T., Yang, S.J.: Criteria, strategies and research issues of context-aware ubiquitous learning. Journal of Educational Technology & Society 11(2), pp. 81–91 (2008)
5. Zapata-Ros, M.: Aprendizaje inteligente. In: Redes Abiertas (blog), 16 de abril (2018)
6. Morales-Rodríguez, M.L., Ramírez-Saldivar, J.A., Hernández-Ramírez, A., Sánchez-Solís, J.P., Martínez-Flores, J.A.: Architecture for an Intelligent Tutoring System that Considers Learning Styles (2012)
7. Morales-Rodríguez, M.L., González, J.J., Florencia-Juárez, R., Sánchez-Solís, J.P.: Emotional Dialogue in a Virtual Tutor for Educational Software (2012)
8. Iglesias-Fernández, C.Á.: Fundamentos de los Agentes Inteligentes. Departamento de Ingeniería de Sistemas Telemáticos (1997)

9. Newell, A., Simon, H.A.: Computer science as empirical inquiry: Symbols and seach. Commun. ACM. 19, pp. 113–126 (1976)
10. Lozano-Bachioqui, E.: La interpretación y los actos de habla (2010)
11. Austin, J.L.: How to Do Things with Words (1962)
12. Searle, J.: Speech Acts: An Essay in the Philosophy of language (1969)
13. Alanis, P.: Caracterización de Expresiones Gestuales Semióticas Influenciadas por la Personalidad y el Estado Emocional para Personajes Virtuales (2011)
14. Torruella, J.: Diseño de corpus textuales y orales (1999)
15. Google: ListNote 8.8 para Android. https://listnote.uptodown.com (2016)

ISSN 1870-4069

Process of Building an Educational and a Military Ontology for the Mexican Context

Cristal Karina Galindo Durán[1], R. Carolina Medina-Ramírez[1], María Auxilio Medina Nieto[2], José Luis García-Cué[3]

[1] Universidad Autónoma Metropolitana – Iztapalapa, Departmento de Ingeniería Eléctrica, Mexico City, Mexico
[2] Universidad Politécnica de Puebla, Departamento de Posgrado, Puebla, Mexico
[3] Colegio de Postgraduados, Campus Montecillo, Texcoco, State of Mexico, Mexico
cdgalindod@gmail.com, cmed@xanum.uam.mx
maria.medina@uppuebla.edu.mx, jlgcue@colpos.mx

Abstract. This article presents a process to build ontologies for the educational and military domains focused on the representation of knowledge in Mexican institutions. The Mexican context is represented by the structure of these institutions, their teaching modalities and considering the educational strategies to enable students to achieve good cognitive and formative levels. The ontologies are used by individuals of the Military Education System as training tools. The article includes the description of some tests for ontology evaluation. The preliminary results show that these tools are acceptable for potential users.

Keywords: knowledge model, ontology, education, military.

1 Introduction

The Semantic Web is an extension of the current web, which has a structure that enables to express the content of pages or documents so that computers can understand and process them, this fact facilitates interaction between computers and users [1]. To make this possible, we need knowledge models (ontologies), descriptions of resources; as well as management languages and knowledge representation techniques.

In particular, ontologies can have different purposes, among which stand out: the publication of information according to a knowledge model, the exchange of information between applications, the disambiguation of concepts, the inference of knowledge and the description of vocabularies [2,3].

In recent years, multiple ontologies have emerged in different fields of knowledge ranging from Physics to Social Sciences [4]; however, the adaptation of ontologies in particular domains is a common task due to the need of covering specific domains, that means that although reutilization is recommendable, often only general concepts are maintained.

This article presents the construction of two ontologies: an educational and a military ontology focused on the representation of knowledge in Mexican institutions. The Mexican context is represented by the structure of these institutions, their teaching

modalities and considering the educational strategies to enable students to achieve good cognitive and formative levels. These ontologies are used by individuals of the Military Education System as training tools, they are also part of the MIIDAS prototype, which is a proposal for the integration of educational resources managed through semantic technologies [5].

The rest of the article is structured as follows: Section II presents some concepts about ontologies. Section III presents a review of the ontologies developed in the educational and military domain. Section IV explains the process of building both ontologies. Section V shows preliminary results of ontologies' evaluation. Finally, conclusions and future work are presented in Section VI.

2 Theoretical and Explanatory Framework

The theoretical perspective of this work involves the topic of ontologies, which we describe below.

Ontologies can be conceived as an explicit and formal specification of a shared conceptualization [14]. From this definition, it can be said that the specification refers to the selection of a particular domain, the explicit refers to the concepts used in the ontology and the restrictions for its use that must be clearly defined, the formal describes the use of comprehensible syntax for computers, the conceptualization points to the representation of knowledge and shared refers to the consensus by domain experts [15].

According to [16], the main elements of ontology are: *class* (concepts), *properties* (relationships), and *individuals* (instances). *Classes* are real-world objects, which can be grouped with elements that have similar characteristics; these classes are the base element of ontology and describe the concepts of a specific domain. The *properties* are relationships and serve to describe relevant features of the entities; these can be of three types: object, data or annotation. *Individuals* are elements which belong to a specific class.

OWL (Ontology Web Language) is the standard language of the Semantic Web to express and codify ontologies [17], this language is based on descriptive logics. OWL language is composed of three sublanguages with different levels of expressivity, these are: 1) OWL-Lite for those cases that need a hierarchical classification and simple restrictions, 2) OWL-DL for those cases that require great expressiveness and a computability guaranteed and 3) OWL-Full for those cases that require maximum expressiveness and complete syntactic freedom; but without guarantee of complete computability. OWL also uses special software modules, called *reasoners*, that can make inferences and to check logic consistency into the knowledge base of ontologies.

3 Related Work

In this section, we present some ontologies that have been developed for the educational and military domains, the purpose is to review the concepts that can be adopted by our own ontologies for knowledge reuse.

An ontology that represents a knowledge model which establishes the relationships between the requirements of the labor markets and the content of the work plans is presented in [6], the ontology is studied during a given training program. In the article [7] describes domain and pedagogical ontologies which together help to enable the search, visualization and navigation of learning objects in Science of the Earth and Geography areas. Curriculum Ontology [8] is an ontology proposed by the UK public radio and television service, which aims to provide data models and vocabularies to describe national curricula in the same country.

In the HERO ontology (Higher Education Reference Ontology) [9] several general aspects of the domain of a university are described, such as the organizational structure, personal (academic and administrative), roles (teaching and research) and even incomings.

Regarding ontologies in the military domain, the following stand out: the C2 ontology [10], this exposes a knowledge model that specifies military command and control concepts of the Department of Defense of the United States.

The military ontology Muninn [11] defines classes and properties of military history. The article [12] describes the implementation of an ontology as a basis for the intelligent information system of the tactical command of the army of Korea. The THOR ontology [13] provides a vocabulary to describe and request the content generated by the combatant.

As summary, the related works shows diverse ontologies from different perspectives, however, they do not consider the features of the Mexican context. That represents the organization of educational institutions in Mexico, their educational approaches, the teaching modalities and strategies used to achieve cognitive and formative levels of individuals of the Mexican Military Education System.

4 Building Ontologies

For the building of the ontologies Methodology 101 [18] was used, proposed by Stanford University, which consists of 7 phases, these phases are: 1) Determine the domain and scope of the ontology, 2) Reuse existing ontologies, 3) List important terms for the ontology, 4) Define classes and their hierarchy, 5) Define the properties of the classes: slots, 6) Define the facets of the slots and 7) Create instances or individuals.

4.1 Building Phases for the Educational and Military Ontology

Before building the ontologies it is necessary to define the domain and the scope of these, in order to determine it, it is essential to ask the competency questions. Table 1 shows the different competency questions for each one of the ontologies.

Once the competency questions have been formulated, the scope of the ontologies can be established. Table 2 shows its respective scope.

After analyzing the scope of the ontologies it is necessary to list the important terms; as well as defining its hierarchy. For the educational ontology, we worked with experts in education and pedagogy that helped to obtain the significant terms.

Table 1. Competency questions for educational and military ontologies. (Source: own research).

Ontology	Competency questions
Educational	-What kind of educational approach do Universidad Autónoma Metropolitana, Unidad Iztapalapa (UAMI) and Escuela Militar de Ingenieros (EMI) use? - What are the didactic strategies established at the UAMI and the EMI? -What are teaching strategies are used to achieve cognitive and formative levels? -What are the test instruments used by the EMI?
Military	-What is the Plan of Assistance to the Civil Population in Cases of Disaster (DN-III-E)? -What are the values promoted by the Military Education System? -What is the responsibility of a captain? -What is a duty in the military field?

Table 2. The scope of the educational and military ontologies. (Source: own research).

Ontology	Scope
Educational	The ontology should describe several aspects of Mexican educational institutions, such as their structure, their teaching modalities, the educational strategies to achieve a cognitive and formative level in the students; as well as the evaluation instruments used.
Military	The ontology should describe the terminology used in the different subjects for the training of military personnel in the various curricula of the Military Education System.

Regarding the Military domain ontology, the taxonomy of concepts and their definitions were obtained by reviewing documents that belong to the Military Educational System. Figure 1 shows the hierarchy of classes of both ontologies.

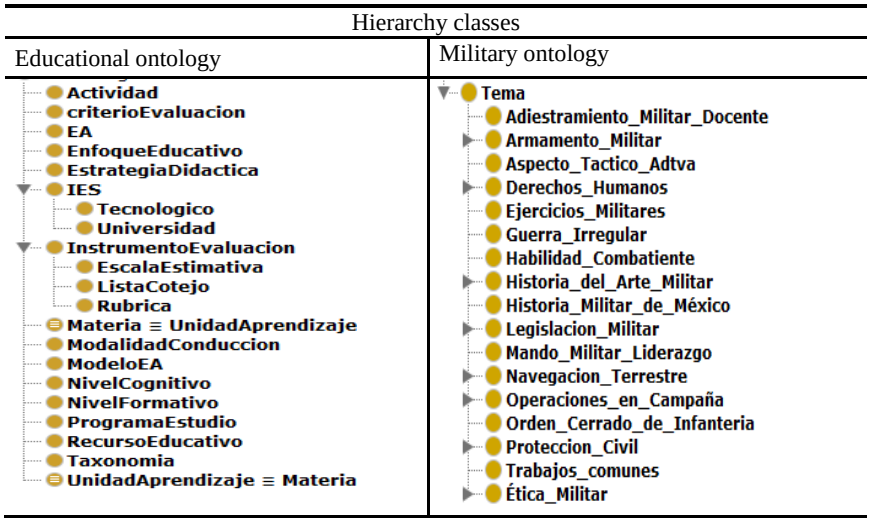


Fig. 1. Hierarchy classes of the educational and military ontology. (Source: own research).

Table 3. Object properties of the educational ontology. (Source: own research).

Property	I	F	T	S	R	Domain	Range
tieneProgEstudio	Y	N	N	N	N	Educational institution (EI)	Degree program
perteneceA	Y	N	N	N	N	Degree program	EI
tieneMateria	N	N	N	N	N	Degree program	Subject
tieneEstiloAprendizaje	Y	N	N	N	N	Learning Style Model	Learning style
perteneceAModelo	Y	N	N	N	N	Learning style	Learning Style Model
estableceEstrategiaApre	N	N	N	N	N	EI	Learning strategy

Symbology
Y: Yes N: Not
I: Inverse F: Functional T: Transitive S: Symmetric R: Reflexive

In summary, the educational ontology has 22 classes, 15 object properties, 18 data properties and 22 annotations. The military ontology has 284 classes and their corresponding annotations. Protégé editor was used to build the ontologies; they are implemented in OWL version 2 language.

5 Ontologies Evaluation

The quality of the built ontologies was estimated by considering three aspects: structural evaluation, functional evaluation and usability. These aspects were tested as follows.

5.1 Structural Evaluation

The structural evaluation consists of analyzing the logical structure of the ontology. Three different reasoners were used to verify the logical consistency as well as the redundancy of information, they were FaCT ++ [19], Pellet [20] and HermiT [21], all of them reported no inconsistencies or redundancies in the educational and military ontologies.

5.2 Functional Evaluation

Functional evaluation refers to the use of ontology; as well as the conceptualization of some domain. This evaluation includes aspects such as the agreement of domain experts, evaluation through competency questions and an estimation of user satisfaction.

Table 4. Competency questions for the educational and military ontology with its corresponding query in SPARQL and its result. (Source: own research).

Ontology	Natural language query	Query in SPARQL	Result
Educational	Which educational institution does the Bachelor of Computer Science belong to?	PREFIX oed: <http://pcyti.izt.uam.mx/pmiidas/ontoeduca#> SELECT ?ies WHERE { ?PEstudio oed:perteneceA ?ies FILTER regex (?PEstudio, "^Licenciatura en Computación")}	UAMI
	How many teaching strategies does EMI establish for teaching practice?	SELECT (count (?Estrategia) as ?numEstrategia) WHERE { ?uni oed:estableceEstrategia ?Estrategia FILTER regex (?uni, " ^EMI ") }	5
Military	What is the duty of a captain?	PREFIX po: <http://pcyti.izt.uam.mx/pmiidas/ontomilitar#> PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#> SELECT ?subject ?object WHERE { ?subject rdfs:comment ?object. FILTER(?subject=< http://pcyti.izt.uam.mx/pmiidas/ ontomilitar.owl#Deber_Capitán>)}	He is responsible for the instruction, good internal management and military spirit of the force under his command.
	What are the institutional values?	SELECT ?subject ?object WHERE { ?subject rdfs:SubClassOf ?object. FILTER(FILTER(?subject=<http://pcyti.izt.uam.mx/pmiidas/ ontomilitar.owl#Valor_Institucional>))}	Spirit of body, loyalty, honor, discipline, patriotism, courage, self-denial and honesty.

Domain Expert Evaluation. This evaluation is carried out by experts in the domain, which assess the compliance of the ontology in terms of concepts, hierarchy, standards and requirements [22]. The educational ontology was evaluated by experts in education, specifically in didactics and pedagogy, while the military ontology was evaluated by expert personnel in the military environment (captains and majors). In general, the suggestions of the experts were based on the classification and equivalence of concepts, inclusion of object and data properties; as well as reviewing annotations.

Evaluation through Competency Questions. This test consists of translating the competency questions posed at the beginning of the construction of the ontologies to the SPARQL query language. Table 4 shows the queries made to the educational and military ontology in natural language, the respective queries in SPARQL as well as the result set.

5.3 Usability Evaluation

To evaluate the usability of the ontologies, a prototype called MIIDAS was made, which is a web application that makes use of these ontologies and allows users to interact with them by means of a graphical user interface. A usability rubric was used as an evaluation instrument to evaluate ontologies.

This rubric was applied to 25 professors and 39 students of the Military School of Engineers, all of them are part of the Military Educational System. As an illustration, Figure 2 shows the average of qualification assigned for navigation and utility (the qualification is a number between zero to ten).

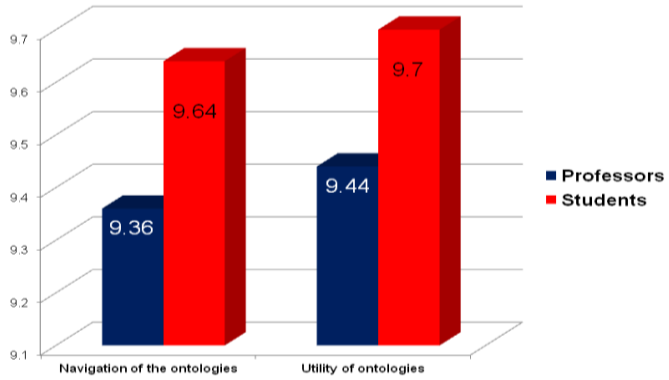


Fig. 2. Average result of navigation and utility for the build ontologies. (Source: own research).

According to the results presented on the evaluation criteria of the ontologies, these were favorable in terms of navigation and usefulness for the student and teaching population, as they reported favorable results for each of these criteria.

6 Conclusions

This article describes the process of building two ontologies, one in the educational and the other in the military domains; both ontologies were built by implementing Methodology 101 and considered the context of the Mexican Military Educational System.

Ontologies presented in this work have been preliminary evaluated in three different aspects: structural, functional and usability, obtaining favorable results in each one. These ontologies can serve as a reference to develop multiple applications and as a point of reference to standardize a vocabulary within Mexican institutions.

As future work, we have to research plans, the first is to extend the educational ontology to support training of teacher's evaluation established by the National Institute for the Evaluation of Education (INEE), and the second is to support the impact of the ontology when this is used to support individual's development.

References

1. Berners-Lee, T., Hendler, H., Lassila, O. The Semantic Web. Scientific American 284(5), pp. 34–43 (2001)
2. Horrocks, I.: Pragmatism and the evolution of semantic systems. In: Pragmatism and the evolution of semantic systems, B. Küppers, U. Hahn and S. Artmann (Ed.), Springer Link, pp. 175–188 (2013)

3. Labra-Gayo, J.: Web Semántica: Comprendiendo el cambio hacia la Web 3.0. 1st ed. Netbiblo, España (2011)
4. Swoogle Semantic Web Search Engine, http://swoogle.umbc.edu/2006/. Date consulted: 2018/08/10
5. Galindo-Durán, C., Medina-Ramírez, R., Medina-Nieto, M., García-Cué, J.: PMIIDAS: Integration of open educational resources based on semantic technologies. Research in Computing Science 144, pp. 151–165 (2017)
6. Vas, R.: Educational Ontology and Knowledge Testing. The Electronic Journal of Knowledge Management 5 (1), pp. 123–130 (2007)
7. Gahegan, M., Agrawal, R., Banchuen, T., DiBiase, D.: Building rich, semantic descriptions of learning activities to facilitate reuse in digital libraries. International Journal on Digital Libraries 7(1), pp. 81–97 (2007)
8. Curriculum Ontology, https://www.bbc.co.uk/ontologies/curriculum. Date consulted: 2018/10/12
9. Zemmouchi-Ghomari, L., Ghomari, A.: Process of Building Reference Ontology for Higher Education. In: World Congress on Engineering, pp. 1595-1600. London, UK (2013)
10. Smith, B., Kristo, M., William, M.: The Ontology of Command and Control. In: Proceedings of the 14th International Command and Control Research and Technology Symposium (ICCRTS) (2009)
11. Military Ontology Specification-1.1: http://rdf.muninnproject.org/ontologies/military.html. Date consulted: 2018/10/13.
12. Yoo, D., No, S., Ra, M.: A Practical Military Ontology Construction for the Intelligent Army Tactical Command Information System. International Journal of Computers Communications & Control 9 (1), pp. 93–100 (2014)
13. LeBlanc, E., Nguyen, D., Balduccini, M., Regli, W.: Military Ontologies for Information Dissemination at the Tactical Edge. In: Joint Ontology Workshops, Buenos Aires, Argentina (2015)
14. Gruber, T.R.: A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition 5(2), pp. 199–220 (1993)
15. Gómez-Pérez, A., Manzano-Macho, D. An Overview of Methods and Tools for Ontology Learning from Texts. The Knowledge Engineering Review 19(3), pp. 187-212 (2005)
16. OWL Web Ontology Language Overview, https://www.w3.org/TR/owl-features/. Date consulted: 2018/10/14
17. OWL Web Ontology Language Reference, https://www.w3.org/TR/owl-ref/. Date consulted: 2018/10/14
18. Noy, N.F., McGuinness, D.L.: Ontology development 101: A guide to create your first ontology. Standford Knowledge Systems Laboratory Technical Report KSL-01-05, https://protege.stanford.edu/publications/ontology_development/ontology101.pdf. Date consulted: 2018/09/28
19. OWL: FaCT++, http://owl.man.ac.uk/factplusplus/. Date consulted: 2018/10/16
20. Pellet, http://pellet.owldl.com. Date consulted: 2018/10/16
21. HermiT Reasoner: Home, http://www.hermit-reasoner.com/. Date consulted: 2018/10/16
22. Lozano-Tello, A., Gomez-Perez, A.: Ontometric: A method to choose the appropriate ontology. Journal of Database Management 15(2), IGI-Global Publisher, pp. 1–18 (2004)

ISSN 1870-4069

Comprehensive Model for Learning

Norma Josefina Ontiveros Hernández[1], Miguel Pérez Ramírez[2], Jesús Ángel Peña Ramirez[1], Sócrates Espinoza Salgado[1], Mario Humberto Tiburcio Zuñiga[1]

[1] Tecnológico Nacional de México/Instituto Tecnológico de Zacatepec, Mexico
[2] Instituto Nacional de Electricidad y Energías Limpias, Mexico
norma.oh@itzacatepec.edu.mx, mperez@ineel.mx, penita14@hotmail.com, socrates.es@itzacatepec.edu.mx, mario.tz@itzacatepec.edu.mx

Abstract. In this paper we discuss about learning contexts and analyze different aspects or factors which intervene in specific learning processes, so that the learning goals can be achieved effectively. The personalized context and flexibility are characteristics own by personalized learning. However, this can be expensive and inaccessible in practice, but it is argued that technology could provide some support to this personalization of learning. Thus, a comprehensive model to learning is delineated here and some examples of dimension are detailed.

Keywords: context of learning, learning process, multidimensional model, learning goals.

1 Introduction

The aim of this article is to analyze some of the characteristics or factors that are part of the learning context and that can determine the success or failure of the teaching-learning process.

The learning process includes three stages [1] where the student: 1) receives the information, 2) processes it, assimilates or understands it and, 3) finally, is able to make use of that information or knowledge through its application within the specific field of the acquired knowledge.

The problem that occurs in many educational institutions is that not all students present the same level of achievement. This could be due to multiple factors, for example, their sociocultural conditions, lack of prior knowledge, deficiency in their capacity for analysis and synthesis, lack of interest, lack of attitude, etc.

Here we offer a brief analysis in which we first propose to analyze the learning context and then integrate the support of technology to try to emulate the flexibility and richness of personalized instruction. Here we are assuming that personalized instruction is efficient in the sense that an instructor, who tries to teach a student, detects the deficiencies and strengths of his student and uses a variety of resources to get the student to reach the predefined learning goals; what is not practical in a group instruction.

The rest of the article is organized as follows: The multidimensional model of learning is shown below. Afterwards, a discussion section includes some reflections on the dimensions described. Finally, some conclusions are included, which is followed by a list of references.

2 Towards a Multidimensional Learning Model

2.1 Learning Context

Learning contexts (LC) are conceived as the sum of factors that intervene in a specific learning process. From this point of view and unlike traditional instruction, based on groups of students and instructors in a classroom and even in distance mode, which is usually considered incomplete and less efficient, here the integral model is followed, where it is assumed that the more dimensions or factors are integrated into a specific learning process, that is the more complete is a specific learning context, the more efficient the instruction is to achieve a specific learning goal. However, the fact that different factors are present in a learning context does not imply that they exist in a disorderly manner. The most obvious factors may be present simultaneously, for example: text, audio, images, etc., but some others (also expressed as dimensions afterwards) may have the flexibility of being introduced at will or on demand, so that they are present when they are required or when they are more opportune so that they can support the reach of some learning goal.

2.2 Related Work: Discussion

Some authors [2] criticize the traditional methods of instruction of the cognitive domain based on textbooks and basic practical lessons, saying that they have several limitations to help students remember or recognize certain knowledge and develop their understanding, their intellectual abilities and his skills.

Intuitively, having a group of students, all of them with different abilities, we can see that the traditional instructional method will be appropriate for the skills of a subset of students in the group, but not for the rest of it. At most there will be some students who, to achieve the same advantage as those of the aforementioned subgroup, will cost them an extra effort, others could simply give up. Here we review the following two strands to learning.

a) Theoretical approaches

Different approaches and theories have emerged to improve learning. For example, some of the most representative are:

Behaviorism. It is a psychology strand proposed by John B. Watson (1878-1958) that defends the use of strictly experimental procedures to study observable behavior and denies any possibility of using subjective methods such as introspection. [3].

Constructivism. It is a position shared by different trends in psychological and educational research. Among them are the evolutionary theory of Piaget, which states

that learning is a restructuring of cognitive structures, the student assimilates what he is learning under the prism of previous knowledge he has in his cognitive structures; the socio-cultural approach of Vygotsky, which affirms that learning is conditioned by the society in which we are born and develop, which implies the possibility of learning with the help of more skilled people; Ausubel's significant learning, which coincides with Piaget, that the starting point of all learning is knowledge and previous experiences and the most important factor that includes learning is what the student already knows and the current Cognitive Psychology, which they maintain that what is developed are fundamentally aspects related to basic cognitive processes, such as attention, memory and activities of storage and retrieval of information, as well as the amplitude of short-term memory, executive strategies and metamemory or meta-knowledge [4]. And others could be included here, but from the intuitions mentioned above, it can be seen that the learning process requires a more integral approach, which could even include all these theories, so that instruction can have an impact on a larger audience.

One of the problems here is that usually, the instructional design is not directed to groups of students with the same skills, rather they are applied to a heterogeneous audience where each student has different abilities.

b) Technology based approaches

There is also the strand of technology. One of the most representatives are the Intelligent Tutoring Systems that based on interaction and progress of students tries to guide students in a "kind of" personalized instruction. Early version perhaps were not well-implemented [5], nonetheless the architecture is still a base to integrate other technologies. For instance, in [6] is presented an e-learning platform for automatic course personalization based on the Felder and Silverman pedagogical approach, based on the correspondence between teaching styles and learning styles. The platform is based on *Diogene* a LMS (Learning Management System) another technological tool. This source also introduces some formalism and ontologies in order to deal with course personalization. Again, on the course personalization, [7] proposes a learner profile design model to facilitate the development of personalized e-learning systems. They use the learner profile design models proposed by IEEE PAPI [8], and proposed an agent-based E-learning system architecture in order to achieve personalization. However, it is not automatic, for instance, it is the user who creates a user profile capturing all his personal information including preference, actions, goals etc. on the other hand agent is in charge of dynamically discover learning objects that fits a user's learning requirements. On the same lines [9] presents a review of user's models and user's modeling approaches for adaptive web systems. Despite focusing adaptation for web systems, it also touches learning topic such as learning styles, affective states, work contexts, domain models and of course ITS, features shared with learning contexts. After an exhaustive review they conclude that "… adaptive educational systems rely mostly on user knowledge and learning goals capitalizing on the modeling and representation techniques established in the field of ITS." There are also integral approaches, for example, Chen et al. [10] propose a theoretical framework based on integrative goals and some multimedia principles.

Here the integrative learning goals for instructional design [11] are based on the idea that the design begins with the identification of learning goals (eg, preparing a cake,

which may include sub processes such as preparing the mix that includes the exact amount of all the ingredients, preheat the oven, bake it, decorate it, etc.). Sometimes, the goals are conceived as objectives that must reflect a certain human performance and sometimes as the capacities that must be acquired to reach a certain performance, in such a way that the goals have to do with a combination of several individual objectives that have to be integrated in a learning goal in such a way that the goals have to do with a combination of several individual objectives that have to be integrated into a learning goal.

2.3 Multidimensional Approach to Learning

The methodology we propose is not complicated. The integral term is not only adopted through the integrative goals, but we believe it is necessary to enrich the learning context through the integration of different factors that can connect instructors and students to achieve the learning goals effectively.

The main idea is that instructors have in mind or are aware that they can or perhaps should, resort and make use of different resources that support them so that the teaching-learning process is effective even when their audience is a heterogeneous group.

Thus, the intuition behind the multidimensional approach, proposes the consideration of as many factors as possible within each specific learning context, with the aim of providing stimulus to the greatest possible number of elements within the instructor's audience. As an example of these factors, we can mention the teaching techniques, the student's state of mind, the learning channels, etc. On the other hand, it is not only about integrating as many factors as possible, but about having the flexibility to apply the necessary factors in the appropriate context and time.

It is worth mentioning that apart from the pedagogical techniques and methodologies that the instructor can use, he could also resort, if possible, to the arsenal of tools that the technology provides him. For example: Learning Management Systems (LMS), e-learning tools, virtual reality, intelligent tutors, etc.

In order to respond to the needs expressed is that we adhere to the multidimensional model of Learning [12].

2.4 Model for the Multidimensional Approach to Learning

Following the integral approaches and the use of technology, in relation to instruction, there are a variety of dimensions or factors that intervene in the learning process (Fig. 1) and that should be considered, if we want to achieve the main goal that is the transfer of knowledge. These dimensions may vary according to different situations, some examples are mentioned here.

• *Student-Instructor Dimensions*: According to the people involved, two dimensions can be identified that must intervene collaboratively to achieve the goal of knowledge transfer. In the case of students, this goal is to accommodate a new piece of information or a new organization of information, within their repository of knowledge in their brains. When this is done, the student could modify their behavior or points of view, increase their skills, etc. For instructors, this goal should be to teach and have evidence

that knowledge has actually been transferred to the students' brains. These sub-goals involve dimensions that are perhaps simply assumed, but decisive, in order to achieve a combined effort to achieve the learning goal, that is, learners must really want to learn and instructors must really want to teach.

• *Dimension of the Instructional Model*: The instructional model is another dimension, different models have been proposed (*e.g.*, behaviorism, constructivism, active learning, etc.), each with strengths and weaknesses. All of them provide some truth and some approach to improve learning (example, instructor-centered learning, student-centered learning, learning focused on the interaction among instructors and students, etc.). There may be cases in which a model is used so effectively, that even students without interest are involved and guided towards a specific learning goal, however, depending on the domain, a model or combination of models must be selected in order to make efficient the instruction.

• *Dimension of the instructional domain*: This is another dimension; it is not the same football training, which includes mainly a physical activity that physics lessons, which could be mainly theoretical. It is clear that each domain demands specific skills from the learners, but it also determines which instructional method may be best to achieve an instructional goal.

• *Dimension of the learning channels*: One more dimension is given by the different types of students, according to the learning channels they prefer to use when they are learning, or that makes learning easier. Usually three types of apprentices are identified, according to their dominant learning channel, these are: auditory, those who learn best by listening; visual, those who learn best through visualization and kinesthetic for those who learn best by manipulating objects. Students also have different moods, different skills, etc., which, in combination with the learning channels, intervene in the efficiency of learning. We do not use just one learning channel, most people learn better by using more than one channel simultaneously. If the instructional design and content includes stimulus elements for these three learning channels, the efficiency will reach a larger audience.

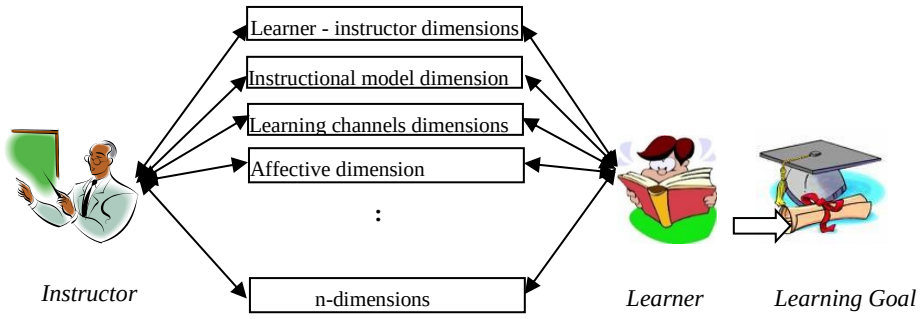


Fig. 1. Multidimensional model. Different dimensions intervene in the goal of knowledge transfer.

• *Affective dimension*: Here we are assuming that the affective state of a student and his/her instructor could also determine whether the learning objective is satisfactorily achieved. As a simple example, a student might have argued with his girlfriend just before attending the lesson in the classroom. This could lead to an effective state, which results in a distraction and lack of concentration in the classroom. As another example, the exposure method used by the instructor could cause a state of apathy on the student or could motivate them to concentrate and effectively achieve the defined learning goals. Emotional states are inherent to students; therefore, the affective dimension should be an integral part of learning.

Other dimensions may be present in the learning process. Thus, a LC could be used to group the different dimensions involved in any learning process. A LC can include Personal Learning Contexts (PLC), which are subsets of dimensions associated with specific people, whether apprentices or instructors. We can also identify the Group Learning Contexts (GLC), which can be restricted to the sum of the PLCs of the trainees and instructors who participate in a specific learning process.

However, in the model proposed here, in the term Learning Context (LC), "context" is borrowed from the Natural Language Processing (PLN) community, where it is defined as a consistent set of propositions describing the set of beliefs of a person. Thus, unlike a learning environment that includes external elements that influence the learning process [13], a learning context is intended to be an internal personal view of an environment and thus a more accurate and complete view of the learning factors that influence such process, individually in each student, that is, where even external aspects really must be considered as each student perceives them and individualizes them within their learning process.

Integral approaches to learning are based on theories such as integrative goals. Thus, following these approaches, identifying, integrating and considering LCs within the learning process would provide a more complete and efficient tool to achieve knowledge transference.

3 Discussion

Here are some reflections on the dimensions described above. These are described by means of examples since it is not a generic recipe, each particular LC imposes its own demands and conditions. The same applies to the dimensions described that are only examples since there will be many more factors that intervene in learning and that have not been included here but will appear in some LC.

• Among the dimensions that are decisive in a LC of the proposed model, are those that refer to the roles of the Student and the Instructor. For example, if you have a teacher without a vocation, with lack of commitment or ethics, etc., or on the other hand you have a student, lacking commitment, without ethics, or who does not have enough energy to learn, then, this dimension will not be fulfilled. There can be multiple reasons why some of the participants do not comply with the condition imposed here, in which the teachers must want to teach and the students must want to learn and therefore, the teaching-learning process cannot be carried out effectively.

• Regarding the dimension of the Instructional Model. It is evident that the teacher must be aware of the diversity of models already proposed and apply them on demand, that is, just as the student must learn to learn, the teacher must learn to teach besides having mastery of his area of knowledge.
• The dimension of the instructional domain is related to the previous dimension, so that knowing different instructional models the teacher should select those that are appropriate according to the content of their subject.
• In the case of the dimension of the learning channels, if an instructor had the possibility of using the technology (e.g. RV), it would be ideal. However, it is more important that you are aware that your group is made up of students of different types and try to design, create and use instructional content that contains visual, auditory and kinesthetic resources, as far as possible, in order to stimulate these channels of their students.
• In the case of the affective dimension it could be very difficult for a teacher to determine the affective state of each one of his students in order to act accordingly and continue to promote their learning. However, you could use standard measures that do not affect your students negatively and emotionally. For example, promoting participation and avoiding at all costs mocking or ridiculing someone for wrong answers or obvious questions. Other examples could be to promote respect in class, celebrate success and even errors to focus on as areas of opportunity.

On the other hand, technology has proven to be useful as a learning tool, since it has contributed to providing tools such as learning objects and repositories, learning management systems (LMS), content management systems (CMS), Intelligent Tutoring Systems (ITS) and Virtual Reality (VR) for training, among others. In addition, there is a key point in the use of technology in comprehensive approaches allows us to integrate different dimensions involved in the learning process, therefore provides the tools to increase the efficiency of the process.

The multidimensional model of Learning partially has been implemented is different Virtual Reality Training Systems. As an example, Virtual Reality allows recreating learning environments; where a user can also navigate and interact with the environment, providing different stimuli that turn it into a useful tool to integrate different aspects, such as those mentioned in the model for the creation of learning contexts, according to the proposed model (Fig. 1). For example, it is able to naturally support the dimension of learning channels.

A study conducted in 2009 [11] shows that we have almost the same preference for the three learning channels:

a) 37% of learning is tactile or kinesthetic, through the movement of things, touching and doing.
b) 29% of the learning is visual, through photos and images.
c) 34% of learning is auditory, through sounds and words.

However, it is known that while we learn we use more than one sensory channel. Within the literature of study strategies, this is known as multimodal study strategy and according to Fleming [14], the majority, approximately 60% of any population, falls

into this category. Each learning style uses different parts of the brain, so the more channels are involved during learning, the more we remember what we learn [13].

Although in the literature of learning [14, 15], different styles are mentioned such as reading/writing, logic, verbal, etc., here we focus on the primary sensory channels [10], whose percentages of preference were listed above. We believe that the importance of these learning channels lies in that they involve some of our basic senses. Here technologies such as Virtual Reality might be able to stimulate, to some degree, these learning channels and might include images, text, animations and audio.

It is worth mentioning here that the multidimensional model of Learning is rather a general model that in the near future might be formalized. Technology might provide the means to implement such model as is the case of both: a) the VR training systems already mentioned and that have proved to improve learning [16]; and b) research as the examples included in the section of related work.

The model presented explicitly demands some flexibility to introduce the described dimensions when these are required, this in accordance with personalized and adaptive instruction, technology can be of great help, at least partially, and it is the intelligent tutoring systems (ITS). The ITS could store information related to the student's progress in such a way that their strengths and weaknesses are registered, in order to know where the student needs support and to try to decide which dimensions would be the most adequate to support the student to reach the learning goals.

Thus far, there is not a comparative point for the products in a) and b), they are different and complementary and both fall under the multidimensional model approach to learning presented above.

4 Conclusions

Traditional methods of instruction can be considered incomplete, in the sense that they do not include different dimensions that intervene in specific learning processes. For example, in some Latin American countries there are educational institutions that evaluate the efficiency of the instructors, based on the percentage of graduated and failed students. This point of view is incomplete, since they do not consider some of the dimensions involved; For example, students may think that the learning process is only the responsibility of the instructors. This position seems to assign all the responsibility to the instructors, which is an incomplete criterion to evaluate the efficiency of learning and worse, could send the wrong message to some students and some educational authorities. Even the use of technology could provide incomplete methods.

Distance learning can be an alternative for problems of lack of access to education; however, if in a distance course the instructional content is only delivered to students along with some instruction to follow, and then some evaluations are applied, the student-instructor interaction may be reduced. It could be appropriate for people with autodidact orientation, but it leaves out other types of students.

On the other hand, the ideal learning context could be almost impossible, unless the instruction is personalized, in which case it could be less practical and surely expensive. We have to settle for including as many dimensions as possible in the learning process,

but perhaps more important is to be aware of the different dimensions involved in specific learning processes successfully integrated in order to make effective the teaching-learning process.

References

1. Asim, S.: Exploring the Relationships among Learning Styles, Annotation Use and Reading Comprehension for Foreign Language Reading in a Hypermedia Environment. Thesis of Master of Arts in English Language Teaching Bogazi University (2003)
2. Chwen, J.C.: The design, development and evaluation of a virtual reality based learning environment. Australasian Journal of Educational Technology 22(1), pp. 39–63 (2006)
3. Tangient LLC. Teorías del aprendizaje: Conductismo. Wikispaces http://uoctic-grupo6.wikispaces.com/Conduc-tismo. Consultada en Agosto 2016.
4. Tangient LLC. Teorías del aprendizaje: Constructivismo. Wikispaces http://uoctic-grupo6.wikispaces.com/ Constructivismo. Consultada en Agosto 2016.
5. Ferster Bill. Intelligent Tutoring Systems: What Happened? E-learning industry. Instructional design. January 21, 2017. https://elearningindustry.com/intelligent-tutoring-systems-what-happened. Available on internet oct the 1st, 2018
6. Enver Sangineto. (2008). An Adaptive E-Learning Platform for Personalized Course Generation. In: Claus Pahl (ed) Architecture Solutions for E-Learning Systems. IGI Publishing (2008)
7. Wei, X., Yan, J.: Learner profile design for personalized eLearning systems. In: The International Conference on Computational Intelligence and Software Engineering (CiSE 2009), pp. 1–4 (2009)
8. IEEE. IEEE P1484.2/D7, 2000-11-28. Draft standard for learning technology. Public and private information (papi) for learners (papi learner)
9. Brusilovsky, P, Millán, E.: User models for adaptive hypermedia and adaptive educational systems. In: Brusilovsky, P., Kobsa, A., Nejdl, W. (eds.) The adaptive web: methods and strategies of web personalization, pp. 3–53. Springer-Verlag, Berlin, Heidelberg (2007)
10. Chen, C.J., Toh, S.C., Fauzy, W.M.: The Theoretical Framework for Designing Desktop Virtual Reality-Based Learning Environments. Journal of Interactive Learning Research 15(2), pp. 147–167 (2004)
11. Gagné, R.M., Merrill, M.D.: Integrative Goals for Instructional Design Chapter 5, pp. 127–140 (1990)
12. Pérez-Ramirez, M., Ontiveros-Hernández, N.J.: Virtual reality as a comprehensive training tool. In: Proceedings of Workshop on Intelligent Learning Environments WILE, pp. 203–215 (2009)
13. Advanogy. Discover your Learning Styles – Graphically! http://wwwlearning-styles-online.com/.Site (August 2016)
14. Neil Fleming Teaching and Learning Styles VARK Strategies. Neil D Fleming. (2006)
15. Advanogy. Overview of Learning Styles http://www. learning-styles-online.com/overview/.Site (August 2016)
16. Ayala-García, A., Galván-Bobadilla, I., Arroyo-Figueroa, G. *et al.*: Virtual reality training system for maintenance and operation of high-voltage overhead power lines. Virtual Reality 20(1), pp. 27–40 (2016)

ISSN 1870-4069

Design of an Effective Assessment-Feedback Scheme through a Virtual Learning Environment

Héctor Jiménez-Salazar, Tiburcio Moreno-Olivos, Alfredo Mateos-Papis

Universidad Autónoma Metropolitana, Unidad Cuajimalpa
{hjimenez,tmoreno,amateos}@correo.cua.uam.mx

Abstract. The teaching-learning process faces big challenges to incorporate information and communication technologies (ICTs) in an effective way. Considering virtual learning environments (VLEs) which provide higher-education students with the possibility to solve assessments focused on solving problems with multiple-choice questions (MCQ), and which offer the student feedback in relation to the answers, this paper presents a proposal to design the medular part of such kind of VLEs: an assessment-feedback scheme. This proposal considers the fundamentals and the results of various works on feedback and on learning-processes. Additionally, the scheme should capture the student's browsing for further analysis.

Keywords: feedback, assessment, virtual learning environment.

1 Introduction

From the seminal work of Sadler [12] to date, several papers have been published, which analyze the topic of feedback in assessment, for example papers about: sustainable feedback, feedback models for learning [12, 13, 6]; formative assessment and feedback, auto-regulated learning [2, 3, 9, 10, 16]; perceptions on the feedback process, practice development of sustainable feedback [15, 4], among others. Also, the use of ICTs on educative processes has allowed the emergence of Learning Management Systems (LMS) or VLEs, and this has powered electronic assessment (e-assessment) and feedback in e-assessment. Several initiatives have emerged to find/create better ways to use ICTs in the teaching-learning process. For example, concerning feedback, some works try to capture the perception of students [1], or personalize the LMS according to the user [11], or analyze the recorded behavior of students when using a LMS, through Weka [14].

The problem that this article addresses is the lack of knowledge, that the authors have perceived, about the effectiveness of the feedback that students receive when they learn a subject though the use of exercises.

In order to study this problem, this article proposes the use of a VLE feedback scheme, as an effective form, not only to analyze the way in which students request feedback when they solve problems, but also to serve as a support tool to help students with their learning process. This scheme should explore the relations among three incidental factors on feedback: 1) The depth

of understanding that the student has when facing a problem, 2) If the problem statement is clear to the student, and 3) The level of certainty that the student has when solving a problem (an exercise).

The remaining sections of this article are organized as follows. Section 2 offers a survey on the relevance of feedback in the learning process, which is remarked by seven principles [9] of good feedback practices in self-assessment, and which form the basis of the proposal of his article. Section 3 presents a brief summary about some implementations of feedback using VLE, which can contrast with the present work. Section 4 presents the aforementioned proposal. The closing section presents the conclusions of this work.

2 Background

An effective feedback, periodically delivered, may help students to know how well they are doing their work. Feedback may help them decide how to be more effective or how to enhance the management of their workload, and at the ending, be satisfied with their learning process.

When a student lacks prior knowledge, necessary to be able to learn a subject, feedback may become a very useful tool. Currently, the increasing number of higher education students that a professor has to attend prevents face-to-face feedback with all students. In such conditions, ICTs can become an important resource to fulfill the purpose of feedback.

A common situation in the progress of a course is to leave feedback for later, or even to postpone it to the end of the course not giving students the opportunity to learn from their mistakes, and without allowing them to take advantage of feedback in the subsequent assessments of the course [16], [5].

It has been found that the remarks of low-value feedback focus on issues related to language (grammar, punctuation, etc.), factual accuracy or bibliographic finesse. This type of feedback is mostly cryptic; use marks on the edge of the sheet, punctuation marks, or imperatives and affirmations (such as "explain") (Lea & Street 2000 cited in [6]). Besides, the students can only get an incomplete or uncertain understanding ([6]). Research on student's perceptions and experiences on feedback comments reveals a similar status.

It is difficult for professors to offer feedback, particularly due to the unfavorable existing conditions to help students, for example, feedback requires a lot of, unavailable, dedication time. Nevertheless, the effect of feedback can benefit, in a significant way, by getting good and timely information to the students, and verifying that they use that information when doing their work.

3 Related Work

This section presents published ideas that deal with feedback, from the point of view of ICT.

In [11] the authors propose an online framework to help in the preparation of a selected LMS so that it can automatically detect the learning style of its

users (the condition which favors the better learning of users) to automatically personalize its learning environment.

In [14] the authors present a scheme that looks for the behaviors that are more significant towards identifying the learning styles of the LMS users, considering affective and performance patterns, for example "how many exercises did the user visit?", or "were these exercises basic, intermediate or advanced?".

In [1], the author investigates the perception, in terms of constructiveness, of undergraduate and graduate accounting students about the difference between: 1- Receiving timely feedback (5 days) against extremely timely feedback (2.5 days or even instant), and 2- Receiving manually written formal assessment feedback against automatically generated feedback.

In [7] the authors evaluate the role of e-assessments feedback through the creation of two study-groups of undergraduate engineering students. Both groups received teaching based on the same traditional way, enhanced with VLE components. The difference lied in the form of e-assessments and feedback given. Group 1 (CG) gave students just one opportunity of e-assessment, indicating only if each answer was correct or not. Group 2 (EG) gave students unlimited opportunities of e-assessments, giving clues as to why each incorrectly-answered question was incorrect. Students of Group 2 had an overall better performance.

4 Feedback in Assessment

Nicol & Macfarland-Dick [9] established seven principles for good practices in feedback. These principles are related to: clarify what good performance is (goals, criteria and expected standards); facilitate the development of self-assessment (reflection) in learning; deliver high-quality information to students about their learning; encourage dialog between peer professors; encourage positive motivational beliefs; provide opportunities to close the gap between current and desired performance; and provide information to professors to help shape teaching. The authors of this work consider that four of the mentioned principles are present in the proposed feedback scheme:

1. It helps to clarify what good performance is (goals, criteria, expected standards).
   The scheme would include features to develop some strategies, such as the following:
   - The application of a questionnaire to the students (in the VLE) to find out to what extent they understand the objectives, criteria and evaluation standards.
   - The offering of examples of performance (in the VLE). Examples are effective because they make explicit what is required and define valid standards for students to compare their work.
2. It facilitates the development of self-assessment (reflection) in learning.
   The feedback scheme proposed should provide a sequence of simple questions to generate confidence in the student when performing the self-assessment

and when looking at the evaluation given by the system; these questions may be repeated several times.

3. It delivers high-quality information to students about their learning.
   Specifically, the scheme provides:
   - Opportunistic feedback to allow students to correct their work while still learning a subject in the course.
   - Corrective advice, which also indicates strengths and weakness.
   - Online assessments and consequential feedback at any time, anywhere, and as many times as students wish.

   The VLE should have assessments (and corresponding automatic feedback) available at all times to reinforce this good practice.
4. It provides opportunities to close the gap between current and desired performance.
   Given a failure to answer correctly an MCQ, the scheme should offer the student indications of the possible reasons for that failure.

The next sub-section gives details of a scheme which, in order to offer feedback to students, captures their browsing.

4.1 Complexity of Distractors

An MCQ in the scheme would be composed of a question and four answers to choose from, which are: the right answer and three distractors. The distractors are classified in the following way:

1. Distractor 1 (D1). The most easy to discard; it requires the use of a single definition to realize that it is wrong.
2. Distractor 2 (D2). It requieres the use of analysis to discard, that is, it requires to combine data, obtained from the question, and concepts, to observe a contradiction.
3. Distractor 3 (D3). This is an answer that has a small mistake, it resembles to be the right one. It is, of course, the hardest one to elucidate to discard.

Each one of the distractors is associated with a browsing level in the feedback sequence (a feedback-level), and a level of error. Thus, D1, D2 and D3 are associated to Knowledge, Application and Test level, respectively. At next subsection feedback levels are explained.

4.2 Feedback Levels

This approach is conceived as a sequence of steps: from the explanation of why the student's response to an exercise was incorrect, to an example that explains how knowledge is used to solve this type of exercise, going through the concepts of the subtopic, related to the exercise. These would be the feedback-levels that the learning platform should provide, that would allow the registration of various variables in the comments offered. The navigation sequence should offer three feedback-levels:

a) Knowledge level is a reminder about the implicit concepts of the exercise, where students can review each concept or continue with other related concept, if required. At the ending, students should be able to decide if they continue to solve more exercises or continue in the browsing sequence with an example that illustrates the application of concepts.

b) Application level shows the application of concepts in a simple example-problem, where there is always the possibility to visualize previous concepts, as many times as desired. In addition, the student should be able to change the example or to choose either to end, to continue with the exercises, or to move to other feedback-level.

c) Test level proposes students to solve an exercise related to the same subtopic, and with the same level of complexity as the one reviewed in the Application feedback-level, offering the opportunity to receive similar feedback if his answer was incorrect.

In the case of success, the student could return to do other exercises or move to any desired feedback-level, for the consolidation of learning.

4.3 Example of a Screen Transit-Plan for the Assessment-Feedback Scheme

The following list presents a transit-plan among the diverse screens of the assessment-feedback scheme proposed. Figure 1 presents a diagram of the transit-plan.

A. A list of exercises appears on the screen.
B. The student chooses one exercise to solve.
C. The statement of the chosen exercise appears on the screen, together with a list of answers to choose from, among which only one answer is correct.
D. The student picks one answer (the student should really work without choosing randomly).
E. The scheme detects if the answer was correct or not. If it was incorrect, according to the answer the system classifies the level of error, to offer the student one of three feedback-levels.
F. This feedback-level presents an explanatory text oriented to the understanding of concepts.
G. This feedback-level presents an example with application of concepts.
H. This feedback-level presents a similar exercise to the one solved, and its solution-approach.
I. The scheme indicates that the solution of the problem was correct.
J. In case of having answered incorrectly, the student might choose to accept, or not, the feedback-levels offered given by the system, but, in any case, he can choose to transit to any one of the options A, F, G or H.

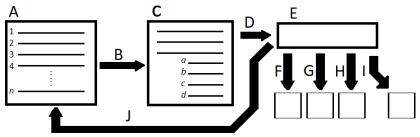


Fig. 1. Relations of the components of feedback in the VLE.

4.4 Example of Feedback at MCQ Solution

The topic of this example is 'Graph Theory', which is offered in the first year of the 'Information Technologies and Systems' undergraduate program. Among the following MCQ, items (a), (b) and (d) are distractors of type 3, 1, and 2, respectively.

If $\mathcal{G}' = (\mathcal{V}', \mathcal{A}')$ is a subgraph of $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ and $v \in \mathcal{V}'$, then it is true that:

a) If v is an isolated vertex in $\mathcal{G}'$, then it is an isolated vertex in $\mathcal{G}$.
b) $\mathcal{G}'$ is connected.
c) If there is a cycle $v - v$ in $\mathcal{G}'$ then there is a simple cycle $v - v$ in $\mathcal{G}$.
d) There is a non-trivial cycle $v - v$ in $\mathcal{G}'$.

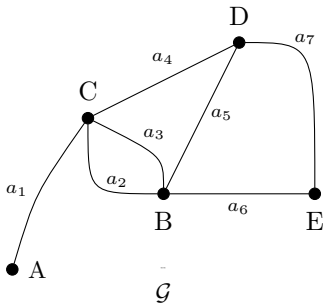
Let us see how would the browsing scheme be. It is important to take into account that there are many previous concepts. The content of each one the three feedback-levels offered could be:

Knowledge Pay attention to the following definitions:

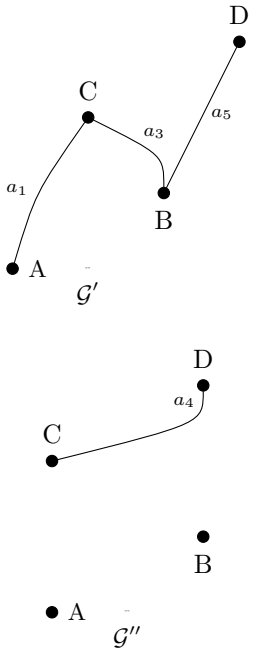
1. A connected graph must satisfy the definition: *A graph* $\mathcal{G} = (\mathcal{V}, \mathcal{A})$ *is* connected *if for all pair of vertices,* $x, y \in \mathcal{V}$ *there exists a path* $x - y$.
2. A subgraph $\mathcal{G}'$ is constructed from a given graph, $\mathcal{G}$, accordingly the definition: *The graph* $\mathcal{G}' = (\mathcal{V}', \mathcal{A}')$ *is a* subgraph *of* $\mathcal{G}$*, iff* $\mathcal{V}' \subset \mathcal{V}$*,* $\mathcal{A}' \subset \mathcal{A}$*, and for every one of its edges* $a \in \mathcal{A}'$ *its extremes hold* $v_1, v_2 \in \mathcal{V}'$.

Application In order to build a subgraph we must choose $\mathcal{V}'$ and $\mathcal{A}'$. We can choose any subset of vertices from $\mathcal{V}$ to define $\mathcal{V}'$.

For example, let $\mathcal{G}$ be given by the following graph:



For instance, we can define $\mathcal{V}'$ as $\{A, C, B, E\}$, or $\{E, D, A\}$, or also $\emptyset$. For the next step, we add the edges as we want. If we defined $\mathcal{V}'$ as $\{A, C, D, B\}$, the edges of this subgraph could be a subset taken from $\{a_1, a_2, a_3, a_4, a_5\} \subset \mathcal{A}$. A choice might be $\{a_1, a_3, a_5\}$, or $\{a_4\}$ or $\emptyset$. For $\mathcal{V}'$ we could build a subgraph $\mathcal{G}'$ with edges $\{a_1, a_3, a_5\}$, or build another subgraph $\mathcal{G}''$ with edge $\{a_4\}$.



We remark that $\mathcal{G}'$ is connected and $\mathcal{G}''$ does not.

Test For the graph $\mathcal{G}$, we now consider another subset $\mathcal{V}'' \subset \mathcal{V}$, for example $\{B, C, E\}$. Construct three subgraphs, two connected and one disconnected.

5 Conclusions and Future Work

This paper presented an informal operation specification of a proposed assessment-feedback scheme, which would be part of the kind of VLEs which provide higher-education students with the possibility to solve assessments focused on solving problems with MCQs, and which offer feedback in relation to a possible wrong answer. In the case of an incorrect answer from a student, the complexity of having discarded options to choose that answer becomes the key to offering the student one of several feedback-levels. One level can present the student with important conceptual content for its study; another level can present the student with useful relationships of concepts; and another level can show the student examples with applications similar to the exercise answered erroneously. This design has basis on the principles of good feedback-practices in self-assessment.

The main subsequent work would be to implement the proposed scheme in a VLE. Subsequently, it would be interesting to explore how students browse and what their regularities are, based on their behavior and their school performance. In addition, it would be important to broaden the scope of this proposal, for example, to have the possibility to, automatically, generate specific exploration spaces for a question from an ontology (a set of related concepts) of the same subject of the question. There is also interest in moving from assessments of MCQs, to assessments with open questions, which would require the application of natural language processing, for the generation of the related automatic-feedback with short answers.

Acknowledgments. This work was partially supported by the Department of Information Technologies of our university, UAM-C.

References

1. Bayerlein, L.: Students' feedback preferences: how do students react to timely and automatically generated assessment feedback? Assessment & Evaluation in Higher Education, 39(8):916–931, Routledge, https://doi.org/10.1080/02602938.2013.870531 (2014)
2. Boud, D.: Sustainable assessment: rethinking assessment for the learning society. Studies in Continuing Education, 22(2):151–167 (2000)
3. Boud, D., Molloy, E.: Rethinking models of feedback for learning: The challenge of design, Assessment and Evaluation in Higher Education, 38(6):698–712 (2013).
4. Carless, D., Salter, D., Yang, M., Lam, J.: Developing sustainable feedback practices. Studies in Higher Education, 36(4), 395–407 (2011)
5. Gibbs, G.: Using assessment strategically to change the way students learn. In: Assessment Matters in Higher Education, Sally Brown & Angela Glasner, editors. Buckingham: Society for Research into Higher Education and Open University Press (1999)
6. Hounsell, D.: Towards more sustainable feedback to students. In: Rethinking Assessment in Higher Education. Learning for the longer term, David Boud & Nancy Falchikov, editors, pp. 101–113. Londres y Nueva York: Routledge (2007)
7. Limniou. M, Smith, M.: The role of feedbak in e-assessments for engineering education. Educatin and Information Technologies, 19:209–225 (2014)
8. Moreno, T.: Evaluación del aprendizaje y para el aprendizaje. Reinventar la evaluación en el aula. Ciudad de México: Universidad Autónoma Metropolitana-Cuajimalpa (2016)
9. Nicol, D., Macfarlane-Dick, D.: Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education, 31(2):199–218 (2006)
10. Nicol, D., Thomson, A., Breslin, C.: Rethinking feedback practices in higher education: A peer review perspective. Assessment and Evaluation in Higher Education, 39(1):102–122 (2014)
11. Núñez, L., et al.: A framework for automatic identification of learning styles in learning management systems. Research in Computing Science, 106:59–68 (2015)
12. Sadler, R.: Formative assessment and the design of instructional systems. Instructional Science, 18:119–144 (1989)

13. Sadler, R.: Formative assessment: revisiting the territory. Assessment in Education, 5(1):77–84 (1998)
14. Salazar, G., et al.: Behavioral patterns for automatic detection of learning styles in learning management systems: a case study. Research in Computing Science, 106:69–77 (2015)
15. Torrance, H., Pryor, J.: Investigating formative assessment. Buckingham, Open University Press (1998)
16. Yorke, M.: Formative assessment and its relevance to retention. Higher Education Research and Development, 20(2):115–126 (2001)

ISSN 1870-4069

Emotion Recognition for Education using Sentiment Analysis

María Lucia Barron-Estrada, Ramón Zatarain-Cabada, Raúl Oramas-Bustillos

Tecnológico Nacional de México / Instituto Tecnológico de Culiacán,
Culiacán, Sinaloa, Mexico
{lbarron, rzatarain}@itculiacan.edu.mx

Abstract. Most systems that recognize emotions that are used in learning systems are oriented towards the detection of basic emotions such as happy, sad or angry. However, an intelligent tutor system must be able to detect, in a student, secondary emotions that have to do with learning as boring or frustrated. In the case of sentiment analysis in text, most of the research focuses on recognizing polarities (positive and negative). In this paper we present a sentiment analyzer for the recognition of polarities and emotions centered on learning using textual phrases written in Spanish in the field of computer science that can be used in an ITS (or any learning environment) to detect sentiments and/or emotions of the students. We also present the results obtained from the sentiment analyzer, which are very promising with 88.26% of accuracy.

Keywords: sentiment analysis, machine learning, deep learning, intelligent learning environments, learning centered emotions.

1 Introduction

The application of the AI in educational environments (intelligent environments) aims to create intelligent teaching software that adapts to the pace of student learning. The objective is always the same, to create a scenario that allows the student to be immersed in an educational environment to help the student learn.

A learning environment is a setting, traditionally a classroom, where students learn [21]. Although there are diverse virtual learning environments, such as intelligent learning environments (ILE) or intelligent tutoring systems (ITS), the main objective of a learning system is to facilitate different learning activities [5]. In this area of research related to intelligent learning systems, various fields are combined such as pedagogy, psychology, cognitive sciences, artificial intelligence, among others, where each of them contributes their vision to the development of the discipline [23]. In the case of modern ILE and ITS, they already incorporate the automatic recognition of emotions [2,6] through the human face, brain signals, and voice. In the case of text recognition or sentiment analysis (SA), most of the work focus on recognizing polarities [13,25] and a few of them recognize basic emotions [19,22].

This work presents a system that carries out a sentimental analysis based on polarities and learning-centered emotions. The emotions that are detected are boring,

frustration, neutral, excitement, and engagement. The recognition of emotions can be used in any ITS or ILE to provide personalized feedback to the student after solving an exercise and to adapt the exercises of the learning system based on the emotions detected.

The main contribution of this work is to use emotional recognition techniques, based on text, and their integration in intelligent learning environments such as an ITS or an ILE and combining traditional algorithms of emotion classification with current deep learning (DL) techniques. Another contribution is the design, construction, and testing of a corpus of texts labeled with emotions centered on learning and a corpus labeled with polarity, both datasets in the domain of computer science.

This paper is organized as follows: Section 2 describes the related work. Section 3 presents the methodology to build the sentiment analyzer. Section 4 shows the results and discussions. Finally, in Section 5 the conclusions and future work are described.

2 Related Work

This section presents a review of sentiment analysis with deep learning techniques and sentiment analysis in learning environments.

2.1 Sentiment Analysis Using Deep Learning

Deep learning has revolutionized the way of solving problems that were previously done with traditional Machine Learning (ML) techniques. For example, in [9] the authors describe a Dynamic Convolutional Neural Network (DCNN) architecture for sentiment modeling of sentences. The network supports sentences of variable length and is able to capture relationships between words. The authors conducted experiments with Twitter phrases with binary and multiclass prediction.

In [11] the author uses a Convolutional Neural Network (CNN) with a convolution layer and pre-trained word vectors over 100 billion words from Google News for sentence-level classification tasks. His model, despite being simple, achieved excellent results for different corpus used. The researchers also concluded that pre-trained vectors are extractors of universal characteristics that can be used for various classification tasks.

In [20] the authors propose a CNN-LSTM model consisting of two parts: a regional CNN and an Long Short Term Memory (LSTM), to predict the qualifications of texts with a valence scale. The regional CNN unlike a conventional CNN that considers a complete text as input uses individual sentences as regions so that it can be weighted according to the valence scale. Finally, this regional information is sequentially integrated into all regions using LSTM for prediction with valence scale. The experimental results show that the proposed method outperforms the lexical, regression-based and ANN-based methods proposed in previous studies.

In [18] they propose simple, trained models with annotations where each sentence is annotated with two classes as negative, positive at the sentence level, and they also model the linguistic role of the lexical feelings with LSTM, the words of denial and the words of intensity. The results show that the models are capable of capturing the

linguistic role of sentimental words, words of denial and words of intensity in sentimental expression.

In [15] they used a lexical-based approach to apply semantic rules and then used a DeepCNN with pre-trained vectors at the character level to capture morphological information of each word to determine how those words are formed and their relationship to others. They then used a Bi-LSTM bidirectional network that produces the representation of the characteristics of the whole sentence from pre-trained vectors at the word level. For the training of the model, 100 billion words of Google News called Word2Vec [14] and TwitterGlove [17] of Stanford University were used. In this work, they report an accuracy of 86.63 %.

2.2 Sentiment Analysis in Learning Environments

In [10] a research paper is presented that proposes to take into account the opinions of the users in order to elaborate an evaluation in this respect using hybrid learning methods based on SVM and Hidden Markov Model (HMM) in an e-learning environment. The authors indicate that sentiment analysis becomes more difficult and challenging when it is done for e-learning blogs.

In [8] a model of sentiment analysis is presented to extract the opinions of the students to evaluate the quality of a course in two steps. In the classification of opinion, machine learning methods have been applied to classify an opinion as positive or negative for the publications of each student. Then, the extraction of opinion is used to extract characteristics, such as teachers, exams and resources, from the content generated by the user for a specific course.

In SA-E [1] the authors analyze the feedback of students who take a class in real time via Twitter messages with two machine learning algorithms: Naïve Bayes and SVM to identify the positive or negative sentiments of the students in real time. The objective was to collect the comments and analyze these data to help improve teaching.

SentBuk [16] is an application that runs on the social network Facebook that extracts information about the user's sentiment. It uses the opinions that students write and classifies them as positive, negative or neutral to detect significant emotional changes. This information is used to enrich the SentBuk system, so that they adopt the tasks that will be proposed to the student, based on their emotional state. These emotions are also used as useful feedback for the teacher of the course. The classification method implemented in SentBuk follows a hybrid approach: it combines lexical and traditional machine learning techniques.

In [4] present an algorithm based on an ANN to predict the student's dropout in MOOC courses using the analysis of sentiment and show the importance of the sentiments of students in this task.

In [12] student opinions are collected in the form of running text and sentiment analysis is performed to identify important aspects together with the orientations using supervised and semi-supervised ML techniques.

In the work of [7], they use sentiment analysis with supervised machine learning algorithms of type SVM, Naïve Bayes, nearest neighbors (KNN) and ANN to find the polarity of the comments of students based on the predefined characteristics of teaching and learning. The study involves the application of a combination of machine learning and natural language processing (NLP) techniques in student feedback data. The results

were compared to find the best performance with respect to several evaluation criteria for the different algorithms.

In [19] they explore NLP and machine learning techniques based on student comments to help university administrators and teachers address problem areas in teaching and learning. The system they developed analyzes student comments in online course surveys and real sources to identify the polarity of feeling, expressed emotions and student satisfaction. They also made a comparison with the results of the direct evaluation to determine the reliability of the system.

3 Methodology for Building and Integration of the Sentiment Analyzer

The proposed methodology for the development of our work is based on the inductive scientific method that is organized into four stages: analysis, design, implementation, and testing. The monitoring of each stage will allow reaching the solution of the development of the system that performs the analysis of feelings.

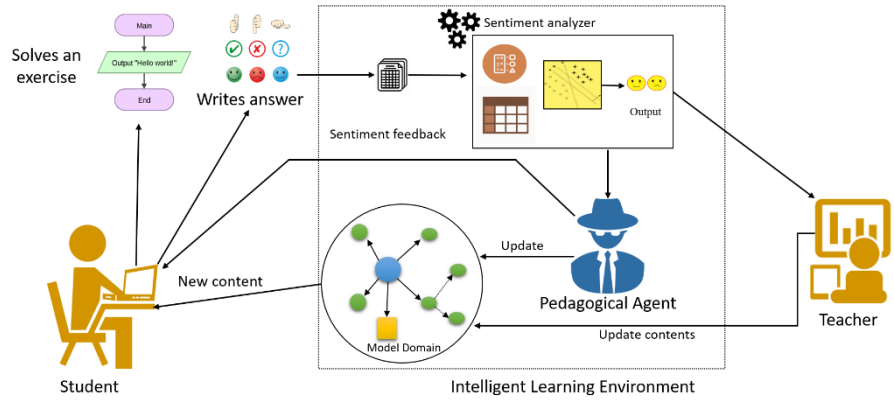


Fig. 1. The general scheme of the system to perform sentiment analysis with application to ILE and ITS.

3.1 Work Analysis and Design

In the analysis stage, the state of the art review was carried out to know which approaches have been used to deal with the problems related to sentiment analysis. A review and analysis of frameworks that are used in sentiment analysis were also carried out. In relation to the design stage, Fig. 1 illustrates the general scheme of the system to perform sentiment analysis in an ILE or ITS for a computer science domain. From the previous diagram, the student solves an exercise in the area of computer science provided by the ILE or ITS by means of responding with texts to the questions. These responses are the input to the SA system (sentiment analyzer) to perform the detection of emotion. For the detection of emotions, an DL model with its own corpus is used. A pedagogical agent uses the detected emotion to perform two activities: first, to send a

personalized feedback to the student based on the emotion detected; second, to present the contents of the student's domain model based on those emotions.

To design the best system architecture to perform SA, we first need to find a model that is suitable for the detection of learning-centered emotions. To find the best model, different tests were performed with different architectures. Fig. 2 shows the design for a DL model, which is a CNN-type network in combination with another LSTM network.

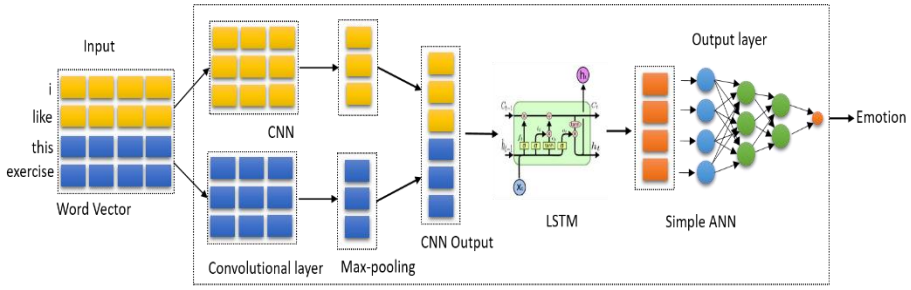


Fig. 2. CNN and LSTM Model.

The model receives as input a vector of words in a format called *one hot encoding* that encodes a text to a numerical format for the DL model to perform its prediction work. Subsequently, a CNN network takes this input vector and divides the information into smaller parts using a convoluted layer to detect patterns or characteristics, and combining this information in the deepest layers of the network. The output of the CNN network represents a feature vector that serves as an input to an LSTM network that processes that information in search of new patterns to pass that information to a simple neural network that is in the output layer and that classifies the text of entry into an emotion.

Once you have the best DL model, the next step is to establish the architecture of the model to perform sentiment analysis (Fig. 3). The emotion recognizer does the following: given an input text, the polarity of the text is first determined using the DL model presented in Fig. 2. Depending on the result of the positive or negative classification, the same model is used to determine whether the Text belongs to a category as engaged or excited or either frustrated or boring.

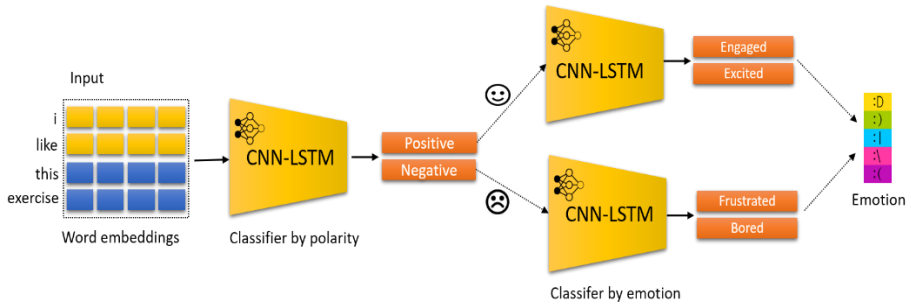


Fig. 3. SA Architecture for emotion recognizing.

3.2 Implementation and Integration with the ILE Java Sensei

Java Sensei [3, 24] is an ILE to teach programming in the Java programming language. The goal of the ILE is to analyze the cognitive and affective states of the student, to personalize their learning rhythm depending on those states. Once the best DL model was obtained to perform SA, the next step is the integration of the module to Java Sensei. The ILE Java Sensei uses a layered design to organize all its modules and components. This design allows the system to be scalable and easy to maintain. Information flows between layers, modules, and components through interfaces. Fig 4. Illustrates the architecture of Java Sensei.

The tutor's layer contains four modules: The ITS module that coordinates the entire tutoring system; the domain module contains the knowledge of the expert; the recommendation module is responsible for adapting the content presented to the student based on the evaluation grades based in Likert-scale provided by the students for each educational resource that they use, and the student module stores all the cognitive and emotional information of the student.

The affective layer contains the affective part of the system. It contains a module for the recognition of facial expressions as well as a module to recognize educational emotions based on the answers of the students' exercises, which is presented in this paper. The fuzzy logic module implements the mechanism of inference of the system based on the emotions detected by the facial expression recognizer and the sentiment analyzer.

Other pedagogical aspects such as the quality of the answers are also taken into account to feed these fuzzy rules. A pedagogical agent uses this information. The web content layer contains the corpus of sentimental text that is used to train the SA module, a repository of exercises and educational resources as well as the rules of the fuzzy logic system. Finally, the data layer stores the information produced by the students.

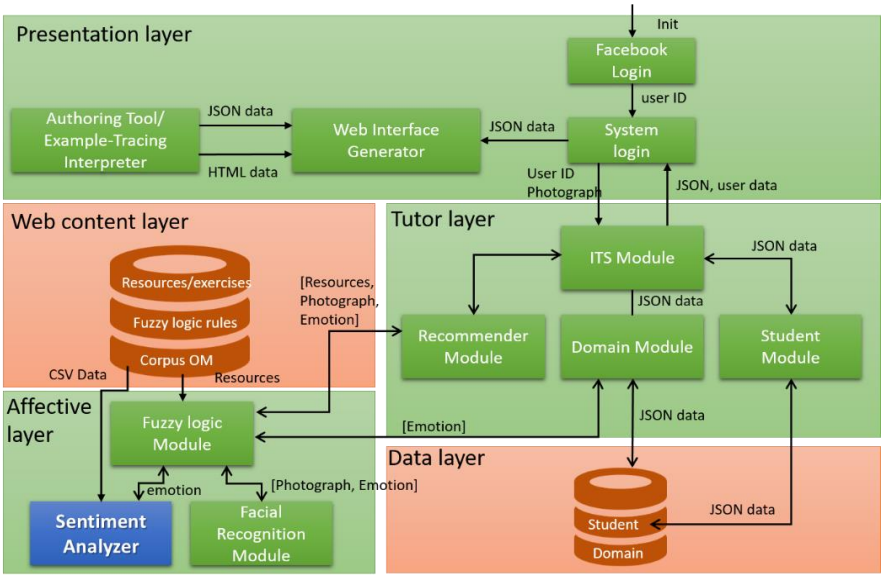


Fig. 4. Java Sensei Architecture with SA integration.

4 Discussion and Future Work

In this section, we present the results obtained from the classification of two datasets: one containing texts labeled with polarities and one containing text labeled with learning-centered emotions. There are a series of metrics that are used to estimate the quality of the classification algorithm that is part of our learning model. The simplest method to calculate the effectiveness of a classifier is the accuracy measure that calculates the average of text documents correctly classified over the total of documents to be classified. To obtain this value we applied a cross-validation technique with a dataset of 15400 texts in Spanish with 90% for the training data and 10% for the data test. Traditional ML classifier algorithms we used for the evaluation were Bernoulli Naïve Bayes, Multinomial Naïve Bayes, Support Vector Machine, Linear Support Vector Machine, Stochastic Gradient Descent, and K-Nearest Neighbors (KNN). For deep learning (see Fig. 3), we created a model based on combining CNN and LSTM neural networks. Table 1 shows the accuracy values obtained according to the size of the corpus. The traditional ML classifier with the highest score was Bernoulli Naïve Bayes, with an accuracy of 76.77%. In general, Bernoulli Naïve Bayes is a robust classifier against irrelevant features and it is suitable for text classifying. On the other hand, the deep learning model reached a precision of 88.26%, surpassing the other classifiers.

Table 1. Accuracy values obtained using different classifiers.

Classifier	Accuracy
Bernoulli NB	76.77%
Multinomial NB	75.31%
SVC	75.79%
Linear SVC	74.69%
SGDC Classifier	76.69%
KNN	68.46%
CNN + LSTM*	88.26%

* Deep learning model.

In a first approach, the system has been tested with 43 students in our research lab. Students must solve several exercises to learn and practice the topics of Java language. Each lesson has 15 exercises. At the end of each exercise, the system asks for student answers. Table 2, shows some examples of recollected answers.

Table 2. Collected opinions with polarity evaluation.

Opinion (Spanish)	Opinion (English)	Evaluation
No entendí este ejercicio	I did not understand this exercise	Negative
Me gusto este ejercicio	I liked this exercise	Positive
No tengo mucho conocimiento sobre este tema	I do not have much knowledge about this topic	Positive
El ejercicio es confuso	This exercise is confusing	Negative
Este ejercicio es sencillo de entender	This exercise is easy to understand	Positive

We recollected 178 student texts. The SA module evaluated 71 texts as positive and 107 as negative. We carried out another experiment with 53 computer-engineering students (45 men and 8 women) at the Instituto Tecnológico de Culiacán. We applied a learning system that covers the topics of Programming Fundamentals. After studying each sub-theme in text and visual or multimedia (video) format, each student registered her/his opinion/answer with a text not greater than 255 characters with respect to the content of the studied subject. Besides that, they labeled his opinion with an emoticon that represented his emotional state: bored, frustrated, neutral, excited or engaged. The total of recorded opinions was 851. Table 3 shows some examples of the opinions recorded by the students who participated in the experiment.

Table 3. Collected opinions with emotion evaluation.

Opinion (Spanish)	Opinion (English)	Evaluation
Me gusto bastante el vídeo.	I really liked the video.	Engaged
No me gusto que las voces fueran de España, además los gráficos de la animación están algo feos.	I did not like that the voices were from Spain, besides the graphics of the animation are somewhat ugly.	Frustrated
El video es bueno aunque creo que le falto profundizar más.	The video is good although I think it needs to go deeper.	Neutral
Vaya, es algo complejo.	Wow, it's something complex.	Bored
Quizás con un ejemplo quedaría más claro.	Perhaps with an example, it would be clearer.	Excited

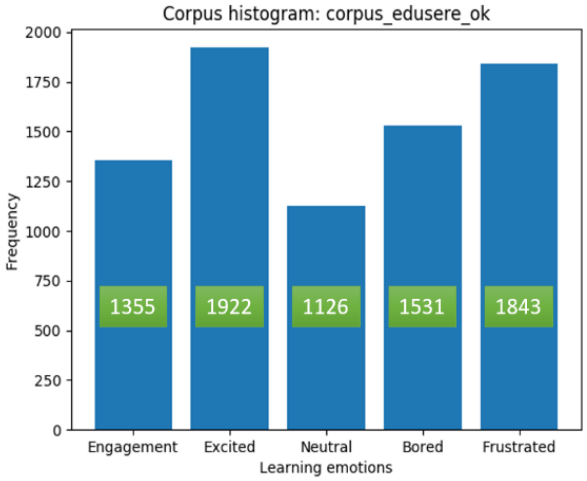


Fig. 5. Distribution of answers (opinions) in learning-centered emotions.

From 851 opinions, we found a tendency to neutral, excited, and engaged emotions, which indicates that new resources must be included that induce students to write

negative comments. Figure 5 shows the distribution of the opinions collected in the first experiment. We trained the previously used classifiers with this new corpus focused on learning, using cross-validation of 90% for training data, and 10% for test data. The algorithm with better precision was Linear SVC with 60%. The level of prediction is low because the size of the corpus is still small and it is not balanced as can be observed in Figure 5.

5 Conclusions and Future Work

Detecting emotions is a complex task even for human beings. Using only text or dialogs to detect emotions is a very hard challenge so that this research needs to be continued. This work presents the building of a sentiment analyzer for detecting emotions related to an educational context. The tests validate that this work had successful results in two type of database: one corpus with texts labeled with polarities and one corpus with texts labeled with learning-centered emotions. The results indicate that convolutional neural networks are a valuable choice in the recognition of polarities. We need more work with respect to classifying secondary emotions. The CNN architecture presented is a good beginning to recognize non-basic emotions and can be applied to recognize emotions from other different sources (e.g. facial expressions). In the future, we want to increase and balance the size of the two databases and we want to test them with more different machine learning methods to increase the number of classifiers. In addition, we will add different types of recognition from other sources such as voice, facial expressions, and brain signals.

References

1. Altrabsheh, N. *et al.*: SA-E: sentiment analysis for education. In: International Conference on Intelligent Decision Technologies. pp. 353–362 (2013)
2. Baker, R.Sj. *et al.*: Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive-affective states during interactions with three different computer-based learning environments. Int. J. Hum. Comput. Stud. 68, 4, 223–241 (2010)
3. Barrón-Estrada, M.L. *et al.*: Sentiment Analysis in an Affective Intelligent Tutoring System. In: Chang, M. *et al.* (eds.) 17th IEEE International Conference on Advanced Learning Technologies, (ICALT 2017). Timisoara, Romania, pp. 394–397 (2017)
4. Chaplot, D.S. *et al.*: Predicting Student Attrition in MOOCs using Sentiment Analysis and Neural Networks. In: AIED Workshops (2015)
5. Coen, M.H. *et al.*: Design principles for intelligent environments. In: AAAI/IAAI, pp. 547–554 (1998)
6. D'Mello, S., Graesser, A.: Dynamics of affective states during complex learning. Learn. Instr. 22, 2, pp. 145–157 (2012)
7. Dhanalakshmi, V. *et al.*: Opinion mining from student feedback data using supervised learning algorithms. In: Big Data and Smart City (ICBDSC), 2016 3rd MEC International Conference on. pp. 1–5 (2016)
8. El-Halees, A.: Mining opinions in user-generated contents to improve course evaluation. In: International Conference on Software Engineering and Computer Systems, pp. 107–115 (2011)
9. Kalchbrenner, N. *et al.*: A Convolutional Neural Network for Modelling Sentences. In: Proc. 52nd Annu. Meet. Assoc. Comput. Linguist. (Volume 1 Long Pap. pp. 655–665

(2014)
10. Kechaou, Z. *et al.*: Improving e-learning with sentiment analysis of users' opinions. In: Global Engineering Education Conference (EDUCON), 2011 IEEE. pp. 1032–1038 (2011)
11. Kim, Y.: Convolutional Neural Networks for Sentence Classification (2014)
12. Kumar, A., Jain, R.: Sentiment analysis and feedback evaluation. In: MOOCs, Innovation and Technology in Education (MITE), 2015 IEEE 3rd International Conference on. pp. 433–436 (2015)
13. Medhat, W. *et al.*: Sentiment analysis algorithms and applications: A survey. Ain Shams Eng. J. 5, 4, pp. 1093–1113 (2014)
14. Mikolov, T. *et al.*: word2vec. Google Sch (2014)
15. Nguyen, H., Nguyen, M.-L.: A Deep Neural Architecture for Sentence-level Sentiment Classification in Twitter Social Networking (2017)
16. Ortigosa, A. *et al.*: Sentiment analysis in Facebook and its application to e-learning. Comput. Human Behav. 31, pp. 527–541 (2014)
17. Pennington, J. *et al.*: Glove: Global vectors for word representation. In: Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP). pp. 1532–1543 (2014)
18. Qian, Q. *et al.*: Linguistically Regularized LSTMs for Sentiment Classification (2016)
19. Rani, S., Kumar, P.: A sentiment analysis system to improve teaching and learning. Computer (Long. Beach. Calif). 50, 5, pp. 36–43 (2017)
20. Wang, J. *et al.*: Dimensional Sentiment Analysis Using a Regional CNN-LSTM Model. In: Proc. 54th Annu. Meet. Assoc. Comput. Linguist. (Volume 2 Short Pap. January, pp. 225–230 (2016)
21. Warger, T., Dobbin, G.: Learning environments: where space, technology, and culture converge. EDUCAUSE Learning Initiative (2009)
22. Wen, M. *et al.*: Sentiment Analysis in MOOC Discussion Forums: What does it tell us? In: Proc. Educ. Data Min. (EDM), pp. 1–8 (2014)
23. Woolf, B.P.: Building Intelligent Interactive Tutors Student-centered strategies for revolutionizing e-learning. Morgan Kaufmann (2010)
24. Zatarain-Cabada, R. *et al.*: An affective and Web 3.0-based learning environment for a programming language. Telemat. Informatics. 35, 3, pp. 611–628 (2018)
25. Zhang, L., Liu, B.: Sentiment Analysis and Opinion Mining. In: Encycl. Mach. Learn. Data Min. May, pp. 1–10 (2016)

ISSN 1870-4069

Fuzzy System Inference and Fuzzy Cognitive Maps for a Cognitive Tutor of Algebra

Pedroza-Mendez Blanca-Estela[1], Reyes-García Carlos-Alberto[2], González-Calleros Juan Manuel[3], Guerrero-García Josefina[3]

[1,]Tecnológico Nacional de México, Instituto Tecnológico de Apizaco, Apizaco, Tlaxcala, Mexico
[2] Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOEP), Puebla, Mexico
[3] Benemérita Universidad Autónoma de Puebla, Puebla, Mexico
blanca.pm@apizaco.tecnm.mx, kargaxxi@inaoep.mx,
{jumagoca, joseguga}@gmail.com

Abstract. A cognitive tutor is an intelligent tutorial focused on supporting students on the resolution of problems. The large majority of the architectures for tutorial systems consider three common and indispensable modules: the student module; the tutor module; and, the domain module. These three modules are in charge of executing part of the more important activities within the tutor, and it is within these where it is necessary to implement models that simulate decision taking the way a human expert would. The student module should be constantly evaluating the progress made by students; the tutor module should provide adequate content, based on both the student's characteristics and the learning goals; and, the domain module should be capable of simulating the knowledge of experts regarding teaching-learning methodologies. Somehow, these three modules should interact in order to achieve the goals of the cognitive tutor. In this paper a proposal of fuzzy logic-based models for each of these modules is presented; the design of a system in which the three modules interact is also shown.

Keywords: fuzzy cognitive maps, cognitive tutor, problems resolution.

1 Introduction

Intelligent Tutoring System (ITS) is a system capable to guide students along a particular domain of knowledge through the solving of tasks tailored to the needs of the student [2]. In turn, a cognitive tutor is a type of ITS with a long-time proven efficacy. Its efficacy is based on its capacity to provide individualized support for the learning of complex cognitive abilities through the practice of problem solving [8]. According to Gonzalez [2], the key components of traditional ITS can be organized in different modules as a student model, domain model, tutor model and interface or communication module, which interact with the user. In this paper, the interest is focused on the three first modules, which contain the representation of expert knowledge in areas related to evaluation processes, teaching and learning methodologies and the detection of cognitive skills of students. These activities have

the purpose of guaranteeing that the implemented educational designs are adequate for learning achievement. These designs are teaching processes whose implementation isn't easy without help from automatized tools such as intelligent tutoring systems. For this research in particular we are focusing on systems for algebra teaching through problems solving, using fuzzy cognitive maps, for represent the mental schemas that model dependencies between key concepts of domain, in this case algebra, and a fuzzy model for the detection of students' expertise level.

Since the student model module requires the implementation of strategies for the detection of students' cognitive skills for problem resolution in the chosen knowledge field (algebra), it was necessary, first, to use an instrument for the constant measurement of such skills. The instrument is based on the one proposed in [11], known as the 3UV (3 uses of the variables) model. This instrument was adapted in order to associate it with Bloom's Taxonomy, which is a cognitive method for educational objectives whose educational goals were classified as six cognitive levels: remember, understand, apply, analyze, evaluate and create [9]. The instrument based on 3UV model, are the rubric of input to the fuzzy inference model (equation 8). The goal of the inference system is to classify the student with values that are not very rigid and that may provide information that is closer to reality than a rigid model would.

However, in order to develop more efficient tools, not only should knowledge representation be considered as a base. One should go in depth into something more complex known as the representation of dependency between key knowledge concepts in a given domain. This is supported by different pedagogical theories such as the Cognitive Load Theory, which computationally can be modeled through Fuzzy Cognitive Maps (FCM). A FCM helps in the decision making process and allow us to represent such dependencies. This tool is in an initial testing phase with students. Therefore, the results presented in this paper are experimental and are related, execution-wise, to map efficiency. For the tutor module, a fuzzy inference model was also considered; this model allows, based on the results of the student and domain models, the determination of the kind of problems on which the student should conduct further work. For its part, the Cognitive Load Theory (CLT), is based on the assumption that the construction and automatization of cognitive schemas for learning are the main goals of teaching. But those objectives may be thwarted by the limited capacity of working memory. Due to this factor, the proper allocation of available cognitive resources is essential for the learning process [4]. Here lies the importance of the representation of the dependency between the key concepts in the knowledge domain to be learned, and of not only using a general representation of knowledge. In other words, FCM must represent how the knowledge of a domain concept of the teaching material, may be affected by the knowledge of another domain concept [1].

2 Basic Concepts

2.1 Modeling Fuzzy Cognitive Maps

Fuzzy Cognitive Map (FCM) is fuzzy-graph structure for representing causal reasoning, analyze inference patterns and they act as a nonlinear dynamical system [5]. In education, the causality characteristic allows the FCMs to be adequate to represent

the dependence between key concepts of some domain of the knowledge in question, allowing to detect the learning material that should be delivered, to some student, with respect to their knowledge level and personal needs [1,10].

For FCM reasoning process, a simple mathematical formulation is usually used. A model implication converges to a global stability, equilibrium in the state of the system. During the inference process, the sequence of patterns reveals the inference model. The mathematical representation of FCMs has the following form [3,5,6]:

$$A_i(k+1) = f\left(A_i(k) + \sum_{j=1}^{N} A_i(k) \cdot e_{ji}\right), \quad (1)$$

where f (·) is a threshold (activation) function. Sigmoid threshold function gives values of concepts in the range [0, 1] and its mathematical type is:

$$f(x) = \frac{1}{1+e^{-m \cdot x}}, \quad (2)$$

where *m* is a real positive number and x is the value $A^{(k)}_i$ on the equilibrium point.

2.2 Fuzzy Cognitive Maps in Education

However, when working with the causal dependency between the knowledge levels that a student possesses on a particular domain subject, a situation may arise in which the increase in the concept of any of the nodes of the fuzzy cognitive map may not be total, as represented in equations 1 and 2. Because of this, in these cases we should use a model that helps determine the causality between the nodes based on the dependency of the knowledge domain of the learning material. Chrysafiady et al. [1] define such model as a tuple (C, W, KL, f), where:

1. $C= \{C_1, C_2, .. . C_n\}$ is the set of concepts of the domain knowledge.
2. $W: (C_i, C_j) \rightarrow w_{ij}$ is a connection matrix, where w_{ij} is a weight of the directed ard from C_i to C_j, which denotes that the knowledge level of the concept C_i affects that of concept C_j.
3. KL is a function that at each concept C_i associates the sequence of its activation degree. In other worlds, $KL_i(t)$ indicates the value of a concept's knowledge level at the moment t.
4. f is a transformation function. For the definition of the transformation function the following limitation has to be taken into account. Only the knowledge level of the most recently read concept affects the knowledge level of a domain concept, each time. Consequently, the KL value of a concept is affected only by the KL value of the most recently read concept, regarding the weight of the directed arc that connects them. Therefore, the transformation function for a FCM, which is used to represent the domain knowledge of the learning material, is defined as:

$$KL_i(t+1)=f(KL_i(t)\pm w_{ji} * p_j * KL_i(t)/100), \quad (5)$$

where p_j is the percentage of the difference on the value of the knowledge level of the most recently read concept C_j, with $p_i = (KL_j(t + 1)\text{-}KL_j(t))*100/KL_j(t)$. Also, the + is used in case of increase and the – is used in case of decrease.

2.3 Analytical Model for evaluation of learning

The areas and Cognitive ability is closely related to learning acquisition and is the basis for developing instructional strategies, and their preferences are the basis for developing the type of mentoring. Rongmei proposes an analytical model, combined with fuzzy logic to somehow categorize students based on their cognitive ability and propose a model that is able to make the appropriate instructional design decision to the student, based on their progress. The evaluation model can be expressed via a triple, such as formula [7].

$$M= (U, V, A), \tag{6}$$

where, U = (u1, u2, u3, u4, u5, u6), are the weight of each of the six points of knowledge of Bloom´s Taxonomy, which is given by the experts. V= (v1, v2, v3, v4, v5), these elements respectively stand for five reviews: excellent, good, medium, passed and fail, which is based on both the results of cognitive ability synthetic evaluation and the test.

A is a matrix where each line has the following form:

$$A_i = (a_1, a_2, a_3, a_4, a_5, a_6), \tag{7}$$

where a_i are values between 0 and 1, and are define as:

$$a_i = \frac{r_{ij}(1)}{r_{ij}(1) + r_{ij}(0) + r_{ij}(-1)}, \tag{8}$$

$r_{ij}(1)$ is defined as the number of correct answers of the student in each of the cognitive abilities, r_{ij} (0) is the number of unanswered questions and r_{ij} (-1) is the number of incorrect answers in each of the cognitive skills.

Finally, M is defined as:

$$M = \sum_{i=1}^{6} g_i * u_i, \tag{9}$$

where

$$G = W \cdot A = (g_1, g_2, g_3, g_4, g_5, g_6), \tag{10}$$

and W defines the weight of each rubric, this is:

$$W = (w_1, w_2, w_3, w_4, w_5, \ldots w_n). \tag{11}$$

3 Description of the Proposed Models

Concretely, the designed domain module for the cognitive tutor is in charge of analyzing the dependencies between the topics that are considered as the main goal in the learning-teaching process, algebra, for this case. For practical purposes, the syllabus for the subject of algebra of a public, higher secondary level educational system in Mexico was chosen. This syllabus is divided into three units; for the modeling of dependencies between the topics, a fuzzy cognitive map for each unit was built. To obtain the values for the edges of each FCM, a survey with 18 higher secondary level math teachers was used. In Figure 3 the FCM of first unit is showed.

Generally, the processes of the three models in which this work is focused can be seen as the iterative process shown in Figure 1. The first process is the one related with the student model module; process 2 is the domain module; and, process 3 is the one that executes the activities of the tutor model module. The iterative process starts with the handling of rubrics based on the 3UV model, which are evaluated through the analytical model described in the section 2.3 so, using this model, the input parameters for the first FIS are obtained. The model for process 1 is a Fuzzy Inference System (FIS) type Mamdami whose parameters are shown in Figure 2, and it uses as input parameters the levels of Bloom's Taxonomy, which work as the fuzzy sets for the FIS. The fuzzy value obtained of process 1 is used how input to the process 2.

The second process is a model of fuzzy cognitive maps, whose goal is to identify the relationships between the concepts of the knowledge domain (algebra) in order to determine its relationship with the students' cognitive skill and, based on these relationships, obtain the input parameters for the second FIS, which determines the problems that should be presented to the student in order to help him(her) improve his(her) problem solving skills. For the iterative process for each FCM and for getting the convergence parameters, the models of sections 2.1 and 2.2 are applied. The relationship between these three processes gives way to an iterative process that constantly relates the results of each of the former ones; so in process 1, a condition is established that gives guidelines to the execution of the three processes, as long as the cognitive evaluation doesn't reach adequate values for each of the three defined fuzzy cognitive maps.

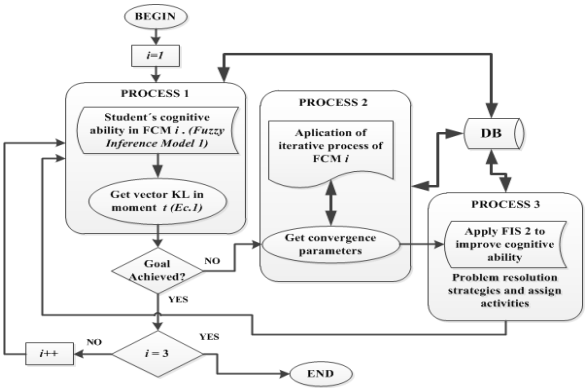


Fig. 1. Iterative process for the three modules of cognitive tutor.

TYPE OF VARIABLE	LINGUISTIC VARIABLE	RANK	LINGUISTIC VALUES	MEMBERSHIP FUNCTION	PARAMETERS
INPUT VARIABLES	Remember	[0,1]	High	Trapezoidal	[0.6627 0.782 1.1 1.9]
			Low		[-0.9 -0.1 0.5542 0.771]
	Understand	[0,1]	High	Trapezoidal	[0.6071 0.726 1.01 1.01]
			Low		[0.00789 0.328 0.541 0.649]
	Apply	[0,1]	High	Trapezoidal	[0.586 0.901 1.007 1.02]
			Low		[-0.36 -0.04 0.446 0.7288]
	Analyze	[0,1]	High	Trapezoidal	[0.5939 0.845 1.01 1.01
			Low		[-0.36 -0.04 0.522 0.6521]
	Evaluate	[0,1]	High	Trapezoidal	[0.5066 0.784 1.01 1.02]
			Low		[-0.34 0.137 0.3876 0.652]
	Create	[0,1]	High	Trapezoidal	[0.496 0.832 1.01 1.01]
			Low		[-0.357 -0.0374 0.422 0.544]
OUTPUT VARIABLE	COGNITIVE LEVEL	[0,1]	Excellent	Trapezoidal	[0.901 0.975 1.02 1.23]
			Medium		[0.671 0.725 0.874 0.922]
			Low		[0.409 0.53 0.612 0.725]
			Passable		[0.025 0.225 0.275 0.475]
			Failed		[-0.225 -0.025 0.025 0.225]

Fig. 2. Parameters of first Fuzzy Inference System of figure 1.

4 Results

The processes of figure 1 were simulated in a program with graphical user interfaces of matlab, version 2017b. In principle, for the purpose of testing, random values were considered for two rubrics with three problems, of each topic of algebra. Subsequently, process 1 is applied to obtain the values of Bloom's Taxonomy and the fuzzy value. The Figure 3 show the results of file in which, the process of simulate the assessment of 5 topics of unit 1, when the counter "*i*" of the figure 1 is equal to 1, is accomplished. Figure 4 shows the results of the file that is responsible for performing the iterative process of the first diffuse cognitive map, which shows the values that the student must reach to get to understand all the topics of unit 1. Finally, Figure 5 shows the process related to decision making related to the type of problems that the student must solve, according to the values obtained in the previous processes.

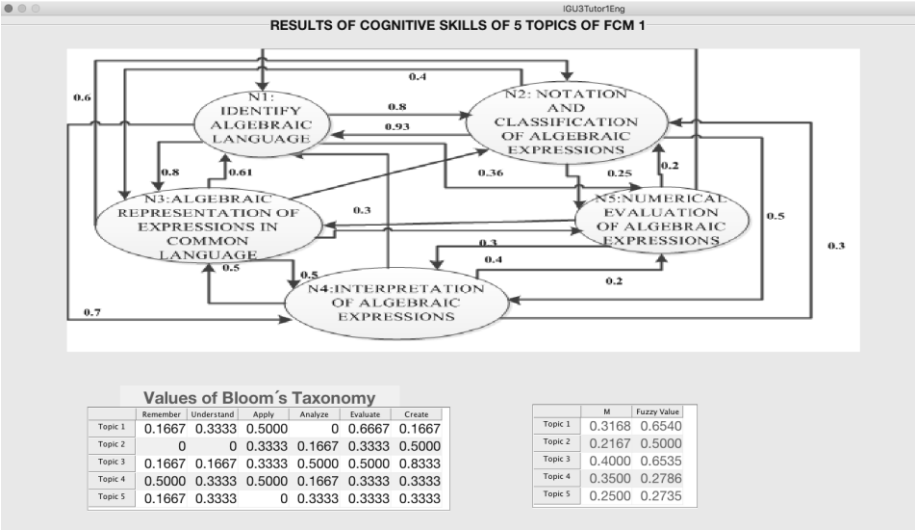


Fig. 3. Evaluation of the 5 topics of the FCM of unit 1.

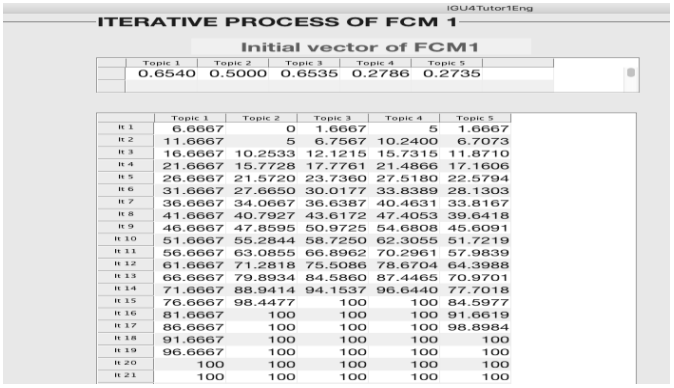


Fig. 4. Results of the iterative process of FCM 1.

4.1 Analysis of Results

In Figure 3, fuzzy values and average values are show, the difference among these values is that, the fuzzy value is representing of best manner, the characteristics of the student, in relation with your performance in each level of the Bloom´s taxonomy. By example for topic 1, the average value is very low, even though the student performs well at the "evaluate" level of the Blooms taxonomy, which is why the fuzzy value is high. The results of Figure 5 are the suggestions that the tutor module will make to the student, which depend on the values of the Bloom´s taxonomy that the student obtains and the dependencies between topics, therefore, for the topic 5 in figure 5, the suggestion is that the student must work with fully resolved examples because their cognitive level is low.

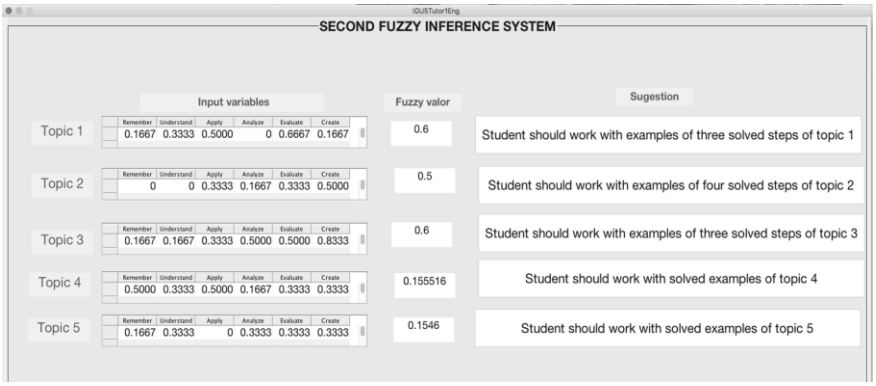


Fig 5. Application of the second Fuzzy Inference System for the tutor module.

5 Conclusions

The dependency between several algebra topics, as well as the relationship with problem solving, was modeled through of fuzzy models. It can be observed from the

results shown that applying a variant of the original model proposed by Kosko, a model that converges into an optimal vector, using several incremental factors for each of the concepts related to the different subjects of algebra is obtained. This allows us to simulate the dynamic behavior of the students' learning process, and to determine the kind of problems that should be offered to them as exercises to enhance their learning.

The execution of the cognitive map convergence process has only been conducted to analyze the diverse combinations of input parameters. However, the goal is to obtain real data from several students that may serve as an input vector for map modeling and, thus, obtain the parameters that will help in the training of the model that will assign the problems and topics in which the student should work in. All this, in order to achieve the final objective, which is that the student can to solve completely and autonomously every of the problems in each topic, which will occur when the vectors associated to evaluation contain values that represent a total comprehension of every one of the concepts. This will be the vector to which all fuzzy cognitive maps should converge. This way, the efficiency of the domain module could be evaluated.

References

1. Chrysafiadi, K., Virvou, M.: A knowledge representation approach using fuzzy cognitive maps for better navigation support in an adaptive learning system. SpringerPlus 2(1), p. 81 (2013)
2. González, C., Mora, A., Toledo, P.: Gamification in intelligent tutoring systems. In: Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality - TEEM '14, pp. 221–225 (2014)
3. Groumpos, P.P.: Fuzzy Cognitive Maps: Basic theories and their application to complex systems. Fuzzy Cognitive Maps, 247, pp. 1–22 (2010)
4. Kalyuga, S.: Cognitive Load Theory: How Many Types of Load Does It Really Need? Educational Psychology Review 23(1), pp.1–19 (2011)
5. Kosko, B.: Fuzzy cognitive maps. International Journal of Man-Machine Studies 24(1), pp. 65–75 (1986a)
6. Papageorgiou, E.I., Salmeron, J.L.: Methods and Algorithms for Fuzzy Cognitive Map-based Modeling. In: Papageorgiou, E.I. (ed.) Fuzzy Cognitive Maps for Applied Sciences and Engineering, pp. 1–28 (2014)
7. Rongmei, Z., Lingling, L.: Research on Internet Intelligent Tutoring System Based on MAS and CBR. In: 2009 International Forum on Information Technology and Applications, pp. 681–684 (2009)
8. Salden, R.J.C.M., Aleven, V., *et al.*: The expertise reversal effect and worked examples in tutored problem solving. Instructional Science 38(3), pp.289–307 (2010)
9. Skiba, D.J.: Bloom's digital taxonomy and word clouds. Nurs. Educ. Perspect. 34(4), pp. 277–280 (2013)
10. Sweta, S.. Lal, K.: Personalized Adaptive Learner Model in E-Learning System Using FCM and Fuzzy Inference System. International Journal of Fuzzy Systems (2017)
11. Ursini, S., Trigueros, M.: How do High School Students Interpret Parameters in Algebra? pp. 361–368 (2004)

ISSN 1870-4069

Methodology for Automatic Identification of Emotions in Learning Environments

Yesenia N. González Meneses[1], Josefina Guerrero García[1], Carlos Alberto Reyes García[2], Iván Olmos Pineda[1], Juan Manuel González Calleros[1]

[1] BUAP, Facultad de Computación, Puebla, Mexico
[2] INAOE, Ciencias Computacionales, Puebla, Mexico
{yeseniaglez0, joseguga0, ivanoprkl, jumagoca78}@gmail.com, kargaxxi@inaoep.mx

Abstract. This paper presents a methodological proposal for automatic identification of emotions in educational environments using machine learning algorithms and physiological and behavioral signal acquisition technologies to identify relations between emotions and learning. Four of the main learning-centered emotions are considered [1]: engagement, boredom, confusion and frustration. It is proposed to make a fusion of data from two physiological and behavioral signal acquisition technologies with the objective of achieving the identification of emotions in the most precise manner. Therefore, considering the stages of the proposed methodology, the first of them is presented and the design of the experiment that will be executed for data collection. The development of an appropriate database with elements belonging to a learning environment for the study of emotions is an essential task.

Keywords: automatic identification of emotions, learning-centered emotions, learning environment, physiological signals, machine learning.

1 Introduction

From a computational approach, the detection and identification of emotions is a relevant problem within the area of affective computation research, since it is the starting point in the study and development of human-machine interaction systems sensitive to the emotions of human beings. In the human-machine interaction area, this is known as classification of emotions, automatic detection or identification of emotions, as it is mentioned in [2] and [3]. This problem is supported by computational models of emotion. In the case of the psychological research that uses computational models, the emphasis will largely be on fidelity with respect to human emotion processes. In AI (Artificial Intelligent) and Robotics related works, evaluation often emphasizes on how the modeling of emotion impacts reasoning processes or leads in some way to improved performances. In Human-Computer Interaction (HCI) work, the key evaluation is whether the model improves human computer interaction such as making it more effective, efficient or pleasant.

To appreciate the transformative role that computational models of emotion can have on research, in [2] they consider three aspects: the impact on emotion research in

psychology, the impact on artificial intelligence (AI) and finally the impact on work in human-computer interaction. Thus, the complexity of the problems that affective computation deals with lies in the fact it is an inter-disciplinary area that encompasses computer science, psychology, and cognitive science [4]. Particularly, in the computational area the challenge is selecting and testing machine learning algorithms that could be integrated in a complete model of recognition of learning-centered emotions.

In this sense, two approaches have been used to recognize human emotions [5]: The objective, using sensors or capturing images; as in the recognition of facial expressions, voice, heart rate, body language, body's thermal activity, muscular activity and brain waves. And the subjective, such as contextual analysis through direct observation, surveys, or interrogations of the individuals themselves.

The automatic recognition of emotions and the human-machine interaction can be considered, then, from the use of physiological signal sensors that allow the acquisition of data through voice analysis, video images of the face, eyes, head, or body movements of the people [5].

Within the human-machine interfaces that capture physiological signals that can help in the recognition of emotions, as mentioned in [6], the brain wave diadems that send information to an electronic device in the form of an electroencephalogram (EEG) are found.

There are also cardiovascular wristbands, which measure heart rate and provide information in the form of an electrocardiogram (ECG); electrodermal activity sensors that measure the level of conductivity of the skin through sweat on the hands; thermal cameras that allow the measurement of the temperature change of the human body associated with the different emotional states.

There are also devices that measure muscle electrical activity in response to a nerve stimulation of a muscle, in the form of an electromyography (EMG). Regarding devices related to the identification of people's behavior -such as body postures and gestures- there are traditional video cameras, webcams or augmented reality that allow the recording of facial expressions and body movements, as well as eye tracking, important to the recognition of emotions. In this type of devices, there are also voice recorders, another medium used to identify emotions, as mentioned in [7].

The data obtained from different devices must be processed and classified based on a specific objective. In this case, they will be used for the recognition of learning-centered emotions and captured in real educational environments when students are executing a learning activity. Therefore, it is expected to contribute to physiological signal processing technics, from their capture, preprocessing, selection of characteristics and, finally, with the identification of learning-centered emotions.

The objective of identifying emotions during educational activities is to corroborate the relationship between learning-centered emotions and the level of learning obtained by students. This relationship is the base for approaching educational strategies that help to improve the levels of learning and, therefore, the educational level in our country.

2 Motivation

The problem involving automatic recognition of emotions has been an area of investigation highly active in the last years. Regardless of this, a clear solution, which is within reach of the most people, is still far away. Several drawbacks have influenced the construction of an appropriate solution from a computational point of view. On one hand, a factor that affects the performance of emotion recognizers in real contexts is the difficulty to generate databases with spontaneous emotions. Generally, works are made with actuated databases which provide portraits of emotions representing prototypical and intense emotions that facilitate the search of correlations and the subsequent automatic classification. This kind of databases are usually captured in a controlled environment, which decreases problems in the processing of information (noise, for example).

In addition, it can guarantee a balanced number of samples per class. As consequence, there have not been good results when translating the knowledge extracted from these databases to real contexts [8]. In contrast, the data bases with spontaneous records show information with emotional content that does not belong to a single class, but a mixture of them. In other cases, there are samples with a very light emotional charge, close to a neutral emotional state. In addition, databases with spontaneous emotions are usually recorded in noisy environments, such as classrooms, study rooms, entertainment areas, offices, factories or in phone conversations, which leads to the inclusion of noise.

Finally, because of the very nature of the problem, it is not possible to ensure a balanced quantity of examples per class. Another challenge to be solved is the extraction and selection of a set of characteristics that allow recognizing emotions in the data captured spontaneously. Although progress in the area has been important, there is still much to be done in realistic contexts. Therefore, it is necessary to propose and explore other approaches that allow reaching a good performance of the recognition of emotions in real world applications. An evident aspect to consider is the fact that the area of application has an important influence on the accuracy of emotion recognition [7], as well as the degree of intrusion of the tools used.

2.1 Problem Statement

Considering the presented problems and the analysis of recent related works, we can identify that in learning environments there is a lack of adequate methodologies to recognize the emotional state of students during learning processes through interaction with a computer. Preferably, that integrates diversified characteristics, obtained even from the fusion of data coming from the use of different technologies of acquisition of physiological and behavioral signals based on a model of emotions that allows to become closer to reality and, overall, to the cognitive learning process, which contributes to the analysis of the emotion-learning relation.

2.2 Research Objective

The objective of this work is to recognize emotions in educational environments using machine learning algorithms and technologies of acquisition of physiological and

behavioral signals to identify emotion-learning relationships. To achieve the above, the following activities are proposed:

- Investigate, analyze and select the model of emotions and technologies for the acquisition of physiological data on which the proposal will be based.
- Analyze and select educational environments to make the physiological and behavioral data capture.
- Investigate, study and select machine learning algorithms for the selection of characteristics for the classification of emotions.
- Design a methodology for the identification of learning-centered emotions.

To validate the work done, tests will be carried out with college students and the recognition accuracy will be evaluated with metrics that allow comparing it with related works in the literature. With this we pretend to corroborate how the proposed methodology allows identifying emotions in educational contexts that can contribute to decision making in emotion-learning relationships.

3 State of the Art

The analysis of the state of the art that is being carried out consists of a collection of articles focused specifically on the recognition of emotions in learning activities and the identification of the emotion-learning relationship. These investigations in automatic recognition of emotions represent approximately 20%, and the remaining proportion (approximately 80%) are works that recognize basic emotions (happiness, sadness, fear, anger, contempt, disgust and surprise). There is a clear difference in the literature found. There are works that, although they make an identification of emotions, give greater relevance to the analysis of the emotion-learning relationship. Others emphasize especially the algorithms for automatic recognition of emotions and the recognition rate reached and, with less importance, analyze the emotion-learning relationship. Considering these two aspects, a review of the state of the art analyzed so far is presented.

3.1 Works Focused on the Computational Problem of Automatic Recognition of Learning-centered Emotions

In the literature reviewed, these works represent approximately 63% of the total of papers reviewed that identify learning-centered emotions. Only the most recent research is mentioned here.

In the work of [10] they use a convolutional neuronal network for the recognition of learning-centered emotions. They run tests using three data-bases: RaFD, database of posed facial expressions containing images of 8 basic emotions and two spontaneous databases created by them, especially with content related to learning-centered emotions. The emotions they recognize are: engagement, excitement, boredom and relaxation.

In [11] they propose a computer-assisted method for special school instructors, where they teach students with mental disorders or emotional problems using a system that employs wearable sensor technologies and intelligent recognition of

emotions. The emotion recognition module starts with the capture of the signal from brain diadems. The filtered data is sent to the characteristic extraction module. After extracting the characteristics, they are processed with two classifiers: support vector machine and near k-neighbors with a cross-validation of 10 iterations. The emotions they recognize are: happiness, calm, sadness and fear.

In [12] they implement a binary local pattern for the recognition of learning-centered emotions. The purpose of the work is to build a database of spontaneous facial expressions corresponding to affective states in learning environments to be used in different intelligent tutorial systems. The data capture technologies they use are video and EEG diadems (Emotiv-EPOC). The learning-centered emotions that they recognize are: frustration, boredom, engagement and excitement. They use a support vectors machine to do the classification of emotions

The work of [2] explains the construction and validation of a database of facial expressions that they collect by taking pictures with a webcam. Each photograph is labeled with the emotions of the users obtained at that moment from the Emotiv-EPOC device. For the recognition of facial expressions, they use a technique based on geometry. These measurements are transformed into characteristics to train a support vector machine. They obtain accuracy, per emotion, of: boredom of 64%, engagement of 64%, excitement of 83% and frustration of 62%. The same model for the recognition of emotions is also used in [13], as part of an affective learning environment based on Web 3.0 to learn how to program in Java.

Finally, in [14] they propose different recommended activities to induce a certain mental state and capture the EEG response for each one of them. They intend to identify the ideal emotional state to learn. These activities are based on psychological research dedicated to measuring the level of attention, concentration and other functions. In the classification process, they use k-means and clustering.

Of the totality of works reviewed we can conclude that the most popular devices for the capture of data are webcams, followed by EEG Emotiv diadems. From the use of these technologies and other less popular ones, 80% of the works create their own databases to train and assess their recognizers. The most used algorithms for classification tasks are artificial neural networks, support vector machine (SVM), clustering techniques, Bayesian classifiers and nearer neighbors (KNN), among others. Considering this analysis, we will begin the processing of the data collected using the aforementioned algorithms on the signals that we capture.

3.2 Works Focused on the Analysis of the Emotion-learning Relationship based on the Automatic Recognition of Learning-centered Emotions

Up next, the most recent research on the analysis of the emotion-learning relationship using computer techniques for the identification of emotions is listed. These works represent 37% of the literature reviewed so far.

In [15], they make a study of the affective states that originate when students learn with technology, using a tutorial to learn the basics of programming in Python. The emotions they can identify are: engagement, confusion, frustration, boredom, curiosity (which were the most frequent affective states), anxiety, happiness, anguish, surprise, disgust, sadness and fear (which were the rarest). With the analysis of their results, they identify the emotion-learning relationship from five different approaches.

In [16] they present the development of an intelligent tutor with recognition and management of emotions for mathematics. And in [17] they develop an affective learning system for algorithmic logic by applying gamification. In both works, a module for the recognition of emotions is integrated. This process is based on the analysis of faces.

The analysis of these works helped us to identify the areas of opportunity regarding our research and to be able to define the design strategies of the methodology and algorithms for the selection of characteristics and classification to be used. The above, to validate the hypothesis that states that the proposed methodology allows us to identify emotions in educational contexts that contribute to the decision-making in the emotion-learning relationship.

4 Proposed Methodology

General methodology of the investigation is shown in Fig.1.

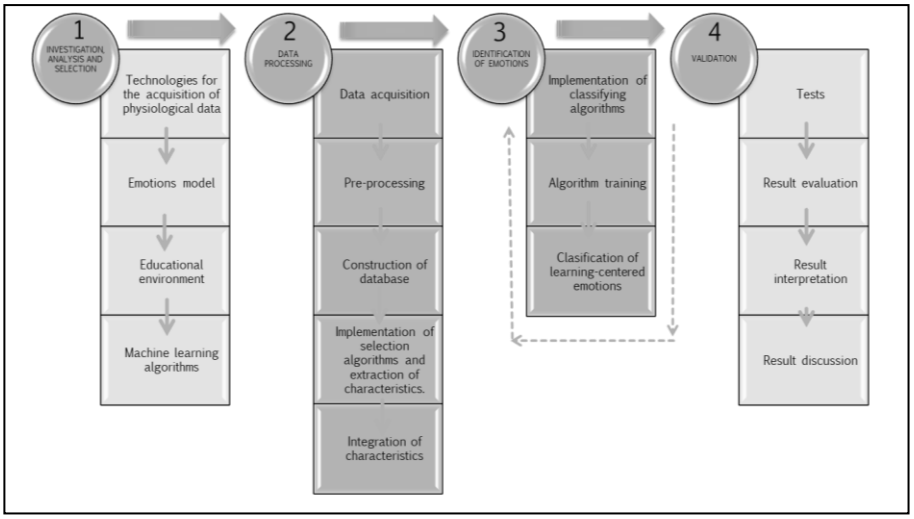


Fig. 1. Proposed Methodology.

The main stages of this methodology are:

1. Research, analysis and selection of technologies for acquisition of physiological data, emotions model, application context and machine learning algorithms, which will be proposed for the development of the methodology.
2. Data processing; it includes the recording and acquisition of physiological and behavioral data to form the data base, the preprocessing required to prepare the data and the implementation of extraction, selection and integration algorithms of relevant characteristics derived from the different signals. We will create a database of student features obtained in real time while doing a learning activity. Which will be formed by two physiological characteristics: the temperature of

the face and the heart rate; and two behavioral characteristics: the facial image and the movements of the superior limbs.

3. Identification of emotions; the selected machine learning algorithms will be tested and trained for classifying learning-centered emotions. We will start with neural networks adjusted to the characteristic vectors formed by the aforementioned data, trying to recognize emotions from a discrete approach. We will also test fuzzy classification algorithms with the objective of recognizing emotions from a continuous focus. In both cases, with the purpose of improving recognition accuracy with respect to other similar jobs in the state of the art.
4. Validation; tests will be carried out to evaluate the complete methodology for the identification of emotions with metrics that measure their execution, precision and accuracy.

In the stages of data processing and identification of emotions we will follow the steps of the KDD process (Knowledge Discovery in Databases), which correspond to each of the activities considered in the stages of the methodology.

5 Discussion

For the development of the proposed methodology, the use of machine learning algorithms is proposed, starting with neural networks and the use of fuzzy techniques. A preliminary analysis of the captured data will allow us to choose the best algorithms according to their characteristics and distribution. A contribution is expected with the proposal of integrating data from different physiological and behavioral signal acquisition technologies.

The data acquisition stage will be carried out through a controlled experiment, in which college students will participate using a tutorial to learn algebra. Data of 100 students will be collected. It is expected to obtain results that contribute to the improvement of the identification of learning-centered emotions with better precision and that can be considered for the decision-making process on pedagogical strategies and learning activities more adapted to each student.

6 Conclusion

The development of this project will contribute to the identification of gaps in the automatic recognition of learning-centered emotions. It will allow identifying and proposing less uncomfortable technologies for students, recommended for the collection of data in the identification of learning-centered emotions; we will start with the integration of video cameras, thermal cameras, heart rate sensor and Kinect for Windows.

The implementation of the process of automatic identification of emotions can be integrated into learning environments within the student modeling module or within another type of environment, which may consider the emotional state to provide feedback to the students' learning.

References

1. D'Mello, S., Graesser, A.: Dynamics of affective states during complex learning. Learn. Instr. 22(2), pp. 145–157 (2012)
2. Armony, J.: Computational models of emotion. In: Proceedings. 2005 IEEE Int. Jt. Conf. Neural Networks, vol. 3, pp. 1598–1602 (2005)
3. Picard, R.W.: Affective Computing (1997)
4. Fuentes, C., Herskovic, V., Rodríguez, I., Gerea, C., Marques, M., Rossel, P.O.: A systematic literature review about technologies for self-reporting emotional information. J. Ambient Intell. Humaniz. Comput., pp. 1–14 (2016)
5. Zatarain, R., Barrón, M.L., Luis, O.J., Martínez, J.A.: Reconocimiento automático y aspectos éticos de emociones para aplicaciones educativas. Komputer Sapines, México, pp. 27–31 (2014)
6. Cowie, R. *et al.*: Emotion recognition in human computer interaction. IEEE Signal Process. Mag. 18(1), pp. 32–80 (2001)
7. Steidl, S.: Automatic Classification of Emotion-Related User States in Spontaneous Children's Speech. Universität Erlangen-Nürnberg (2009)
8. González-Hernández, F., Zatarain-Cabada, R., Barrón-Estrada, M.L., Rodríguez-Rangel, H.: Recognition of learning-centered emotions using a convolutional neural network. J. Intell. Fuzzy Syst. (2017)
9. Mehmood, R., Lee, H.; Towards Building a Computer Aided Education System for Special Students Using Wearable Sensor Technologies. Sensors 17(317), pp. 1–22 (2017)
10. Zatarain-Cabada, R., Barrón-Estrada, M.L., González-Hernández, F., Oramas-Bustillos, R., Alor-Hernández, G., Reyes-García, C.A.: Building a Corpus and a Local Binary Pattern Recognizer for Learning-Centered Emotions. Adv. Comput. Intell., vol. II, pp. 524–535 (2017)
11. Zatarain-Cabada, R., Barron-Estrada, M.L., González-Hernández, F., Rodríguez-Rangel, H.: Building a Face Expression Recognizer and a Face Expression Database for an Intelligent Tutoring System. In: Proc. - IEEE 17th Int. Conf. Adv. Learn. Technol. ICALT 2017, no. 2161–377X/17, pp. 391–393 (2017)
12. Zataraín, R., Barrón, M.L., González, F., Reyes-García, C.A.: An Affective and Web 3.0 Based Learning Environment for a Programming Language. Telemat. Informatics (2017)
13. Arana-Llanes, J.Y, González-Serna, G., Pineda-Tapia, R., Olivares-Peregrino, V., Ricarte-Trives, J.J., Latorre-Postigo, J.M.: EEG lecture on recommended activities for the induction of attention and concentration men tal states on e-learning students. J. Intell. Fuzzy Syst. (2017)
14. Bosch N., D'Mello, S.: The Affective Experience of Novice Computer Programmers. Int. J. Artif. Intell. Educ. 27(1), pp. 181–206 (2015)
15. Barrón, M.L., Zatarain, R., Hernández, Y.: Intelligent Tutor with Emotion Recognition and Student Emotion Management for Math Performance. Rev. Electrónica Investig. Educ., vol. 16, pp. 88–102 (2014)
16. Zatarain-Cabada, R., Barrón-Estrada, M.L., Ríos-Félix, J.M.: Affective Learning System for Algorithmic Logic Applying Gamification. Chapter Lect. Notes Comput. Sci. Oct. 2017, vol. 576, no. August (2017)
17. Fayyd, U.M., Shapiro, G.P., Padhraic, S.: From data mining to knowledge discovery: An overview. Adv. Knowl. Discov. Data Mining, AAAI Press / MIT Press (1996)

ISSN 1870-4069

Study of Spontaneous and Acted Learn-Related Emotions Through Facial Expressions and Galvanic Skin Response

Andres Mitre-Ortiz, Hugo Mitre-Hernandez

Center for Research in Mathematics,
Human-Centered Computing Lab,
Quantum: Knowledge City, Zacatecas, Mexico
{andres.mitre, hmitre}@cimat.mx

Abstract. In learning environments emotions can activate or deactivate the learning process. Boredom, stress and happy –learn-related emotions– are included in physiological signals datasets, but not in Facial Expression Recognition (FER) datasets. In addition to this, Galvanic Skin Response (GSR) signal is the most representative data for emotions classification. This paper presents a technique to generate a dataset of facial expressions and physiological signals of spontaneous and acted learn-related emotions –boredom, stress, happy and neutral state– presented during video stimuli and face acting. We conducted an experiment with 22 participants (Mexicans); a dataset of 1,840 facial expressions images and 1,584 GSR registers were generated[1]. A Convolutional Neural Network (CNN) model was trained with the facial expression dataset, then statistical analysis was performed with the GSR dataset. MobileNet's CNN reached an overall accuracy of 94.36% in a confusion matrix, but the accuracy decreased to 28% for non-trained external images. The statistical results of GSR with significant differences in confused emotions are discussed.

Keywords: facial expression recognition, GSR, MobileNet, learn-related emotions, CNN.

1 Introduction

Emotions as stress [7] or happy [23] can appear during deep learning activities; other emotions as relaxation and boredom can deactivate the learning activity [23]. Consequently, emotion recognition is a useful tool to adjust the learning environment.

Behaviors (e.g., facial expressions and movements) and physiological data (e.g., signals from the brain, heart, skin, muscle, or eyes.) are non-verbal expressions of the human body, which can give relevant information about the emotions

[1] Hispanic Facial Expressions and Galvanic Skin Response (HFEGSR) dataset available at: https://github.com/andresmitre/Study-of-Spontaneous-and-Acted-Learn-Related-Emotions-Through-FER-and-GSR

of a person; these outputs are more notorious in video stimuli [24, 20, 14] than image stimuli [21, 15] from the International Affective Picture System (IAPS) or sound stimuli from the International Affective Digitized Sounds (IADS)[1] because of the simultaneous perception activities. The information generated from stimuli is collected in datasets.

There are several acted [17, 16] or spontaneous [17, 19] FER datasets, but no for all emotions related to learning as stress and boredom, such emotions can be found in physiological signals datasets –e.g., Electrocardiogram (ECG), Heart Rate (HR) and Galvanic Skin Response (GSR)– GSR being the most representative signal in boredom [9] and other negative emotions [3] –e.g. stress– for correctly classification in emotion recognition. In this paper we introduce: (i) a technique in a controlled experiment using video stimuli to produce GSR and FER datasets of the next basic emotions related to learning: happy, stress, and boredom; (ii) a classifier of spontaneous and acted emotions with our FER dataset; and (iii) a GSR dataset analysis of significant differences in spontaneous and acted emotions.

The paper is organized as follows: in section 2 related works are presented; in section 3 the experiment is detailed; section 4 discusses the results; section 5 describes the contributions and future work.

2 Related Works

CNN, has a high computational cost due to its network depth and the need of large dataset, where tuning hyperparameters is difficult as it is slow to train the CNN and there are numerous parameters to configure [22]. Concerning emotional classifiers, Hierarchy CNNs in [13] tested the accuracy on multiple FER datasets: 61.6% for SFEW2.0 [6], 72.72% FER-2013 [10], 87.71% TFD [27] and 95.38% GENKI4K [30]. On the other hand, CNN's classifiers such as MobileNet [26] and ShuffleNet [31], reduce computational cost sacrificing accuracy. MobileNet is built primarily from depth-wise separable convolutions used in inception models [12] to reduce the computation in the first few layers. In [11], MobileNet achieved a 79.4% with 5.60 millions less Mult-Adds (computation) compared to FaceNet [25] who reached a 83% accuracy. MobileNet allows using low-cost technology (less processing) in educational environments.

In the physiological process, the GSR signal has better classification accuracy than other physiological signals in discrimination of happiness or neutral states [5], boredom [9] and stress [28]. Features collected from GSR signal may contribute to the accuracy of learning emotions classification with FER datasets.

Most acted [17, 16] and spontaneous [17, 19] FER datasets do not include physiological signals during the data collection procedure. Moreover, GSR datasets [14] offer the video of participants' face without restriction of video recording. An interesting technique to collect FER images was described in [2], which the configured approach with EEG device triggers the facial image collection during high emotional rates. We used this principle to collect FER images during more representative GSR data.

3 Experiment

The objectives of the experiment are: (i) the creation of the FER and GSR datasets of acted and spontaneous emotions related to the learning process. (ii) the emotion classification with a CNN throughout FER. (iii) Find significant differences in the GSR Percentage of Change (P.C) between spontaneous/acted and neutrality.

3.1 Materials and Method

To create the FER and GSR datasets, the materials we used were: Haar Feature-based Cascade [29] –a machine learning based approach where a cascade function is trained from many positive an negative images– for the face recognition developed in OpenCV; DSRL Rebel T3 Cannon Camera in order to capture the facial images; Grove GSR Sensor –measure the resistance of the human skin, the resistance is measured with two electrodes of 1/4"(6 mm) dimension of nickel material attached to an electronic circuit– for the physiological data; and a 22in LCD monitor.

Table 1. Selected scenes from films for audiovisual stimuli.

Emotion	Film title	Scene description
Neutral	The lover	Marguerite gets in a car, gets off and walks to a house
Happy	When Harry met Sally	Sally fakes an orgasm in a restaurant
Stress	Irreversible	Woman raped and brutally beaten
Boredom	Merrified and Danckert	Two men ironing clothes.

An audiovisual stimuli was shown to the participant to obtain the spontaneous emotion, it consisted of a visualization of video clips from films validated by participants in diverse work related emotion. Table 1. describes the film related to the stimuli and a brief description of the scene. The stimuli from Table 1 was validated by 100 participants in neutral and happy stimuli [20], boredom was approved by two studies (study 1: 241 participants, study 2: 416 participants) [18], and stress validated by 41 participants [4]. On the other hand, within the acted emotions, facial expression was imitated by the participants using FACS [8] and others proposals, for an interval of 10 seconds.

3.2 Participants

A sample of 22 subjects of Hispanic ethnicity participated in the study: 9 females and 13 males with a range from 18 to 62 years of age.

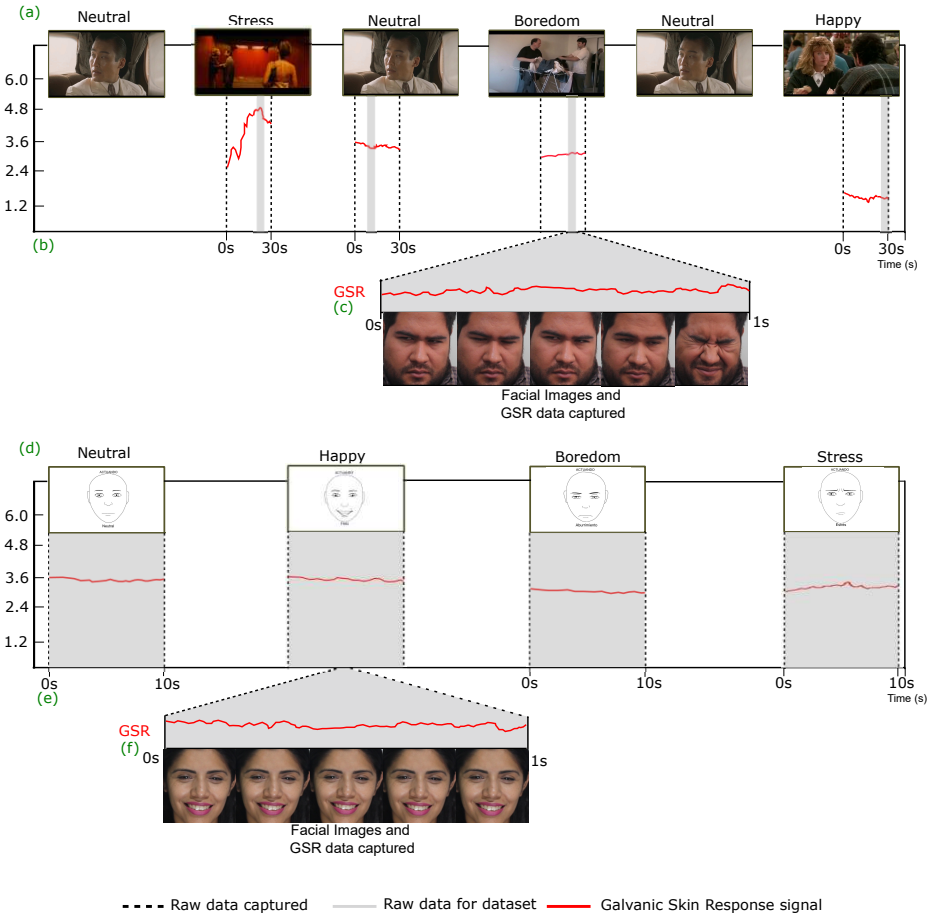


Fig. 1. Procedure of data collection example during spontaneous and acted sessions. Left timeline: (a) films sequence of spontaneous stimuli (neutral, stress, boredom, happy). (b) Timeline procedure for spontaneous stimuli per participant, obtaining the raw data (facial images and GSR) of the session captured within an interval of 30 seconds (gray bars) where the strongest appeared emotion, the GSR signal is noted with the red. (c)(f) Facial images and GSR signal chose for the dataset within a second at 10 FPS. The second was determined according to the biggest difference between the actual emotion recorded, and the baseline (neutral state). Right timeline: (d) acted images sequence (neutral, stress, boredom, happy), the participants imitated the images with the facial expressions described in [8] and other pictorial instructions. (e) Timeline procedure for acted stimuli of 10 seconds per images.

3.3 Procedure

For the creation of the dataset, we proposed a technique that consisted of two sessions: spontaneous and acted. Firstly, the procedure of the experiment

Table 2. MobileNet Body Architecture.

Type/ Stride	Filter Shape	Input Size
Conv / s2	3 x 3 x 3 x 32	224 x 224 x 3
Conv dw / s1	3 x 3 x 32 dw	112 x 112 x 32
Conv / s1	1 x 1 x 32 x 64	112 x 112 x 32
Conv dw / s2	3 x 3 x 64 dw	112 x 112 x 64
Conv / s1	1 x 1 x 64 x 128	56 x 56 x 64
Conv dw / s1	3 x 3 x 128 dw	56 x 56 x 128
Conv / s1	1 x 1 x 128 x 128	56 x 56 x 128
Conv dw / s2	3 x 3 x 128 dw	56 x 56 x 128
Conv / s1	1 x 1 x 128 x 256	28 x 28 x 128
Conv dw / s1	3 x 3 x 256 dw	28 x 28 x 256
Conv / s1	1 x 1 x 256 x 256	28 x 28 x 256
Conv dw / s2	3 x 3 x 256 dw	28 x 28 x 256
Conv / s1	1 x 1 x 256 x 512	14 x 14 x 256
5x Conv dw / s1	3 x 3 x 512 dw	14 x 14 x 512
Conv / s1	1 x 1 x 512 x 512	14 x 14 x 512
Conv dw / s2	3 x 3 x 512 dw	14 x 14 x 512
Conv / s1	1 x 1 x 512 x 1024	7 x 7 x 512
Conv dw / s2	3 x 3 x 1024 dw	7 x 7 x 1024
Conv / s1	1 x 1 x 1024 x 1024	7 x 7 x 1024
Avg Pool / s1	Pool 7 x 7	7 x 7 x 1024
FC / s1	1024 x 1000	1 x 1 x 1024
Softmax / s1	l—Classifier	1 x 1 x 1000

consisted in providing instructions to the participants about the session and how to position their middle and index phalanges over the GSR electrodes. During the session, participants watched a series of films (stimuli) while raw data was recorded –photos taken at 10 FPS and GSR recording at 10Hz.

Figure 1 illustrates the sequence of spontaneous and acted sessions. Fig. 1(a) shows the spontaneous session; the stimuli consisted in the following sequence: Neutral → Stress → Neutral → Boredom → Neutral → Happy; where at the end of every stimuli, neutrality was induced to the participant to generate the correct desired emotion. Fig. 1(b) indicates the raw data taken in an interval of 30 seconds from the scene where the strongest emotion appeared. Fig. 1(c) facial expression and GSR reading captured at 10 FPS and Hz respectively for spontaneous emotion. For the acted emotions (Fig. 1, bottom timeline), participants were instructed to carefully read the instructions for each emotion by performing the imitation during 10 seconds (Fig. 1(d)). Fig. 1(e) indicates the raw taken in an interval of 10 seconds. Fig. 1(f) facial expression and GSR reading captured at 10 FPS and Hz respectively for acted emotion.

The emotion classification was developed with MobileNet Architecture (Table 2) [11] on TensorFlow; 1,840 images were used for the CNN, the images from the FER dataset were resized at 224 x 224 for the input layer, afterward CNN

classified the images into 8 different classes: “Happy_A”,“Happy_S”, “Stress_A”, “Stress_S”, “Boredom_A”, “Boredom_S”, “Neutral_A” and “Neutral_S” (“_A”, stands for Acted emotion and “_S”, stands for spontaneous emotion). For the external images, a total of 80 images (10 images per class) were used to test the CNN; the images were taken from social networks sites, where users declared their emotional state.

4 Results and Discussions

The classification of the trained dataset brings excellent results: training precision: 100%, validation accuracy: 96.5%. On the other hand, the classification for external images with previously trained dataset shows an average poor performance of 28.75%. To evaluate performance of MobileNet, we analyzed classifier based on confusion matrix analysis (fig. 2), where a correct classification was done with an overall accuracy of 94.6% (fig. 2(a)).

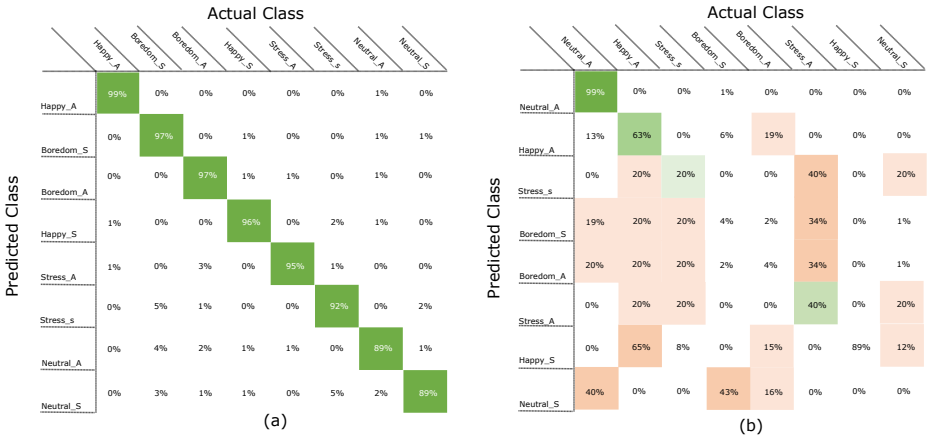


Fig. 2. Confusion Matrix of classifier based on the performance of the CNN. (a) Confusion Matrix of the dataset trained for the CNN, the classification shows an overall accuracy of 94.36%. (b) Confusion Matrix of external images, the classification shows an accuracy of 28.75%. The letters A and S in (a, b) represents the state of emotion, where ”A” stands for acted emotion and ”S” for spontaneous emotion.

Signed Wilcoxon Rank test was used to compare means of change percentage variable of each emotion. For the test with P.C. of spontaneous stress with acted, it was found that Z = -0.141, P > 0.05, no significant differences were found. The test with P.C. of spontaneous happy with the acted one, it was found that Z = 4.46, P = 0.000; P.C. of spontaneous happy (M = 0.196) is greater than happy acted (M = -0.38). Regarding boredom, the result of the test was: Z = -5.167, P = 0.000; P.C. of spontaneous boredom (M = 1.565) is greater than P.C. of

acted boredom (M = 0.006). Therefore GSR can contribute on the recognition of the happy spontaneous confused by a 65% with acted one; in addition the recognition of boredom spontaneous with acted boredom of 2% (fig. 2(b)).

5 Conclusions and Future Work

A classifier of learn-related emotions is a powerful tool for intelligent learning environments, the recognition of emotional states could improve the reinforcement learning of a focus group for education (e.g., educational task and tools) and gaming (e.g., reaction on different levels of immersion). The contribution of this work lies on: i) a new technique to generate a dataset that employs effective stimuli for FER and GSR. ii) Two dataset, FER: 1,840 facial expressions images; GSR: 1,548 registers –baseline difference between emotions and neutrality. iii) Trained model of spontaneous and acted emotions. iv) The possibility of GSR may enhance the classification of the CNN. Future work will consist an including other emotions related to learning (e.g., confusion), as well as incorporating the GSR data into the RGB value channels of the trained images, furthermore, re-train the CNN and contrast results.

References

1. Bradley, M.M., Lang, P.J., Bertron, A., Zack, J., Gintoli, S., Axelrad, J., Cason, J., Brollia, T., Hayden, S., Thorne, B., Karlsson, M., Bittiker, A.: The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective Ratings of Sounds and Instruction Manual. Tech. Rep. 2, University of Florida, Gainesville, Fl (2007)
2. Cabada, R.Z., Estrada, M.L.B., HernÃ¡ndez, F.G., Bustillos, R.O., Reyes-GarcÃa, C.A.: An affective and web 3.0-based learning environment for a programming language. Telematics and Informatics 35(3), 611 – 628 (2018), sI: EduWebofData
3. Colomer Granero, A., Fuentes-Hurtado, F., Naranjo Ornedo, V., Guixeres Provinciale, J., Ausín, J.M., Alcañiz Raya, M.: A comparison of physiological signal analysis techniques and classifiers for automatic emotional evaluation of audiovisual contents. Frontiers in computational neuroscience 10, 74 (2016)
4. Cousijn, H., Rijpkema, M., Qin, S., Marle, H.J.F.V., Franke, B., Hermans, E.J.: Acute stress modulates genotype effects on amygdala processing in humans 107(21), 9867–9872 (2010)
5. Das, P., Khasnobish, A., Tibarewala, D.N.: Emotion recognition employing ecg and gsr signals as markers of ans. In: 2016 Conference on Advances in Signal Processing (CASP). pp. 37–42 (June 2016)
6. Dhall, A., Goecke, R., Lucey, S., Gedeon, T., et al.: Collecting large, richly annotated facial-expression databases from movies. IEEE multimedia 19(3), 34–41 (2012)
7. D'Mello, S., Graesser, A.: Dynamics of affective states during complex learning. Learning and Instruction 22(2), 145–157 (2012)
8. Ekman, P., Rosenberg, E.L.: What the face reveals: Basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). Oxford University Press, USA (1997)

9. Giakoumis, D., Tzovaras, D., Moustakas, K., Hassapis, G.: Automatic recognition of boredom in video games using novel biosignal moment-based features. IEEE Transactions on Affective Computing 2(3), 119–133 (July 2011)
10. Goodfellow, I., Yoshua, B., Aaron, C.: Deep learning. MIT Press (2016)
11. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications (2017)
12. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)
13. Kim, B.K., Roh, J., Dong, S.Y., Lee, S.Y.: Hierarchical committee of deep convolutional neural networks for robust facial expression recognition. Journal on Multimodal User Interfaces 10(2), 173–189 (2016)
14. Koelstra, S., Muhl, C., ..., M.S.I.T., 2012, U.: Deap: A database for emotion analysis; using physiological signals. IEEE Transactions on Affective Computing 3(1), 18–31 (2012)
15. Lang, P., Bradley, M., Cuthbert, B.: Technical report a-8, international affective picture system (iaps): affective ratings of pictures and instruction manual (university of florida, gainesville, fl) (2008)
16. Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D.H.J., Hawk, S.T., van Knippenberg, A.: Presentation and validation of the radboud faces database. Cognition and Emotion 24(8), 1377–1388 (2010)
17. Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I.: The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In: Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on. pp. 94–101. IEEE (2010)
18. Markey, A., Chin, A., Vanepps, E.M., Loewenstein, G.: Identifying a Reliable Boredom Induction. Perceptual and Motor Skills 119(1), 237–253 (2014)
19. Mavadati, S.M., Mahoor, M.H., Bartlett, K., Trinh, P., Cohn, J.F.: Disfa: A spontaneous facial action intensity database. IEEE Transactions on Affective Computing 4(2), 151–160 (2013)
20. Megías, C.F., Mateos, J.C.P., Ribaudi, J.S., Fernández-Abascal, E.G.: Validación española de una batería de películas para inducir emociones. Psicothema 23(4), 778–785 (2011)
21. Nourbakhsh, N., Chen, F., Wang, Y., Calvo, R.A.: Detecting Users' Cognitive Load by Galvanic Skin Response with Affective Interference. ACM Transactions on Interactive Intelligent Systems 7(3), 1–20 (2017)
22. Pérez-Espinosa, H., Avila-George, H., Rodriguez-Jacobo, J., Cruz-Mendoza, H.A., Martínez-Miranda, J., Espinosa-Curiel, I.: Tuning the parameters of a convolutional artificial neural network by using covering arrays. Research in Computing Science 121, 69–81 (2016)
23. Rainhard Pekrun, R.P.P.: Control-value theory of achievement emotionsNo Title. In: International Handbook of Emotions in Education, chap. 7, p. 22 (2014)
24. Sauter, D.A., Fischer, A.H.: Can perceivers recognise emotions from spontaneous expressions? Cognition and Emotion 32(3), 504–515 (2018)
25. Schroff, F., Kalenichenko, D., Philbin, J.: Facenet: A unified embedding for face recognition and clustering. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 815–823 (2015)
26. Sifre, L., Mallat, S.: Rigid-motion scattering for image classification. Ph.D. thesis, Citeseer (2014)

27. Susskind, J.M., Anderson, A.K., Hinton, G.E.: The toronto face database. department of computer science, university of toronto, toronto, on. Tech. rep., Canada, Tech. Rep, 3 (2010)
28. Villarejo, M.V., Zapirain, B.G., Zorrilla, A.M.: A stress sensor based on galvanic skin response (GSR) controlled by ZigBee. Sensors (Switzerland) 12(5), 6075–6101 (2012)
29. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001 1, I–511–I–518 (2004)
30. Whitehill, J., Serpell, Z., Yi-Ching Lin, Y.C., Foster, A., Movellan, J.R.: The Faces of Engagement: Automatic Recognition of Student Engagementfrom Facial Expressions. IEEE Transactions on Affective Computing 5(1), 86–98 (2014)
31. Xiangyu, Z., Xinyu, Z., Mengxiao, L., Jian, S.: Shufflenet: an extremely efficient convolutional neural network for mobile devices. In: Computer Vision and Pattern Recognition (2017)

ISSN 1870-4069

Synaptix: A Web Platform based on Gamification Techniques for the Study of Clinical Cases

Omar López Chávez, Ignacio N. Márquez, Luis-Felipe Rodríguez, Jorge G. Mendoza León

Sonora Institute of Techonology, Department of Computer and Design, Mexico
omarlopch@gmail.com, ignacio_nmarquez@hotmail.com, luis.rodriguez@itson.edu.mx, jorge.mendoza@itson.edu.mx

Abstract. In this paper we present a Web platform designed to allow medical students and practitioners to study clinical cases on their own. The main objective of this proposal is to address key problems inherent in the traditional study of clinical cases by providing a tool that implements techniques and elements of gamification, simulation, and serious games. The proposed platform offers an improved learning experience through a virtual environment that provides an alternative method for training and interpretation of clinical cases for medical examinations. Medical students and practitioners can play the role of a real doctor in a simulated office. In particular, the platform allows medical students and practitioners to learn through their mistakes without hurting human beings. In addition, this platform is designed to allow users to add new clinical cases and make them available for study. The platform was validated in a local hospital by 8 medical practitioners. The participants indicated that the platform design, the tutorials included, and the ease of use factors are satisfactory.

Keywords: medical student, clinical case, gamification, simulation, serious game.

1 Introduction

Gamification is the process of changing a set of traditional actions to an attractive gaming experience for the user [13]. Matallaoui *et al.* [11] define gamification as "the use of game design elements in a context not related to the game, is an interdisciplinary tool, where users are motivated to achieve certain behavioral or psychological results". Gamification also enables the development of immersive games in virtual environments in which users are encouraged to perform desired actions. In the academic field, gamification serves as a tool to facilitate teaching and learning processes through collaborative environments [2].

Serious games (SGs) have a high impact as an instructional tool that benefits from traditional game concepts and information and communication technologies. Serious games have allowed the implementation of simulations and realistic

virtual environments, where players can experience adventures while acquiring, practicing, and verifying knowledge. This represents a significant opportunity for 21st century educators and trainers to improve their educational tools [1].

The traditional method for studying clinical cases is through information sources such as books, articles, and automated tests. However, this method usually leads medical students to a state of saturation, stress, and anxiety given that, for example, the quantity and complexity of clinical cases to study in exam periods is high. Moreover, the feedback received from this type of information sources is limited. The study of clinical cases by medical students and practitioners also takes advantage of the monitoring of patients in hospitals. This study of clinical cases requires observation and analysis for long periods of patients with diverse conditions, diseases, signs, and symptoms [9]. However, in hospitals such as those known as "third level hospitals", only critical patients have long stays. The majority of patients are hospitalized for short periods of time, which makes it difficult the generation and access of medical students to a greater amount of knowledge that improves their learning experience and the acquisition of skills related to clinical cases that depend on hospitalized patients.

Lifshitz [8] indicates that clinical cases can not be learned through memorization or readings or through distance education strategies. In fact, the analysis of clinical cases has a very strong affective component because it implies confrontation with illness and suffering. The discussion about the limits of the teaching of clinical cases has not been solved, but this type of learning usually requires abilities for communication, physical examination, treatment, and clinical reasoning. Lifshitz [8] also proposes a well defined structure to organize and study clinical cases: 1) the approach to the patient, 2) the collection of information, 3) the analysis of the information collected, 4) clinical procedures, 5) the diagnostic decision, 6) the decision therapist, and 7) the decision Prognosis. In this context, although there is great interest in the design and application of guidelines in clinical practices for the prevention and care of diverse health situations, greater attention must be paid to its implementation and effectiveness in various practical scenarios [5].

The simulation of clinical cases involves a set of techniques that facilitate medical students and practitioners the acquisition of knowledge and skills. In particular, techniques and methods from fields such as artificial intelligence, virtual and augmented reality, and human-computer interaction have enabled the development of platforms that incorporate virtual scenarios, simulation models, and multimedia materials to simulate different real situations (e.g., for the analysis of clinical cases) [4]. Although simulation platforms do not replace the real scenarios, these allow students to learn and practice in controlled media, contributing to improve their skills and decrease the anxiety when performing an exam or procedure. This type of platform also accelerates learning and enriches the true interactions with the patients, helping medical students to avoid states of saturation experienced with traditional learning methods [9].

In this paper, we present a platform designed to allow medical students and practitioners to study clinical cases on their own. This platform represents an

attempt to address key problems inherent in the traditional study of clinical cases by providing a tool that implements techniques and elements of gamification, simulation, and serious games. Its design takes into account the elements and strategies that according to medical students and practitioners are required to learn and practice clinical cases [10]. The paper is structured as follows. In Section 2 we discuss related work. The proposed platform and corresponding validation are presented in Section 3 and Section 4, respectively. Finally, concluding remarks and future work are discussed in Section 5.

2 Related Work

The website *The New England Journal of Medicine* presents interactive medical clinic cases designed according to the following interaction process: presentation of the case, medical history of the patient, information of the physical examination, and finally, performs a test to provide feedback and solutions of correct and incorrect answers [6]. However, although this website shows to users the percentage of the result and the studies carried out, key elements of gamification and serious games are not considered such as *dashboard*, *unlock*, and *challenge*. Moreover, the interaction is based on text and 2D graphics, leaving aside the implementation of 3D scenarios and simulations. The inclusion of additional clinical cases by users is not allowed.

Nevin *et al.* [12] developed the Kaizen-Internal Medicine (Kaizen-IM) software that includes elements of gamification. This tool involved a large number of residents in a medical contest that facilitated the acquisition of new knowledge in the academic period 2012-2013 in two training programs IM (internal medicine) in the USA: the residency program in Internal Medicine at the University of Alabama at Birmingham (UAB) and the University of Alabama Program at Huntsville (UAH). The data was recorded at participant level and question. The analyzes focused on the acceptance, use, and determination of the factors associated with the loss of players (attrition) and the retention of knowledge. The Kaizen-IM data provided information on modifiable factors associated with student attrition and retention of knowledge that can serve to further enhance the educational benefits of this strategy for students. This tool is an attempt to demonstrate the benefits of incorporating elements of gamification in the learning process of medical students. However, this software does not implement 3D simulations, but is based on strategies based on questions and answers.

Leba *et al.* [7] proposed an application for the training of medical students in the field of anatomy with computed tomography images, using elements of gamification, simulation, and serious games. It was designed in the context of an educational software. This proposal represents an attempt to support modern and practical methods of examination based on real cases useful for medical students and teachers. However, this application for training was not validated in a case study, the authors proposed only design guidelines.

3 Synaptix

The Sypnatix platform was designed for the study of clinical cases that follow the clinical practice guide of the federal government of Mexico [3]. Synaptix implements elements and techniques of gamification, simulation, and serious games. In particular, its design is based on a previous study by Marquez *et al.* [10] in which data was collected from medical students and practitioners in order to 1) identify key elements and learning strategies for the study and practice of clinical cases, and 2) define how these elements and strategies should be taken into account in the design of a learning platform that incorporates elements of gamification for the study of clinical cases (see Figure 1) [10].

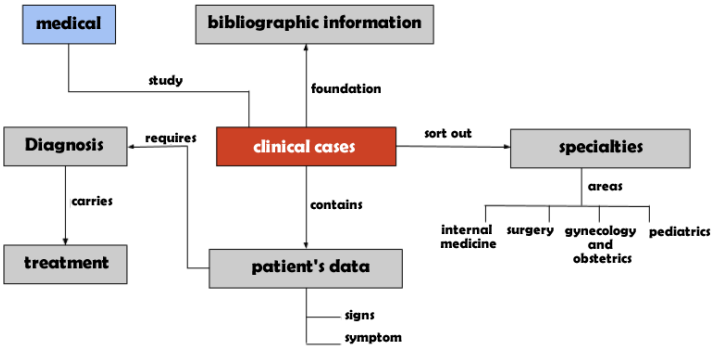


Fig. 1. Elements involved in the study of clinical cases [10].

The Synaptix platform incorporates the following elements of gamification (which are common elements reported in the literature [14]):

- Points: represent the way in which the player is observed, classified, and guided. In Synaptix, the user starts with 0 points. The points are cumulative and depend on the correct answers provided by the user.
- Badges: mark the fulfillment of goals and the constant progress of the game. Badges are activated as the user completes clinical cases on Synaptix.
- Levels: a marker for players to know where they are in a gaming experience. In Synaptix, levels are activated as progress is made in solving clinical cases.
- Dashboard: an ordered list of names and its corresponding score. Synaptix makes available information about the achievements of each user.
- Unlock: allows players to access another achievement after certain requirements are met. In Synaptix, objectives are unlocked as the user resolves clinical cases.
- Challenges: offer players an address so they know what to do within the world of the gamified experience. In Synaptix, the challenges are associated with obtaining badges.

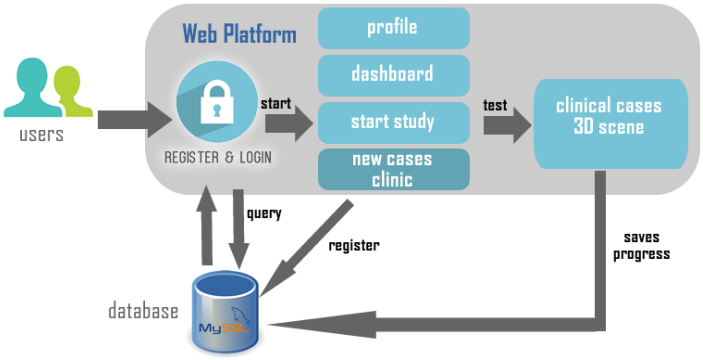


Fig. 2. The architecture of Synaptix.

Figure 2 shows the main components of the Sypnatix's architecture and their relationship: 1) the elements of gamification mentioned above, 2) a 3D virtual scenario, and 3) components related to the data management. These components create an interactive experience by enabling medical students to acquire, practice, and verify knowledge, which are key elements of SG and simulation.

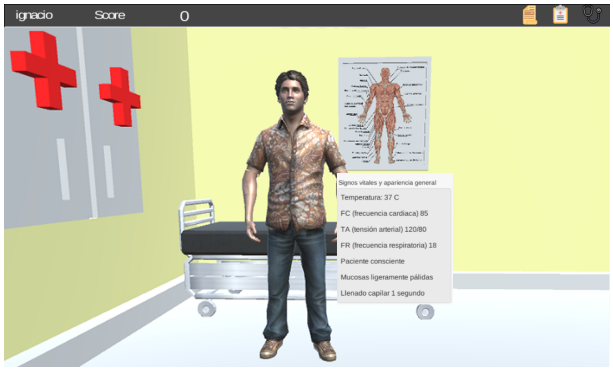


Fig. 3. Simulation of a doctor office and a virtual patient.

The virtual scenario provided by Synaptix for the study and practice clinical cases is shown in Figure 3. This virtual scenario simulates a medical office and a virtual patient. Medical students interact with the virtual patient by physically examining, for example, its lungs and head. Additional information related to clinical records or results of clinical studies (data that is sometimes taken into account in clinical cases) is also displayed in the scenario. After the medical student analyzed the data provided in the clinical case (i.e., physically examined the virtual patient and its medical records), a series of questions and possible

answers are displayed to evaluate whether the diagnosis and treatment suggested by the student are correct. Synaptix provides feedback to medical students once an answer is submitted. If the answer is incorrect, feedback about medications or treatments is provided. The points obtained and the assigned badges are activated in the user's profile as shown in Figure 4.

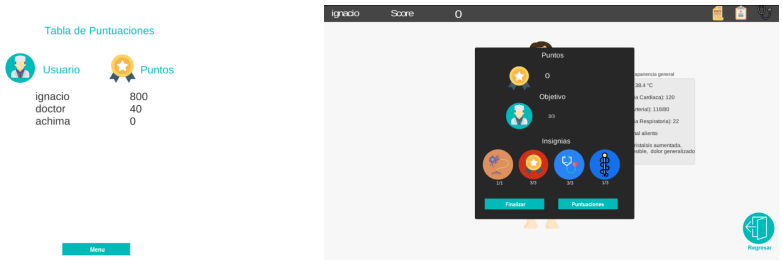


Fig. 4. Dashboards that show the scores and points obtained by medical students.

Although Synaptix is designed for the study of clinical cases related to the areas of internal medicine, surgery, gynecology and obstetrics, and pediatrics, currently only clinical cases of internal medicines are included. Nevertheless, Synaptix allows medical practitioners to include new clinical cases which are then available for their analysis by medical students. The tools used in the development of the Synaptix Web platform were Unity game development platform, Php scripting language, and the Mysql database server. In order to carry out the tests, it was posted on a web server: http://arevolution.com.mx/synaptix/

4 Validation

Synaptix was validated in a private local hospital by 8 medical students and practitioners (4 female and 4 male). The validation session consisted of an introduction by the authors about the functionality and characteristics of the platform. Afterwards, the participants used the platform to practice available clinical cases and answered a questionnaire. The instrument consisted of 29 items to measure 1) the design of the platform, 2) the instructions provided by Synaptix, and 3) easy of use factor. Figure 5 and Figure 6 show the results obtained in this evaluation phase.

The comments provided by participants include the following: 1) provide further feedback or suggest additional information sources to the user once a clinical case is carried out, 2) allow the user to make more questions to the virtual patient, 3) indicate the specific areas the medical student needs to reinforce in order to achieve better results, 4) provide greater details in the diagnosis, 5) include better medical images, and 6) include more references about the clinical cases presented. Finally, some participants emphasized the importance of badges and individual scores as incentives.

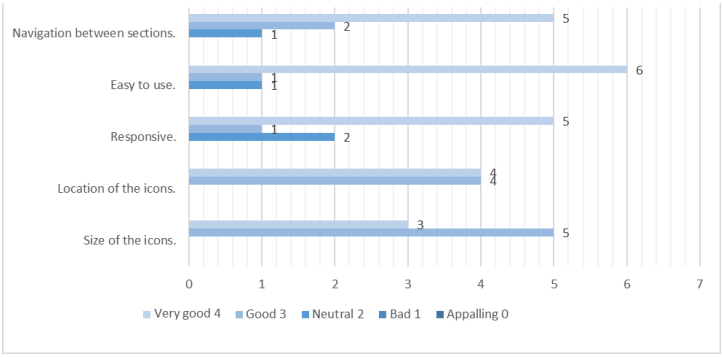


Fig. 5. Results of evaluating the design of Synaptix.

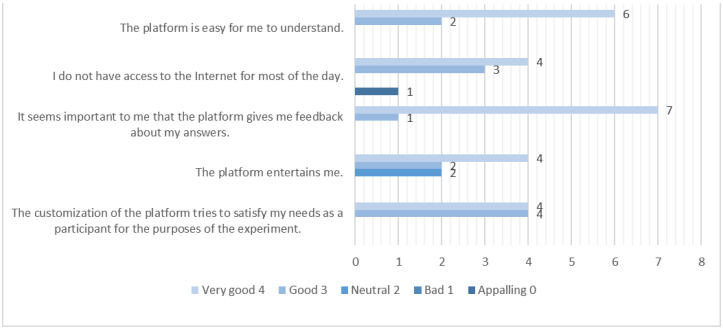


Fig. 6. Results of evaluating the easy of use of Synaptix.

5 Conclusions and Future Work

The main contribution of this paper is the design and implementation of a platform that 1) incorporates elements of gamification, serious games, and simulation, and 2) takes into account the elements and strategies that according to medical students and practitioners are required to learn and practice clinical cases. The Synaptix Web platform is a gamified tool that attempts to serve as an alternative method for the study of clinical cases, avoiding states of saturation, fatigue, and anxiety in medical students and practitioners. Synaptix was validated in a private local hospital by 8 medical students and practitioners. The results demonstrated that participants find that the factors associated with the *design* and *easy of use* of the platform is in general satisfactory. Future research involves an evaluation to measure complex individual's aspects such as motivation, learning, and user engagement as well as the incorporation of virtual

reality and augmented reality components in order to create more immersive scenarios. Furthermore, AI techniques will be incorporated to enable the virtual patients to develop some human-like behaviors.

Acknowledgment. This work was supported by PFCE 2018.

References

1. Arnab, S., Lim, T., Carvalho, M.B., Bellotti, F., De Freitas, S., Louchart, S., Suttie, N., Berta, R., De Gloria, A.: Mapping learning and game mechanics for serious games analysis. British Journal of Educational Technology 46(2), 391–411 (2015)
2. Burke, B.: Gamify: How gamification motivates people to do extraordinary things. Routledge (2016)
3. Cenetec: Clinical practice guide. `https://cenetec-difusion.com/gpc-sns/` (2018)
4. Chen, L., Day, T.W., Tang, W., John, N.W.: Recent developments and future challenges in medical mixed reality. In: IEEE International Symposium on Mixed and Augmented Reality (ISMAR). pp. 123–135. IEEE (2017)
5. DiCenso, A., Guyatt, G., Ciliska, D.: Evidence-Based Nursing-E-Book: A Guide to Clinical Practice. Elsevier Health Sciences (2014)
6. Group, N.: interactive medical case. `https://www.nejm.org/multimedia/interactive-medical-case` (2018)
7. Leba, M., Ionica, A., Apostu, D.: Educational software based on gamification techniques for medical students. In: Proceedings of the 5th International Conference on Applied Informatics and Computer Theory (AICT). pp. 225–230 (2014)
8. Lifshitz, A.: La nueva clínica. Intersistemas Editores (2014)
9. Makransky, G., Bonde, M.T., Wulff, J.S., Wandall, J., Hood, M., Creed, P.A., Bache, I., Silahtaroglu, A., Nørremølle, A.: Simulation based virtual learning environment in medical genetics counseling: an example of bridging the gap between theory and practice in medical education. BMC medical education 16(1), 98–107 (2016)
10. Márquez, I., Mendoza, J.G., Rodríguez, L.F.: Identificación de elementos clave en el estudio de casos clínicos para su gamificación. In: Prieto, M., Pech, S., Francesa, A. (eds.) Tecnologías y Aprendizaje. Investigación y Práctica, pp. 132–139 (2018)
11. Matallaoui, A., Herzig, P., Zarnekow, R.: Model-driven serious game development integration of the gamification modeling language gaml with unity. In: 48th Hawaii International Conference on System Sciences (HICSS). pp. 643–651. IEEE (2015)
12. Nevin, C.R., Westfall, A.O., Rodriguez, J.M., Dempsey, D.M., Cherrington, A., Roy, B., Patel, M., Willig, J.H.: Gamification as a tool for enhancing graduate medical education. Postgraduate medical journal pp. 121–130 (2014)
13. Robson, K., Plangger, K., Kietzmann, J.H., McCarthy, I., Pitt, L.: Is it all a game? understanding the principles of gamification. Business Horizons 58(4), 411–420 (2015)
14. Vargas-Enríquez, J., García-Mundo, L., Genero, M., Piattini, M.: Análisis de uso de la gamificación en la enseñanza de la informática. In: Actas de las XXI Jornadas de la Enseñanza Universitaria de la Informática. pp. 105–112. Universitat Oberta La Salle (2015)

ISSN 1870-4069

Towards Personalized Summaries in Spanish based on Learning Styles Theory

Uriel Ramírez[1], Yasmín Hernández[2], Alicia Martínez[1]

[1] Tecnológico Nacional de México, CENIDET, Cuernavaca, Mexico
[2] Instituto Nacional de Electricidad y Energías Limpias, Gerencia de Tecnologías de la Información, Cuernavaca, Mexico
{uriel.ramirez18ce, amartinez}@cenidet.edu.mx,
yasmin.hernandez@ineel.mx

Abstract. Today, advances in information technologies have generated perhaps the largest and fasted exponential growing of electronic texts. On the Internet there are many electronic documents, such as books, technical documents, news articles, blogs, chats, emails and many other digital files. As a result, a user who wants to read and understand this information in a short time will find it a hard task. In this paper, we have conducted an important work in automatic text summarization. Also, we have considered the particular needs of readers. Thus, a model for personalized summarization base on learning styles theory is proposed.

Keywords: automatic text summarization, concept graph, domain specific summarization, multi-document summarization, natural language processing.

1 Introduction

At the presents, advances in information technologies have generated perhaps the largest and fasted exponential growing of electronic texts. Previous studies indicate that the excessive handling of information could cause inefficiency in work, in addition to anxiety and stress [1]. Therefore, in many fields, domains and applications exist the necessity for developing tools that summarized information with different purposes. For example, in Education field the summaries are very needed.

The automatic generation of text summaries consists in taking a text and condensing it in a way sensitive to the needs of the reader [2]. Thus, a useful summary for a reader will be one that takes into account the particular characteristics of readers, such as: knowledge, interests, age, learning styles, etc.

The investigations carried out on Automatic Text Summarization started more than five decades ago with the previous work of Edmundson [3].

Thenceforth, several theories in linguistics and artificial intelligence have been proposed. Such as, superficial techniques [4,5,6], graph-based techniques [7,8,9,10], algebraic reduction techniques [11], statistics [12], etc. However, these researches are still improving. These researches are looking for methods to generate summaries similar to those generated by humans.

Learning styles theory states that every individual has a particular way to learn which includes strategies and preferences. This theory emphasizes that individuals perceive and process information in different ways. Accordingly, learning styles theory states that learning has more to do with a process focusing the learning style than with the intelligence of individuals. Several learning styles models have been proposed, Felder-Silverman learning Styles Model, which is a well-known and broadly used learning styles model [13,14].

In this research work, a model for personalized summarization base on a student model has been proposed. This model will represent the current state of a student. Also, the model includes his/her learning style, among other characteristics. For to do this, we have worked on model the affect and the learning styles of students in order to provide them with more adequate instruction [15,16,17]. Besides, we propose a framework based on conceptual graph for abstractive text summarization. The working language is Spanish.

The paper is organized as follows: Section 2 presents the Background on Automatic Text Summarization and learning styles models. Section 3 describes our approach for a model personalized summarization. Finally, Conclusions and future work are discussed in Section 4.

2 Background and Related Work

2.1 Automatic Text Summarization

The Automatic Text Summarization is the process of extracting or collecting important information from original text, and showing as summary [18]. The Summaries according its purpose can be classified as: generic, domain-specific and query-based on the other hand, the summaries according its output can be abstractive or extractive [19].

As it says, single-document summarization generates a summary from single document input. Conversely, multi-document summarization generates a summary from multiple documents input. Generally, these multiple documents discuss the same topic. Summarization of generic purpose summarizes all documents regardless of topics or domains [19]. While the domain-specific summary focuses on a domain of interest. For example, politics, finance, IT or weather news. The query-based summary deals with the information that is requested from users. In general, queries are sentences in natural language or keywords related to a particular topic [20].

Extractive summarization extracts relevant sentences from the original documents and link them to generate the summary with no changes in the original sentences. In the abstractive summarization, source documents are analyzed and understood using linguistic method to analyze and interpret the document [18,21].

Current research proposes several and diverse methods for automatic text summarization such as statistical [22], machine learning [23,24], text connectivity [25,26], conceptual graphs [27,28,29], algebraic reduction [30], clustering and probabilistic models [31,32,33] and methods adapted to the reader [34,35].

2.2 Concept Graphs

Conceptual graphs are structures for knowledge representation based on first-order logic. Graphs are a natural, simple, and fine-grained semantic representation which can be used to describe texts [27].

Generally, a graph can be denoted in the form G = (V, E), where V represents vertices in the graph and E represents edges between each vertex. In the context of text documents, vertices represent sentences and edges are the weight between two sentences. Consequently, in a graph representing a documents each sentence is a vertex and the weight between each vertex corresponds to the similarity between the two sentences [36].

Figure 1 presents a conceptual graph standing for sentence Manuel hit the piggy bank with a hammer. The concepts are *Manuel*, *Hit*, *Hammer* and *Piggy bank*. These concepts are connected by the relations *agent*, *inst* and *pacient*.

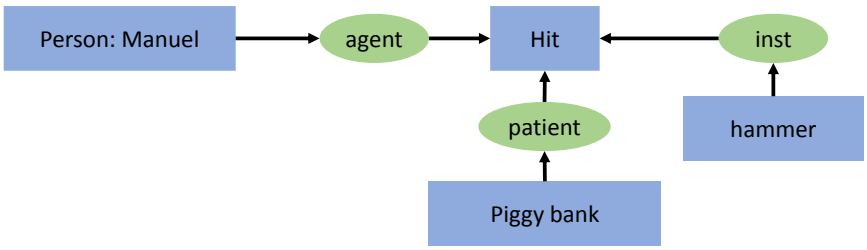


Fig. 1. Conceptual graph for sentence *Manuel hit the piggy bank with a hammer* [36].

2.3 Evaluation Measures for Text Summarization

An important part of the automatic text summarization is the evaluation of the process and results. There are two main tendencies: i) intrinsic evaluation which seeks to measure the quality of the summary, and ii) extrinsic evaluation which measures the performance of the summary in a particular task.

The intrinsic evaluation is based on a standard to be compared with the generated summary. The standard can be an existing dataset (text/summary) or an *ideal* summary generated by a human. On the other hand, extrinsic evaluation requires to select a task which use the summary, and to measure the impact of using the generated summary instead of the original text. In this evaluation, the problems are the selection of the task and the indicators for the measurement [37].

The most common measures to evaluate the quality of the summaries are i) the Measure Recall-Oriented Understudy of Gisting Evaluation (ROUGE) [38] , and ii) Standard measures of information retrieval, such as precision, recall and F-measure. These measures are based on the comparison of *n*-grams between the automatic summary and a manually generated summary. There are several variants of ROUGE: ROUGE-*n*, ROUGE-*L,* ROUGE-*SU.*

2.4 Learning Styles Models

Learning theories describe how people learn concepts and abilities. Diverse learning theories have been proposed, all of them states different, and sometimes, opposite viewpoints. The learning styles theory states that individuals have a particular way to learn which includes strategies and preferences. This theory highlights that individuals perceive and process information in different ways. Consequently, learning styles theory states learning of individuals has more to do with a process focusing the learning style than with the intelligence of individuals[39,40].

Several learning styles models have been proposed, such as Felder-Silverman learning Styles Model [13,14]; which proposes sensing-intuitive, visual-verbal, active-reflective, and sequential-global as categorizations of learning styles. Learning theory of Kolb works on two levels: a four-stage cycle of learning and four separate learning styles. This theory is concerned with internal cognitive processes of learners.

In this sense, the Honey and Alonso questionnaire on learning styles (CHAEA, for the Spanish *Cuestionario de Honey y Alonso de Estilos de Aprendizaje*) has been used as a measurement tool to identify the different types of learning styles presented by students [41]. The instrument consists of 80 items presented in the form of questions, of which 20 correspond to learning styles: active, reflective, theoretical and pragmatic. The affirmative answers are added and the total score is obtained for each style, which is compared with the respective scales.

3 Model for Personalized Summarization

Currently, there are different types of summaries; these will depend on the purpose that the reader has to generate them. The automatic generation of abstracts can generate summaries without taking into account the reader for whom they are intended. This can result in poor understanding or a loss of readers' time. On the other hand, the summaries can be adapted to the peculiarities of the reader, such as: previous knowledge, areas of interest, information needs, learning styles, among other characteristics, to support a better understanding and also to comply with the required information. However, most of the work that has been done revolves around the generic summary [42].

As previous work, we have worked with intelligent training environments for electrical industry where instruction is presented according to particular needs of students [18,19]. In this sense, we have a student model which represents previous knowledge, affective state, personality and learning styles of students [20]. In these training environments, students have to read many handbooks, regulation documents and technical documents. We have notice that in some specific training situations, students do not want to read. They prefer to conduct the practices in a trial and error way. For this reason, we think if students have a personalized summary, they could grasp some aspects that could help them to follow their exercises or lessons. For example, a personalized summary could be useful before to start a lesson to decide which topics to learn, or before to take a quiz or to conduct a practice to review the required material. However, we want to focus the particular needs of students considering their individual characteristics in summarization process.

Additionally, in the academic field, graduate students need to read many research works in order to conduct their work and in order to have success in their studies. In this context personalized summaries are also needed.

Considering these two different domains. We propose a model for personalized summarization in intelligent learning environment. The summarization model will generate abstractive, rather than extractive summaries, for Spanish multi-document source. Fig. 2 shows the architecture of the intelligent learning environment with the model for personalized summarization.

In a high level, the intelligent learning environment consists of student model, tutor module, expert module and domain knowledge. In order to generate the personalized summaries, the model for text summarization was integrated. The personalization process is based on a student model.

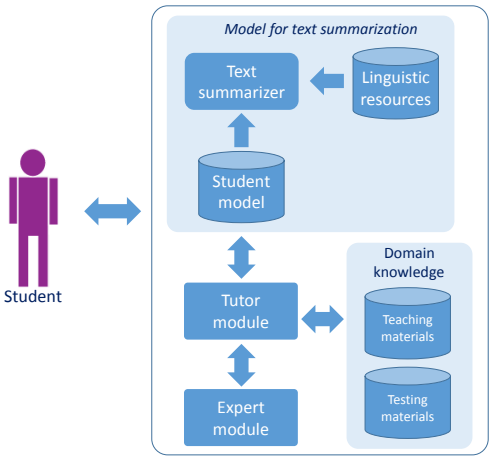


Fig. 2. Architecture of the intelligent learning environment with the model for personalized summarization.

The student model is built with base on the student interaction and its performance in lessons, practices and quizzes. The student model contains following information about students: previous knowledge, affective state, personality, learning styles, age and schooling.

In an initial approach, we wanted to provide students with an adaptive summary according to their learning style. Our proposal for learning styles is based on the CHAEA instrument [43].

The summarization model also integrates linguistic resources built taking into account knowledge about specific topics.

This proposal will be applied in the electrical domain, and some topics on Computer Science. We will use techniques based on conceptual graphs for representing the sentences, entities and their relationships.

Our proposed method aims to generate a summary from a set of documents. However, if the documents are not related, a non-representative or incoherent summary could be obtained. Therefore, we propose first grouping related input documents, and then generating the summary.

The method of summarizing consists on several tasks such as: pre-processing, transformation, synthesis and sentences selection. Fig. 3 shows our proposed method for summarization. This proposed method includes four stages: i) Preprocessing stage, ii) Transformation stage, iii) Synthesis stage, and iv) Sentences selection stage.

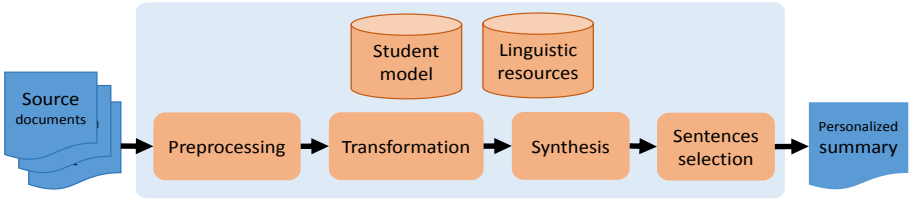


Fig. 3. Proposal for the summarization model including preprocessing, transformation, synthesis and sentences selection.

Preprocessing stage: this stage has aims analyzing the collection of documents, such as: segmentation of the text, extraction of tokens, elimination of empty words, identification of the lexical root of the words and morph-syntactic analysis. The segmentation of text consists on defragmenting the text in paragraphs and sentences. In the extraction of tokens, each sentence is composed of tokens that are not more than each of the parts of the sentence, that is, words, numbers, punctuation marks, etc.

In addition, disambiguation mechanism to discern between different meanings is also necessary to apply in this proposed method. Therefore, it is necessary to translate the lexicon of the document into concepts of the knowledge base of the domain. The disambiguation algorithm as linguistic resources must take into account the domain of the text to be summarized. The disambiguation algorithm will analyze the input sentences; and, as result, a list of concepts or meanings of the knowledge base will be obtained.

Transformation stage: Preprocessing and transformation is based on the model proposed by Miranda, Gelbukh, and Sidorov [37]. For the transformation, this model uses linguistic resources in English language, whereas our model will use linguistic resources in Spanish language, and a user model. In this stage, the representation in the form of a graph will be constructed for each sentence of the document. Capturing the semantic structure and the relationships between terms.

Synthesis stage: The synthesis method is based on a set of operations in Conceptual graphs (GCs): generalization, union or association, weighting and pruning [37].

The generalization operation combines two GCs according to their common elements. For example, the following graphs can be read as G1: Peter buys a crocodile (Peter buys a crocodile) and G2: Mary buys a bird (Mary buys a bird). The GCs are compared and, subsequently, the minimum common concepts to unite them are determined. The hierarchy for crocodile and bird, Animal is the minimum common concept between both concepts; and Person is the least common concept for Peter and Mary. So G3 is the resulting graph after combining the two graphs. G3 can be read as "Two persons buy two animals" (Two people buy two animals).

The union is the operation unites two related concepts of two GCs. This operation supports and improves the results of the weighting process.

The weighting is the operation that filters the most important nodes of the graph. To determine the importance of the nodes, the HITS algorithm is used. HITS is an iterative algorithm that takes into account the degree of input and the degree of output of the node to determine its importance.

Pruning is the operation that is applied to reduce graphs. This operation takes into account the results of the weighting, the verbal patterns to remove the irrelevant nodes and the compression rate or threshold to establish how many nodes should be included in the resulting summary.

Sentences selection stage. This stage aims to select the sentences that will be part of the final summary. Our hypothesis is that knowing the model of the student it will be possible to create a personalized summary. The summary will be constructed from a set of rules and algorithms that will be based on the student's learning style. Subsequently, the sentences are selected according to the order of the documents in which they appear first, and then in the order in which they appear inside the document, to avoid inconsistencies. As a result, the collection of documents is summarized.

In order to evaluate our model, the resulting summary will be evaluated with ROUGE metrics [38]. We use ROUGE metrics because they are automatic metrics which do not need human judges to measure the quality of the summaries and because they are standard metrics used by several research projects.

4 Conclusions and Future Work

The exponential growing of digital documents has motivated the development of technologies as the natural language processing. A field of this area is the automatic text summarization, which is the process of extracting or collecting important information from original text, and showing as summary. Despite there have been diverse efforts, automatic text summarization is in an initial stage. Most researches are interested in to generate a standard summary where particulars needs of readers are not considered.

We propose a model of a personalized summarization according on a student model. This due to the model intends integrate to an intelligent learning environment.

For the time being, we propose to personalize summaries based on the learning styles theory. However, we want examine other ways to personalize the summary; such as, schooling level, age, among other characteristics.

An important contribution of our model will be the linguistic resources. Which will be constructed in the Spanish language. Currently, there is no great variety of resources for the Spanish language.

In this paper, we presented our proposal method for generating Personalized Summaries in Spanish based on Learning Styles Theory. Although this work is in an initial state, we are working on the generation of the user's model.

Acknowledgments. This research work has been partially funded by European Commission and CONACYT, through the SmartSDK project.

References

1. Klingberg, T.: Information Overload and the limits of working memory. Stockholm, Sweden (2009)
2. Steinberger, J., Ježek, K.: Evaluation measures for text summarization. Comput. Informatics 28(2), pp. 251–275 (2009)
3. Edmundson, H.P.: New Methods in Automatic Extracting (1969)
4. Hahn, U., Mani, I.: Challenges of automatic summarization. Computer 33(11), pp. 29–36, Long Beach, Calif. (2000)
5. Ulrich, J., Carenini, G., Murray, G., Ng, R.: Regression-Based Summarization of Email Conversations. Intelligence, pp. 334–337 (2009)
6. Wan, S., McKeown, K.: Generating overview summaries of ongoing email thread discussions. In: Proc. 20th Int. Conf. Comput. Linguist. - COLING '04, p. 549–es (2004)
7. Plaza-Morales, L.: Uso de Grafos Semánticos en la Generación Automática de Resúmenes y Estudio de su Aplicación en Distintos Dominios: Biomedicina , Periodismo y Turismo. p. 352 (2011)
8. Mihalcea, R., Tarau, P.: TextRank: Bringing order into texts. In: Proc. EMNLP, vol. 85, pp. 404–411 (2004)
9. Brin, S., Page, L.: Reprint of: The anatomy of a large-scale hypertextual web search engine. Comput. Networks 56(18), pp. 3825–3833 (2012)
10. Mittal, N., Agarwal, B., Vijay, N., Gupta, A., Upadhyay, N.K.: Semantic Enhanced Text Summarization. Vol. 1, no. 1, pp. 26–29 (2014)
11. Vicente, M., Lloret, E.: The University of Alicante at MultiLing 2015 : approach , results and further insights. No. September, pp. 250–259 (2015)
12. Aries, A., Zegour, D.E., Hidouci, K.W.: AllSummarizer system at MultiLing 2015 : Multilingual single and multi-document summarization. In: Proc. MultiLing 2015, vol. 2015, no. September, pp. 237–244 (2015)
13. Graf, S., Kinshuk: Advanced Adaptivity in Learning Management Systems by Considering Learning Styles. In: IEEE/WIC/ACM Int. Jt. Conf. Web Intell. Intell. Agent Technol., vol. 3, p. 235–238 (2009)
14. Felder, R., Silverman, L.: Learning and teaching styles in engineering education. Eng. Educ., vol. 78, no. June, pp. 674–681 (1988)
15. Model, L.S.: Learning styles theory for intelligent learning environments - Adapting the Instruction. In: Proc. 3rd Int. Conf. Comput. Support. Educ., pp. 456–459 (2011)
16. Hernández, Y., Pérez-Ramírez, M., Zatarain-Cabada, R., Barrón-Estrada,L., Alor-Hernández, G.: Designing empathetic animated agents for a b-learning training environment within the electrical domain. Spec. Issue Intell. Affect. Learn. Environ. New Trends Challenges Educ. Technol. Soc. 19(2) (2016)
17. Hernández, Y., Arroyo-Figueroa, G., Sucar, L.E.: A Model of Affect and Learning for Intelligent Tutors. Univers. Comput. Sci., 21(7), pp. 912-934 (2015)
18. Gaikwad, D.K., Mahender, C.N.: A Review Paper on Text Summarization. Int. J. Adv. Res. Comput. Commun. Eng. 5(3), pp. 154–160 (2016)
19. Kumar, Y.J., Goh, O.S., Basiron, H., Choon, N.H., Suppiah, P.C.: A review on automatic text summarization approaches. J. Comput. Sci. 12(4), pp. 178–190 (2016)
20. Vilca, G.C.V.: Generación automática de resúmenes abstractivos mono documento utilizando análisis semántico y del discurso. Pontificia Universidad Católica del Perú (2017)
21. Acero, I., Alcojor, M., Díaz, A., Gómez, J.M., Maña, M.: Generación automática de resúmenes personalizados. Proces. del Leng. Nat. 27(33), pp. 281–290 (2001)
22. Jeong, H., Ko, Y., Seo, J.: Efficient keyword extraction and text summarization for reading articles on smart phone., vol. 34, pp. 779–794 (2015)
23. Wong, K.F., Wu, M., Li, W.: Extractive Summarization Using Supervised and Semi-

supervised Learning. In: Proc. 22nd Int. Conf. Comput. Linguist. 1. Assoc. Comput. Linguist. 2008., no. August, pp. 985–992 (2008)
24. Fattah, M.A.: A hybrid machine learning model for multi-document summarization. Appl. Intell. 40(4), pp. 592–600 (2014)
25. Pal, A.R., Maiti, P.K., Saha, D.: An approach to automatic text summarization using simplified lesk algorithm and wordnet. Int. J. Control Theory Comput. Model. 3(4), pp. 15–23 (2013)
26. Atkinson, J., Munoz, R.: Rhetorics-based multi-document summarization. Expert Syst. Appl. 40(11), pp. 4346–4352 (2013)
27. Miranda-Jiménez, S. *et al.*: Conceptual Graphs as Framework for Summarizing Short Texts. Int. J. Concept. Struct. Smart Appl. 2(2), pp. 55–75 (2014)
28. Martinez-Romo, J., Araujo, L., Fernandez, A.D.: SemGraph : Extracting Keyphrases Following a Novel Semantic Graph-Based Approach. J. Assoc. Inf. Sci. Technol. 67(1), pp. 71–82 (2016)
29. Chatterjee, N., Sahoo, P.K.: Random Indexing and Modified Random Indexing based approach for extractive text summarization. Comput. Speech Lang. 29(1), pp. 32–44 (2015)
30. Lee, J.H., Park, S., Ahn,C.M., Kim, D.: Automatic generic document summarization based on non-negative matrix factorization. Inf. Process. Manag. 45(1), pp. 20–34 (2009)
31. Mei, J.P., Chen, L.: SumCR: A new subtopic-based extractive approach for text summarization. Knowl. Inf. Syst. 31(3), pp. 527–545 (2012)
32. Wang, M., Tang, X.: Extract summarization using concept-obtained and hybrid parallel genetic algorithm. In: Proc. - Int. Conf. Nat. Comput., no. Icnc, pp. 662–664 (2012)
33. Nagesh, K., Murty, M.N.: Obtaining single document summaries using latent dirichlet allocation. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 7666 LNCS, no. PART 4, pp. 66–74 (2012)
34. Kumar, C., Pingali, P., Varma, V.: Generating Personalized Summaries Using Publicly Available Web Documents. In: 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, pp. 103–106 (2008)
35. Díaz, A., Gervás, P.: User-model based personalized summarization. Inf. Process. Manag. 43(6), pp. 1715–1734 (2007)
36. Rao, P.R.K., Lalitha Devi, S.: Enhancing multi-document summarization using concepts. Sadhana - Acad. Proc. Eng. Sci. 43(2), pp. 1–11 (2018)
37. Miranda, S., Gelbukh, A., Sidorov, G.: Generación de resúmenes por medio de síntesis de grafos conceptuales. Rev. Signos 47(86), pp. 463–485 (2014)
38. Lin, C.Y.: Rouge: A package for automatic evaluation of summaries. In: Proc. Work. text Summ. branches out, no. 1, pp. 25–26 (2004)
39. Funderstanding: Nothing found for About_Learning Cfm. Engaging kids, learning theories. [Online]. Available: https://www.funderstanding.com/about_learning.cfm. [Accessed: 12-Sep-2018].
40. Kazu, I.Y.: The Effect of Learning Styles on Education and the Teaching Process. J. Soc. Sci. 5(2), pp. 85–94 (2009)
41. Alonso, C.M., Gallego, D.J., Honey, P.: Los estilos de aprendizaje. Procedimientos de diagnóstico y mejora. Procedimientos diagnóstico y Mejor. Bilbao Mensajero (1994)
42. Nenkova, A., McKeown, K.: Automatic Summarization. Found. Trends® Inf. Retr. 5(2), pp. 103–233 (2011)
43. Alonso, C.M., Gallego, J., Honey, P.: CHAEA. Cuestionario Honey-Alonso de Estilos de Aprendizaje. Universidad de Deusto. Instituto de Ciencias de la Educación (ICE), 2017. [Online]. Available: http://diged.usac.edu.gt/sfpu/cuestionario/chaea.

ISSN 1870-4069

Towards a Learning Ecosystem for Linemen Training

Guillermo Santamaría-Bonfil

CONACYT-INEEL, Cuernavaca, Morelos, Mexico
guillermo.santamaria@ineel.mx

Abstract. In this paper I present our ongoing work on developing a Learning Ecosystem for Training Linemen in Maintenance Maneuvers. First, challenges involved in training Linemen are introduced. Then, I discuss opportunities for creating a Learning Ecosystem for Linemen training using the Experience API standard and Learning Analytics. Although presented experimental results are reduced, these already show the value of Learning Analytics in exploiting data from already adopted technologies and new educational data sources for enhancing Linemen training.

Keywords: experience API, learning analytics, Learning ecosystem, learning manangement systems, self reports.

1 Introduction

Electric utilities are companies in charge of electricity generation, distribution, and power infrastructure maintenance. The latter is carried out by highly skilled workers called *Linemen*. Due to safety, technical, and business reasons the Linemen Maintenance Training (lMT) is mandatory [14]. The former involves physical preparation, knowledge on physics and electricity theory, theoretical and physical knowledge of maintenance maneuvers execution, and hygiene and industrial rules [3]. The adoption of cutting-edge technologies for the lMT is necessary to increase its effectiveness while avoiding risks.

The objective of this proposal is to pose a Learning Ecosystem (LE) for lMT to improve trainees' maintenance maneuvers apprenticeship. Loosely speaking, this will combine multiple educational technologies to exploit data generated by these to enhance learning within the environment systems [8]. For such endeavor I propose to employ systems based on the Experience API (xAPI) standard. The latter will allow formal and informal distributed learning experiences to be standarized and collected into a single repository [9]. Then, learners' traces collections will be wrangled, analyzed, and modelled using Learning Analytics (LA) to enhance the lMT [11]. In this work we detail how LA will be used for building visualization tools that enable educational stakeholders, specially tutors and course designers, to obtain valuable information of the overall educational process.

In section 2, I present the related work around lMT using intelligent learning environments and new perspectives. In section 3, the proposed Learning Ecosystem is shown. Section 4, discusses the experimental results for LMS and xAPI data using LA. Finally, conclusions, and future research is discussed.

2 Related Work

The lMT has used several technologies, the most used have been non-immersive Virtual Reality Training Systems (VRTS). VRTS have shown to be sucessful in lMT by reducing linemen accidents [13]. Another technological venue in utilities staff training is the usage of LMS along with Shareble Content Objects Repositories [1]. Moreover, ITS have been proposed to support and improve the learning processes within the electric domain by employing blended learning, affective estimation, and open learner models [6]. In these cases, data logs generated from users activities and its exploitation in favor of personalized education has been oversighted.

The xAPI standard allows users learning experience data be collected from heterogeneous learning systems by traducing these into *Activity Statements* (AS). AS are validated and stored in a Learning Record Store (LRS) database to exploit data for increasing training effectiveness [9, 8]. Data integration from educational systems under the xAPI umbrella, and its exploitation for tailoring personalized instruction conforms a Learning Ecosystem [8]. Some recent LE examples are a generic framework based on xAPI and GIFT [8], a Live Fire Training LE proposed by the U.S. Army [4], Transmedia Learning [12], and a xAPI-based framework for collecting and monitoring Self Regulated Learning [10]. In particular, the last two references rely on the usage of Self Reports (SR) for self monitoring, tutor monitoring, measure participants attitudes towards training, and so on. An un explored venue, is the usage of SR altogether with the xAPI standard for gather and exploit emotional self reports such as the Positive And Negative Affect Schedule (PANAS) or the Discrete Emotions Questionnaire [5].

Learning Analytics is consider a multidisciplinary paradigm for manipulating, modelling, and visualizing data from different educational sources to address: learners behaviors and performance, measuring social impact in learning, students' performance prediction, emotional states assessing, identifying student's learning strategies, provide decision making tools for educational stakeholders, and so on [11, 10]. Fig. 1 shows the general data science process, the same framework employed by LA. It starts with a real-world problem, data is then collected and manipulated to conform a data set suitable for Machine Learning (ML) modelling. In parallel, an Exploratory Data Analysis (EDA) is performed for visualizing data beyond formal modelling or hypothesis testing. Afterwards, findings are communicated and validated by educational stakeholders. This part is critical in LA since it allows stakeholders ponder ML models usage in decision making. In accordance to the former, data-based products are built (e.g. model for predicting students performance) and launched, then, the process iterates into new products or into refinement of previous ones.

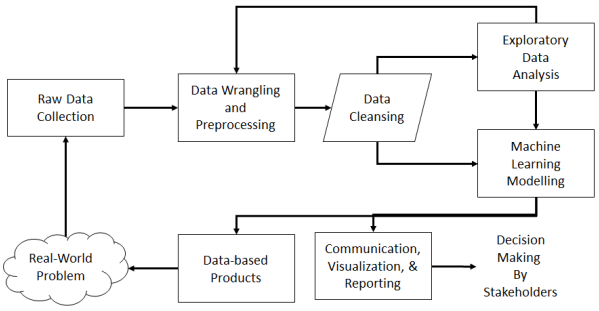


Fig. 1. The Data Science Processed used for performing Learning Analytics.

3 Proposed Learning Ecosystem

The LE proposed for lMT is conformed by an LMS, Traditional Training, a VRTS, and SR. In particular, affective trainees' assessment is done through SR. For each case, students learning experiences are mapped into AS, and collected into a LRS. LA is used for performing knowledge discovery on trainees' trace data, build ITS-inspired components, and provide insight to educational stakeholders. The proposed LE is shown in Fig. 2a.

3.1 xAPI Activity Statements

Any xAPI AS is syntactical similar to English. In its simplest form, AS are composed by an **Actor**, a **Verb**, and a **Object**. The *Actor*, corresponds to a unique id associated to a specific subject (e.g. Lola the trainee). The *Verb*, such as in any language, classifies an actor's activity using a unique internationalized resource identifier (IRI). *Objects* can be of several types and must contain an unique id property for unambiguous identification [9]. For example, a trainee assigned to execute step 1 from maneuver 1 in her lMT may have the following statement generated:

Lola the trainee (*actor*) **executed** (*verb*) **step 1-maneuver 1** (*object*).

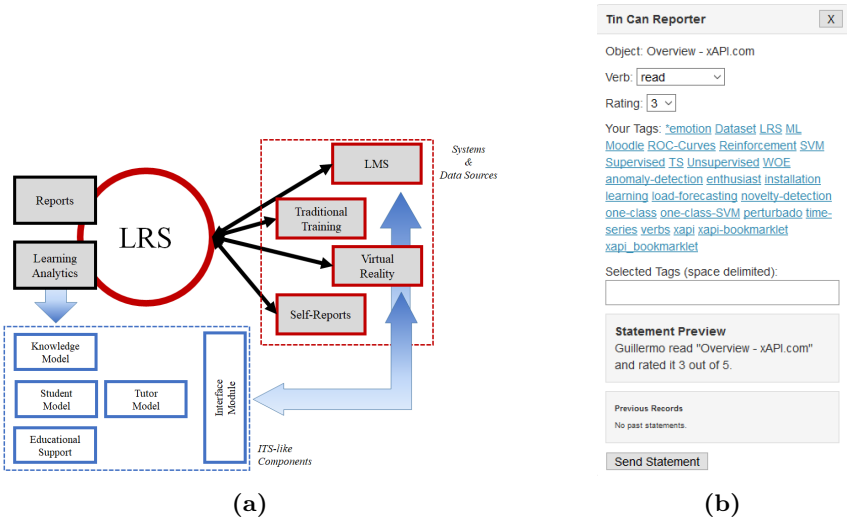
3.2 Data Sources and Technological Requirements

VRTS and Traditional Training are already used by utilities to carry out lMT and other staff training [13]. Traditional training is stored from text documents to adhoc computational systems. Maintenance maneuvers execution are practiced and evaluated using a VRTS, data generated by the former is stored in a local relational database. Consequently, first data from these sources needs to be gathered and mapped into its xAPI AS form, for its latter ingestion into an LRS. On the other side, LMS are currently not been employed for lMT, thus, an LMS is required. LMS is a type of technology vastly explored for educational purposes, thus, we do not detail it. However, it is worth mentioning that, a great

advantage of LMS is that several platforms already generate their internal logs using the xAPI AS format. Thus, data from such systems can be immediately reported into an LRS.

Self Reports

For this endeavor a bookmarklet, a small software stored as a bookmark in a web browser, can be employed. If the SR bookmarklet complies with the xAPI activity reporting standard, it will allow to capture basic learning experiences, and even personalized AS using non-basic fields provided by the xAPI standard. Since these already comply with the xAPI activity report format, learning experience can be immediately stored into the LRS. An example of a SR bookmarklet is shown in Fig. 2b.



(a) **(b)**

Fig. 2. Proposed Learning Ecosystem. On (a) a diagram of the LE, on (b) an example of a SR software.

Emotional SR requires the usage of the *Extension* xAPI field. In short, this field allows to define new attributes using a combination of a key (or property) and a value using the URI scheme. This is used within other xAPI fields such as the *Object*. Thus, using the extension field, we can employ the Discrete Emotions Questionnaire (DEQ) [5] for SR. DEQ distinguishes eight state emotions, i.e. anger, disgust, fear, anxiety, sadness, happiness, relaxation, and desire. Thus, once the DEQ is manipulated to fit into xAPI statements, any student may report its emotional state using the SR bookmarklet. For instance, using the previous xAPI example, if *Lola the trainee* felt *happiness* while executing step 1 from maneuver 1 in her lMT, may report her affective state generating a statement such as:

Lola executed step 1-maneuver 1 $feeling$ **happiness**.

Learning Record Store
In accordance to its core specification, an LRS is a cloud-based system which is in charge of storing and retrieving learning data exclusively formatted as xAPI statements [2]. In its most basic setup it only provides functions for store and retrieve xAPI statements. However, several LRS providers may also include tools and dashboards to visualize, combine, and manipulate AS data.

3.3 Learning Analytics

LA is a multidisciplinary paradigm, among the disciplines employed by it stands Text Mining (TM), Natural Language Processing (NLP), Web Data Mining, Exploratory Data Analysis, Machine Learning, among others [11]. In particular, for EDA we can employ *Word Clouds* (WC). This are visual text data representations which depict words importance (e.g. frequency) within a document(s) using several schemas e.g. larger words represents most frequent words whereas less important words have smaller font sizes. WC are used to quickly perceive the most (relative) important terms. Thus, by using EDA methods, we expect educational stakeholders can distill valuable information for adapting systems, course design, or identify profile characteristics of students at-risk [13].

Machine Learning can be used to build unsupervised and supervised models for the purpose of personalized education. On one hand, employing utilities documents along TDM and clustering algorithms we can build a knowledge structure that can be used as a coarse domain model [7]. On the other hand, we can relate the proficiency scores of each trainee to the errors committed using a Bag of Errors scheme [13].

4 Preliminary Experimentation

In the following I present the initial experiments towards a LE for lMT. The experimental configuration and software used are detailed in the following. Then, LMS and xAPI SR data are visualized. The corresponding LA experimentations for domain and student modelling have already been carried in [7] and in [13], respectively.

4.1 Experimental Setup

Due to a lack of an LMS and SR for lMT a proxy is used. Data from students enrolled in a *Machine Learning course*[1], is used. The course was carried out from 01-29-2018 to 06-01-2018 with 9 posgraduate students (master and PhD) which generated 2527 activity records. Activities and emotional SR were generated for the same course. To incentive students to SR, activities and emotional reporting

[1] https://www2.ineel.mx/posgrado/maestrias/mce.html

accounted for 10% each of the final course grade. SR were generated using the following instructions:

- *Activity:* during a learning session related to the course, if a student found a material which she considered important, useless, clear or confusing, that *object* would be reported using the SR tool.
- *Emotion:* at the start of a learning session related to the course, a student reported her emotional state by employing one of the emotions defined by the DEQ. Table 6 of [5] was used to clarify each emotional state.

SR were generated by Rustici Software LLC bookmarklet [2]. This allows to capture 4 basic learning experiences (i.e. experience, read, bookmark, and tweet), rate an Object, and personalized AS using tags (e.g. extension field). It is shown on Fig. 2b. Using it 420 AS were collected: 208 correspond to activities whereas 212 correspond to emotional SR. For the LMS Moodle^{TM} was used whereas the Watershed^{TM} LRS was selected. All LA experimentation was carried out using R and RStudio^{TM}.

4.2 Results

In the first instance, data from the LMS is visually explored using a standard bar plot shown in Fig. 3a. This presents on the x-axis the frequency of items visited by the students in the LMS. On the y-axis the different items in the course ordered by the total frequency of visits. Colors depict *Actors*. We can appreciate in the figure that most students have the same distribution of visits per item. Also, the top visited items corresponds to *homework* and *course project* tasks, whereas the least visited items corresponds to the *announcements* forum and *archives* used as examples for several course topics.

In the second instance we analyze SR using word clouds. Activities and DEQ SR are shown in Fig. 3b and Fig. 3c, respectively. Results are thrilling, for instance, observe on of Fig. 3b the word cloud generated by websites reported by students in the aforementioned course. As expected, the most visited website is youtube which offers a vast amount of ML courses and related content. This is followed by the course LMS, and the MOOC *DataCamp*. Other appreciable websites are *Google Scholar* and *Wikipedia*. Among the other least visited sites are machine learning blogs (e.g. *Data Science Central*), and scientific search engines such as *ScienceDirect, Springer* and *IEEE Xplore Digital Library*. In regard to the word cloud generated for DEQ SR, we can observe that the most frequent emotions reported were *desire* and *anxiety*. In accordance to [5], *desire* is a pregoal, positive affect which presumably promotes reward acquisition. On the other side, *anxiety* is regarded as a negative, high arousal emotion associated which *fear*, which is evoked by vague, potential threats [5]. Other important emotions reported were *relaxation* and *enthusiasm*, both considered positive affect which presumably assists in promoting reward enjoyment. The least popular emotions were *anger* and *disgust*. These results seems that, students were eager

[2] https://xapi.com/bookmarklet/

to learn the concepts of the course, and, although were anxious about the results they felt fulfilled after accomplishing tasks.

Fig. 3. Self-Reporting using the xAPI Bookmarklet. On (a) the Rustici xAPI bookmarklet is presented. Word clouds correspond to (b) Objects reported and (c) DEQ Emotions.

5 Conclusions and Future Work

In this work I presented a proposal for building a LE for lMT. This goes towards generating a standardized system for storing lMT learning experiences data from lMT systems, and its exploitation for adaptive learning. To date I have focused upon the tools for building the LE and some LA applications. Results for SR and LA are encouraging, thus, the next steps involve establishing a better experiment design for SR and emotional SR, establish contact with utilities managers to construct lMT courses within the proposed LE infrastructure, and use LA to exploit trainees trace and AS data for personalized instruction.

Acknowledgments. I thank the Cátedras CONACYT program and the research network RedICA for supporting this research.

References

1. Argotte, L., Arroyo-Figueroa, G., Noguez, J.: SI-APRENDE: An Intelligent Learning System Based on SCORM Learning Objects for Training Power Systems

Operators. In: Dev. Concepts Appl. Intell., vol. 363, pp. 33–38. Springer-Verlag Berlin Heidelberg (2011)
2. Berking, P.: Choosing a learning record store (LRS). Tech. rep., Advanced Distributed Learning Initiative (2016), http://adlnet.gov/adl-assets/uploads/2015/11/Choosing-an-LRS.pdf, accessed: 2018-09-29
3. Caha, V., Miletic, Z., Peric, S., Raljevic, D.: Live work training centres cooperation and development. In: 2017 12Th International Conference on Live Maintenance (ICOLIM). pp. 1–3 (2017)
4. Durlach, P., Washburn, N., Regan, D.: Putting Live Firing Range Data to Work Using the xAPI. In: Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2015. pp. 1–11. No. 15019 (2015)
5. Harmon-Jones, C., Bastian, B., Harmon-Jones, E.: The Discrete Emotions Questionnaire: A New Tool for Measuring State Self-Reported Emotions. PLOS ONE 11(8), e0159915 (aug 2016)
6. Hernández, Y., Pérez-Ramírez, M., Zatarain-Cabada, R., Barrón-Estrada, L., Alor-Hernández, G.: Designing empathetic animated agents for a B-learning training environment within the electrical domain. Educational Technology and Society 19(2), 116–131 (2016)
7. Hernández, Y., Santamaría-Bonfil, G., Pecero, V.: Text Mining for Domain Structure Analysis in a Training System for Electrical Procedures. In: 10th Workshop on Intelligent Learning Environments, WILE 2017. pp. 1–7 (2017)
8. Hruska, M., Medford, A., Murphy, J.: Learning ecosystems using the generalized intelligent framework for tutoring (GIFT) and the experience API (xAPI). In: Proceedings of the 17th International Conference on Artificial Intelligent in Education. vol. 1432, pp. 9–16 (2015)
9. Kevan, J.M., Ryan, P.R.: Experience API: Flexible, Decentralized and Activity-Centric Data Collection. Technol. Knowl. Learn. 21(1), 143–149 (apr 2016)
10. Manso-Vazquez, M., Caeiro-Rodriguez, M., Llamas-Nistal, M.: An xAPI Application Profile to Monitor Self-Regulated Learning Strategies. IEEE Access 6, 42467–42481 (2018)
11. Peña-Ayala, A.: Learning analytics: A glance of evolution, status, and trends according to a proposed taxonomy. Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 8(3), 1–29 (2018)
12. Raybourn, E.: A new paradigm for serious games: Transmedia learning for more effective training and education. Journal of Computational Science 5(3), 471–481 (2014)
13. Santamaría-Bonfil, G., Hernández, Y., Pérez-Ramírez, M., Arroyo-Figueroa, G.: Bag of Errors : Automatic inference of a student model in an Electrical Training System. In: MICAI 2017: Mexican International Conference on Artificial Intelligence. pp. 1–15 (2017)
14. Vanrobayes, B., Roussel, F.: Live work training and skills development at RTE. In: 2017 12th International Conference on Live Maintenance (ICOLIM). pp. 1–3. IEEE (apr 2017)

ISSN 1870-4069

Using CTF Tournament for Reinforcing Learned Skills in Cybersecurity Course

Hugo Gonzalez, Rafael Llamas, Omar Montaño

Universidad Politécnica de San Luis Potosí,
Academia de Tecnologías de la Información y Telemática,
San Luis Potosí, SLP, Mexico
hugo.gonzalez,rafael.llamas,omar.montano @upslp.edu.mx

Abstract. In this paper, we present an experience on using CTF tournament as a gamification process to reinforce learned knowledge and skills on cybersecurity course. It is important to highlight that part of those skills learned were defined by the students during the course. This strategy was used to improve the students' engagement, at the same time to fulfill as much as possible the expectations for the class. The results of the tournament and final surveys at class showed that it was a good experience for the majority of the participants. The results showed that it helped to reinforce the learned skills, and applying them to specific challenges. The students felt motivated and productive when they were able to solve a challenge.

Keywords: CTF tournament, learning skills, cybersecurity.

1 Introduction

Learning through practice is a common philosophy for some teachers, including the authors of this work. Some courses in different fields in computer science include well-established practices and activities. In those activities, the student is guided to achieve some very specific goal related to the course. Sometimes there is a gamification component if the instructor set a realistic game to test the students in a real-world case scenario in a controlled environment. However, this is not always the case.

In cybersecurity field, there had been several studies talking about setting up capture the flag (CTF) games as part of the course [4], or use it to teach basic cybersecurity skills and generate engagement on the topic for high school students [14]. Other scholars discussed the participation in national wide CTF tournaments to improve skills of students [8]. However, in this work we propose a slightly different approach: Using an extra class CTF tournament to reinforce knowledge and skills learned during a cybersecurity course.

There are mainly two types of CTF, one focus on attacks and defense, and another type that employs a jeopardy style. Jeopardy style CTF are developed as a set of challenges in different domains as such as network analysis, web hacking, forensics analysis, reverse engineering, cryptography. For each domain, a series of

challenges are designed with increasing complexity than the previous. The points awarded for a solved challenge reflects the complexity of that challenge. Easy challenges will motivate participants on taken next levels of the competition.

The remainder of this paper is organized as follows: In Section 2 we describe our experience in running a CTF tournament. In Section 3 we present the results, Section 4 briefly present related work. Finally Section 5 concludes this work.

2 Our Experience

At the Polytechnic University of San Luis Potosi, we offer one shared course about Information Security for two of our programs: Information Technologies and Telematics. Previous experiences showed us that the technical skills, knowledge, and interest about cybersecurity could be very different among the students of a class, from the one not really interested in the topic to the one that expends nights learning and practicing cybersecurity topics by herself. Those differences create an environment where engagement can be low and expectations can be very high from different students.

To address that matter, we decide to improve the engagement of the students and fulfill the expectations of the class as much as possible. As the use of gamification had been proposed before [9], we focus our efforts in a CTF tournament along the flexibility that we have to adapt locally the syllabus of the course. Previous effort on a CTF tournament was poorly received by the students, with only one team registered. This time we took a different approach, the students were involved in define part of the current course syllabus with the objective to increase the engagement in the class and to improve the response on the tournament.

Adaptation process of the syllabus Our course content is organized in three parts: The **first part** is an introduction to cybersecurity, it includes a cybersecurity path as a professional. The **second part** is about information security management systems (ISMS) and their impact at organizations. The **third part** is about advanced topics such as Ethical hacking, and Forensics Analysis.

During spring 2018 term we implemented "research, propose and vote" methodology, a democratic activity to add other topics of interest to the content of the course. At the end of the second part of the course, we asked the students to read the news and search for interesting topics in the cybersecurity field. Then we ask the students to propose one or two topics to be included in the list of options. Finally, we promoted a voting exercise, where each student cast three votes for the topics listed that should be covered in the course. The objective was to increase the engagement of the students and fulfill their expectations on the course. At the same time, we want to make them feel that they were helping to shape the course content on their own interest. Some advanced topics such as low-level attacks were proposed during this activity, surprising the instructors. We combined some of the topics to cover most of the expectations for students with advanced skills.

Part of the activity is shown in Figure 1, after the students proposed the topics and voted for their choices. From the list of options, five topics with the major votes were selected to prepare five laboratory practices to perform during the Third part of the curse. It is important to highlight that the instructors tried to minimize their influence on the selected topics, only discarding non-related topics. However, it was expected to have at least one laboratory practice on forensics and other related to ethical hacking.

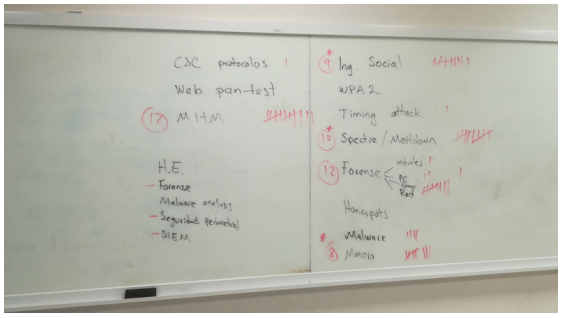


Fig. 1. Cybersecurity topics proposed by the class and voting results.

The selected five topics and their objectives are:

1. Man in the Middle Attack. Perform an attack on a mobile device to capture credentials and other interesting info.
2. Network forensic analysis. Given a file with a network traffic capture, use tools to answer 5 questions about the activities captured.
3. Social engineering. Watch two videos about social engineering attacks and develop a script to apply social engineering to the caller next time they get a phone call that they are not expecting.
4. Spectre and Meltdown attacks. Read an article about these new attacks and execute a proof of concept code in different computers to observe the results.
5. Analysis of a monero miner. Perform a light and superficial analysis of a malware known to mine Monero cryptocurrency to identify the wallet used by the attacker.

CTF design The objective is to create a CTF tournament that includes most of the topics from the adapted syllabus as challenges, including other topics of interest. It is important to note that this tournament should be open to all students at the University during the Information Technology Engineering academic event. But with the original purpose in mind, the students from Information Security courses should receive a special invitation to participate in the tournament to reinforce their knowledge and learned skills. The CTF tournament software selected is Mellivora [15], which is a Jeopardy-style management system

and scoreboard. The challenges should be designed and prepared independently from the platform. Eight categories were created with two to four challenges each. Three of these categories are directly related to the practices in the Third part of the course, four are related to other general topics such as programming. The category *Firmware* is beyond the material covered in the course, this is a challenge to discover and learn new tools and methods under the pressure of the tournament with limited time. Categories, challenges, points available and percentage of teams that solved them are shown in Table 1.

The name of each category intents to be self-descriptive, the categories and challenges that were not solved during the tournament are commented below. Note that the full description of the challenges and the challenges themselves are available upon request to the authors. Help and guidance to create or organize CTF are also offered by the authors.

Firmware category included two challenges, each of them is a firmware for a different device with a modification inside that includes a flag. One modification is a web shell in the admin page. The other firmware contains an extra file with the flag in it. **Programming** category is related to terminal interaction, a student needs to interact with a program 100 times before it returns the flag. These challenges were related to **network programming** category, but instead of sockets, one needs pipes or other methods to interact with other terminal programs in Linux. Both of the challenges include a timer to avoid students to solve them manually. Interacting with the terminal is not a common topic, and the students focus on the rest of the challenges.

The categories that were completely solved are: **Forensics** and **Network Forensics**. The students showed great performance in this categories as they had some similar exercises at class. Also, extra credits for the Information Security course were offered to the team that solved these challenges.

CTF tournament As mention before, the tournament was held in conjunction with an academic event with talks and tutorials some weeks before the end of the term. The tournament lasted two days, access to the challenges were only possible in a classroom assigned to this activity. Five teams of four students and one team of three students were registered. Three teams included at least one female student. Only one registered team was not taking the Information Security course at that term but already took the course before.

Four pictures from the scoreboard and students are showed int Figures 2,3,4 and 5. *ch0co* team won the tournament with 1050 points. **Firmware** category that includes modified firmware for IoT devices were not solved by any team, as this challenge was out of the scope from the course. However, students made comments about these challenges: " [the challenges] were fun to try". The **Web** category was also out of the scope from the course, but most of the students had previous experience with web systems and completed this set of challenges that they considered "not that complicated".

The excitement showed by team members when they capture a flag was contagious and invigorating. Flag after flag they were demonstrating their acquired knowledge during the Information Security course, reinforcing their skills

Table 1. CTF categories and challenges.

Category	Challenges	Points	solved rate
Firmware	Openwrt compromised image	90	0%
	Dlink firmware compromised	110	0%
Forensics	Evidence file	40	100%
	A different flag in the system	90	100%
MITM	Go http test	80	0%
	APK hidden flag	100	50%
Network Forensics	Insecure communications	30	100%
	Binary on the network	60	100%
	Weird binary on the network	100	50%
Network programming	Operations over the network	60	17%
	Something like echo	70	17%
	Strings manipulation over the network	90	17%
Programming	Console interactions	50	0%
	Random strings	70	0%
	Calculate the result	90	0%
Reversing	Go go parameters	70	33%
	Obfuscated Javascript	70	50%
	DotNet mistake	100	50%
Web	Hidden flag	20	83%
	Find the flaw	60	66%
	Enjoy the milk	80	83%
	Official client	80	50%

by playing the game. Some students stated that they learned new skills and abilities during the tournament.

3 Results

The results obtained from this experience are two-fold. First, the students felt engaged and motivated on the class after the "research, propose and vote" process. Second, the students that participated in the CTF tournament described a great experience from the game, solving challenges and applying previous and new skills. The impact of the first result was measured using a survey. In that survey, we asked questions related to the feelings and perception from the students about the process to add topics to the syllabus' course. Also with the engagement and feelings toward the course, expectations and the topics covered. The results of the survey are summarized as follows:

1. The students felt that they are contributing to decide their path on the learning process. They would like to have more courses with a flexible curriculum where they can propose new and interesting topics.
2. The students reported that they were more engaged and motivated in the class, most of them felt that their expectations about the course contents were fulfilled in the class.

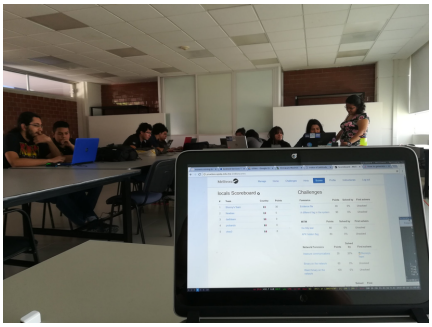


Fig. 2. First flag captured.

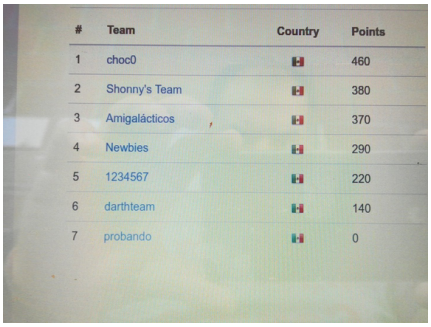


Fig. 3. Scores at the end of the first day.

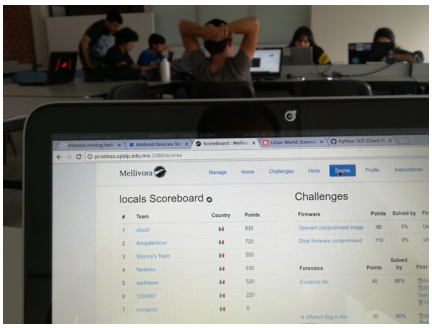


Fig. 4. Working hard on the second day.

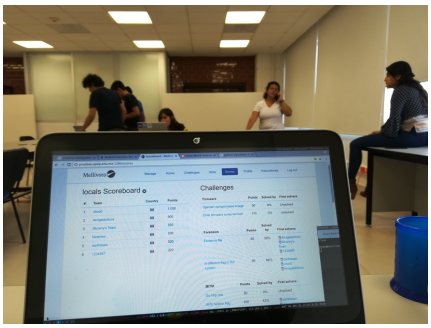


Fig. 5. Final scores.

3. Majority of the students felt well working on the practices. Only two students expressed they were not comfortable with the practices because they found them very technical, and found themselves lacking skills to complete the exercises.

In a blank space to express themselves, some students commented in the survey that they really enjoyed the CTF tournament and that it should be offered every year.

In relation to the CTF tournament, we collected the following information:

More than 60% of students taking Information Security courses registered for the CTF tournament. They expend over 20 hours in the classroom working on the challenges. We consider that the engagement of the students in cybersecurity activities is growing in our university, we believe that this is in part because the gamification strategy introduced in the course. A few students expressed the intention to follow cybersecurity careers.

The overall results of the tournament showed that all teams solved basic challenges on **Forensics**, reinforcing the skills learned at class. Furthermore,

some challenges out of the topics from the class in **Web** category were also solved, showing that students can apply their skills to different problems.

The winning team of the tournament solved 65.2% of all the challenges during the two days competition. Their comments were related to open the game to play outside of the University facilities because they could not play at night from home. That team was suggested to play in the HackDef 2018 CTF pre-qualification tournament [10] that will be held in August.

As feature work, we already started to plan the next CTF game with the format proposed by Chothia et al. [4, 5]. The game will be deployed as an independent virtual machine and will employ an intelligent agent to interact with the students while they are solving the challenges to obtain the flags. The challenges will be related to the topics selected by the class in the term in the course.

4 Related Work

Deterding et al. [7] define "gamification" as the use of game design elements in non-game contexts. Among other tasks gamification is used to engage users and help with the learning process [11]. Recently Li and Kulkarni [13] concluded that gamification is a very effective way of learning.

Gondree et al. [8] discuss the cybersecurity competitions and games and how it is necessary to adopt a common vocabulary to express the game's goals and characteristics. They also discuss about competitions like iCTF, DC3 Forensic Challenge, CyberPatriot, CCDC, PlaidCTF, CSAW CTF where training and education are their main role. At least three frameworks to deploy Capture the Flag contests are freely available as open source software [16, 12, 1].

Chothia et al. had been working in innovation and development of improved course materials. In 2015 the authors presented an offline CTF system which includes 5 learning activities [4]. The students can download the system and play in a controlled environment. In 2016 the authors developed a new course on penetration testing using IoT devices, this course was thought at Birmingham University with great engagement and response from students [6]. In 2017 the authors included a storytelling, intelligent component on a virtual machine to teach a course on information security in 11 weeks [3]; students chose their own adventure in the game. In 2018 the authors focused on spear-phishing, challenging the students to produce realistic spear-phishing attempt where a rules engine will decide if it was successful [5].

Chain et al. offered an innovative platform to capture the flag based on cloud offense and defense. After design and implement their platform, they use a survey to evaluate the pertinence of the exercises, time and difficulties of the tasks. [2].

5 Conclusions

Designing and developing the challenges was a fun exercise for the instructors, and playing them was a great activity for the students to reinforce their

skills. The results from the tournament showed that students were learning new skills that they can apply in real-world scenarios. Gamification in the classroom through CTF tournaments should be implemented and practiced in cybersecurity-related courses. As part of the course or as an extra activity. It is important to mention that all the participant students comment that it was fun, challenging and that they learned and enjoyed the extra class activity.

References

1. Boesen, S., Weiss, R., Sullivan, J., Locasto, M.E., Mache, J., Nilsen, E.: Edurange: Meeting the pedagogical challenges of student participation in cybertraining environments. In: Proceedings of the 7th USENIX Conference on Cyber Security Experimentation and Test. pp. 9–9. CSET'14, USENIX Association, Berkeley, CA, USA (2014), http://dl.acm.org/citation.cfm?id=2671214.2671223
2. Chain, K., Kuo, C., Liu, I., Li, J., Yang, C.: Design and implement of capture the flag based on cloud offense and defense platform. In: 2018 IEEE International Conference on Applied System Invention (ICASI). pp. 686–689 (April 2018)
3. Chothia, T., Holdcroft, S., Radu, A.I., Thomas, R.J.: Jail, hero or drug lord? turning a cyber security course into an 11 week choose your own adventure story. In: 2017 USENIX Workshop on Advances in Security Education (ASE17). USENIX Association (2017)
4. Chothia, T., Novakovic, C.: An offline capture the flag-style virtual machine and an assessment of its value for cybersecurity education. 2015 USENIX Summit on Gaming, Games, and Gamification in Security Education (3GSE 15) (2015)
5. Chothia, T., Paiu, S.I., Oultram, M.: Phishing attacks: Learning by doing. In: 2018 USENIX Workshop on Advances in Security Education (ASE 18). USENIX Association (2018)
6. Chothia, T., Ruiter, J.d.: Learning from others' mistakes: Penetrating testing iot devices in the classroom (2016)
7. Deterding, S., Dixon, D., Khaled, R., Nacke, L.: From game design elements to gamefulness: Defining "gamification". In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. pp. 9–15. MindTrek '11, ACM, New York, NY, USA (2011), http://doi.acm.org/10.1145/2181037.2181040
8. Gondree, M., Peterson, Z.N., Pusey, P.: Talking about talking about cybersecurity games. ;login: USENIX magazine 41(1), 36 – 40 (2016)
9. Gonzalez, H., Llamas, R., Ordaz, F.: Cybersecurity teaching through gamification: Aligning training resources to our syllabus. Research in Computing Science 146, 35–43 (2017)
10. Hacker Defender Academy: Hackdef ctf 2018. http://www.hackdef.net, \url{http://www.hackdef.net}
11. Hamari, J., Shernoff, D.J., Rowe, E., Coller, B., Asbell-Clarke, J., Edwards, T.: Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. Computers in Human Behavior 54, 170 – 179 (2016), http://www.sciencedirect.com/science/article/pii/S074756321530056X
12. Joe Moloch: root-the-box framework. https://github.com/moloch--/RootTheBox/, \url{https://github.com/moloch--/RootTheBox/}

13. Li, C., Kulkarni, R.: Cybersecurity education through gamification. American Society for Engineering Education 123th Annual conference and exposition (2016)
14. McDaniel, L., Talvi, E., Hay, B.: Capture the flag as cyber security introduction. In: System Sciences (HICSS), 2016 49th Hawaii International Conference on. pp. 5479–5486. IEEE (2016)
15. Nakiami: Mellivora is a ctf engine written in php. `https://github.com/Nakiami/mellivora`, `\url{https://github.com/Nakiami/mellivora}`
16. The Computer Security Group at UC Santa Barbara: ictf framework. `https://github.com/ucsb-seclab/ictf-framework`, `\url{https://github.com/ucsb-seclab/ictf-framework}`

Electronic publication
Available online: http://www.rcs.cic.ipn.mx