

UPCommons

Portal del coneixement obert de la UPC

<http://upcommons.upc.edu/e-prints>

Aquesta és una còpia de la versió *author's final draft* d'un article publicat a la revista *Soft computing*.

La publicació final està disponible a Springer a través de <http://dx.doi.org/10.1007/s00500-016-2399-0>

This is a copy of the author 's final draft version of an article published in the journal *Soft computing*.

The final publication is available at Springer via <http://dx.doi.org/10.1007/s00500-016-2399-0>

Article publicat / Published article:

Arguedas, M. [et al.] (2016) A model for providing emotion awareness and feedback using fuzzy logic in online learning. " Soft computing". Doi: 10.1007/s00500-016-2399-0

Soft Computing

A MODEL FOR PROVIDING EMOTION AWARENESS AND FEEDBACK USING FUZZY LOGIC IN ONLINE LEARNING

Full Title:	A MODEL FOR PROVIDING EMOTION AWARENESS AND FEEDBACK USING FUZZY LOGIC IN ONLINE LEARNING
Keywords:	Fuzzy Logic; Affective Learning; Students' Emotive States; (APT) Affective Pedagogical Tutor; Affective Feedback
Corresponding Author:	Marta Arguedas Universitat Oberta de Catalunya SPAIN
Corresponding Author's Institution:	Universitat Oberta de Catalunya
First Author:	Marta Arguedas
First Author Secondary Information:	
Order of Authors:	Marta Arguedas Fatos Xhafa Luis A. Casillas Thanasis Daradoumis Adriana Peña Santi Caballé
Funding Information:	
Abstract:	Monitoring users' emotive states and using that information for providing feedback and scaffolding is crucial. In the learning context, emotions can be used to increase students' attention as well as to improve memory and reasoning. In this context, tutors should be prepared to create affective learning situations and encourage collaborative knowledge construction as well as identify those students' feelings which hinder learning process. In this paper, we propose a novel approach to label affective behavior in educational discourse based on fuzzy logic, which enables a human or virtual tutor to capture students' emotions, make students aware of their own emotions, assess these emotions and provide appropriate affective feedback. To that end, we propose a fuzzy classifier that provides a priori qualitative assessment and fuzzy qualifiers bound to the amounts such as few, regular, and many assigned by an affective dictionary to every word. The advantage of the statistical approach is to reduce the classical pollution problem of training and analyzing the scenario using the same dataset. Our approach has been tested in a real online learning environment and proved to have a very positive influence on students' learning performance.
Section/Category:	Methodologies & Application

A MODEL FOR PROVIDING EMOTION AWARENESS AND FEEDBACK USING FUZZY LOGIC IN ONLINE LEARNING

Abstract Monitoring users' emotive states and using that information for providing feedback and scaffolding is crucial. In the learning context, emotions can be used to increase students' attention as well as to improve memory and reasoning. In this context, tutors should be prepared to create affective learning situations and encourage collaborative knowledge construction as well as identify those students' feelings which hinder learning process. In this paper, we propose a novel approach to label affective behavior in educational discourse based on fuzzy logic, which enables a human or virtual tutor to capture students' emotions, make students aware of their own emotions, assess these emotions and provide appropriate affective feedback. To that end, we propose a fuzzy classifier that provides a priori qualitative assessment and fuzzy qualifiers bound to the amounts such as few, regular, and many assigned by an affective dictionary to every word. The advantage of the statistical approach is to reduce the classical *pollution* problem of training and analyzing the scenario using the same dataset. Our approach has been tested in a real online learning environment and proved to have a very positive influence on students' learning performance.

Keywords Fuzzy Logic - Affective Learning - Students' Emotive States – (APT) Affective Pedagogical Tutor – Affective Feedback

1 Introduction and motivation

Our research work aims at investigating the effectiveness of an emotion labeling model to detect emotions in educational discourse (text and conversation) in a non-intrusive way making emotion awareness explicit both at individual and group level (Arguedas et al. 2014). Studies have shown that emotional experiences influence student's motivation, learning strategies and achievement whereas such emotional experiences are influenced by personality and classroom characteristics (Frasson & Chalfoun 2010; Goetz et al. 2006).

Given that people are able to express a wide range of emotions, which vary in intensity, duration, context, etc. during activity time, our model is based on dimensional categories of emotions (Scherer 2005). It also

makes use of affective dictionaries expressing the emotional weights of words as a function of affective dimensions (pleasure, arousal, etc.). Due to the nature of affective dimensions, which consists of aspects manifesting continuity, this means that dimensions show a wide range of high-precision alternatives; thus every dimension requires a previous treatment. Therefore, each dimension must be preprocessed to obtain fuzzy values corresponding to the magnitudes of each emotional dimension. Such fuzzy values are the qualitative awareness that can be bound to such continuity. Actually this is an analog to discrete conversion.

Once the conversion into qualitative emotional awareness is performed, discrete affective states as well as knowledge regarding emotional performance become available to both teachers and students. Diverse academic and business scenarios could be supported and improved by emotional awareness, which would allow performing different micro-adjustments to discourse or behaviors every aspect of computer-mediated human interaction. Traditional human interaction allows participants to gather a whole range of inputs to build a rich image from reality. However, human interaction supported by digital channels would imply a loss of information, even though sophisticated digital channels are used. Hence, every additional mechanisms aimed at providing supplementary information would entail a further understanding of the new functionalities they offer.

Indeed, among the opportunities to add supplementary information is the possibility to provide emotional awareness from text. The reader may consider a virtual scenario in which different people are interacting, such as a virtual chat room. These people are mainly using text messages and emoticons to communicate. It would be interesting to both participants and the chat monitor person/system to be able to have a mechanism to analyze chat text in order to discover the emotions underlying conversation as well as to produce specific emotional feedback to support and monitor participants' needs.

The aim of this study is to present an effective approach to label affective behavior in educational

discourse based on fuzzy logic, which will enable a human (or virtual) tutor to capture students' emotions, make students aware of their own emotions, assess these emotions and provide appropriate feedback to the involved actors. In order to address these challenges, this paper is organized as follows. In Section 2, we present a comprehensive analysis of the state of the art of approaches to modeling emotions, mainly in learning environments, focusing more on two important issues that concern affective dictionaries and fuzzy logic in Sentiment Analysis. Based on this analysis, in Section 3, we present the conceptual design of our model. Section 4 describes the context, the experiment and the data used for the analysis. In Section 5 we show how our model is implemented in a real case setting. Finally, in Section 6, we present the results obtained, whereas Section 7 provides a critical discussion of the results. Finally Section 8 concludes with the directions for future work.

2 Related work

Our goal is to classify and later label up the students' emotional state through the analysis of their educative discourse in a virtual learning environment. To that end, before starting the design of our model, we have extensively revised the literature related to three topics that concern its development.

First, we reviewed various existing models for the classification of the different emotional states. Secondly, we studied the different affective dictionaries that have been compiled so far, to identify how each one of them provides the use of information about the affective weights of words composing the educational discourse. Finally, we checked out the related works about the application of fuzzy logic in the Sentiment Analysis field

Emotion models

In Artificial Intelligence, affective computing is the branch of studies and development systems that can recognize, interpret, process, and simulate human affects, whereas its main goal is to simulate empathy. Regarding the approach proposed in this study, our aim is to make the machine capable of interpreting the emotional state of humans and adapting its behavior to them, giving an appropriate response to their emotions (Tao & Tan 2005).

There are two leading models describing how humans perceive and classify emotion, namely, dimensional and categorical models (Feidakis et al. 2013;

Daradoumis et al. 2013). Categorical models classify emotions into basic, secondary, tertiary, etc. (Ekman & Friesen 1971; Ortony et al. 1988; Pekrun 1992), while dimensional models specify gradual emotions as arousal, valence, control, intensity, duration, frequency of occurrence, etc. (Russell 1983; Kort & Reilly 2002; Scherer et al. 2013; Kuncheva 2000).

In the learning context, emotions can be used to increase student's attention as well as to improve memory and reasoning (Isen 2000). In this context, tutors must be prepared to create affective learning situations and encourage collaborative knowledge construction as well as identify those students' feelings which hinder learning process (Ibarrola 2000).

According to Kort & Reilly (2002) emotion measurement tools and techniques fall into three main categories:

-Psychological (subjective report using verbal or pictorial scales or questionnaires, etc.). For instance, the PAD (Pleasure-Arousal-Dominance) emotional state model is a psychological model developed in Mehrabian & O'Reilly (1980) to describe and measure emotional states. PAD uses three numerical dimensions to represent all emotions. Its initial use was in a theory of environmental psychology, the core idea being that physical environments influence people through their emotional impact. The PA part of PAD was developed into a circumplex model of emotion experience, and those two dimensions were termed "core affect". The D part of PAD was re-conceptualized as part of the appraisal process in an emotional episode (a cold cognitive assessment of the situation eliciting the emotion). A more fully developed version of this approach is termed the psychological construction theory of emotion. The PAD model has been used to study nonverbal communication such as body language in psychology (Mehrabian 1972). It has also been applied to the construction of animated characters that express emotions in virtual worlds (Becker et al. 2007).

-Physiological (use of sensors to capture biometric signals). For instance, Gonçalves et al. (2016) present a multimodal approach by using multiple sensors to collect and assess users' emotion at interaction time while interacting with a game.

- Behavioral (observation or capturing of motor-behavioral activity e.g., facial expressions, sentiment

analysis of text input, mouse and keyboard logs, etc.). Sentiment analysis tools have been used in this work to provide means for the identification of the attitude holder and the polarity of the attitude as well as for the description of the emotions and sentiments of the different actors involved in the text. Plutchik offers an integrative theory based on evolutionary principles (Plutchik 2001). Emotions are adaptive—in fact, they have a complexity born of a long evolutionary history—although we conceive emotions as feeling states. According to Plutchik (2001), the feeling state is part of a process involving both cognition and behavior and containing several feedback loops.

As mentioned before, our goal in this work is to develop tools that provide teachers with useful information about students' emotional state, so that they can assess these emotions, and eventually provide appropriate affective feedback to students. To this end, we choose a mixed model composed by three dimensions (Mehrabian & O'Reilly 1980) and eight emotional labels (Plutchik 2001).

Affective Dictionaries

In the Sentiment Analysis field, textual information includes, among others, subjective expressions that describe people's sentiments, appraisals or feelings toward entities, events and their properties (Liu 2012).

The majority of studies apply classification algorithms to obtain the contextual polarity and the frequent terms of a document, that is to say, are simply based on the appearance and frequency of terms or the text valence in order to determine if it is a positive or negative critique. (Li et al, 2015). A new trend in emotion research consists in performing lexical analysis of texts with the aim of identifying the words that can predict the affective states of the authors (Calvo & D'Mello 2010).

Within this area, some affective dictionaries have been developed and are widely used. These dictionaries provide a lexical repository in different languages. In particular, we have carried out a major review of affective dictionaries in Spanish language concerning both emotion models (dimensional and categorical).

SentiWordNet is a lexical resource for opinion mining (Baccianella et al. 2010). SentiWordNet assigns to each synset (i.e., synonyms which are grouped into unordered sets) of WordNet Affect (a linguistic resource

for a lexical representation of affective knowledge, (Strapparava & Valitutti 2004)) three sentiment scores: positivity, negativity and objectivity. The method used to develop SentiWordNet is an adaptation to synset classification of a previous method for deciding the PN-polarity (identifies whether a term that is a marker of opinionated content has a positive or a negative connotation) and SO-polarity (describes a given situation or event, without expressing a positive or a negative opinion on it) of terms (Esuli & Sebastiani 2006). The method relies on training a set of ternary classifiers, each of them capable of deciding whether a synset is Positive, Negative, or Objective. However, SentiWordNet is not available in Spanish.

The development of the framework Affective Norms for English Words (ANEW) (Bradley & Lang 1999) is an instrument for the dimensional perspective of emotions based on works of Wundt (1896) and Osgood et al. (1957). From this perspective, three basic dimensions are proposed, through which the entire range of human emotions can be organized: valence (which ranges from pleasant to unpleasant), arousal (which ranges from calm to excite) and dominance or control (ranging from in control to out of control). The ANEW list provides normative values in these dimensions for 1,034 words and there is a Spanish adaptation of the ANEW made by Redondo et al. (2007).

Whissell's Dictionary of Affect in Language, originally designed to quantify the Pleasantness and Activation of specifically emotional words, was revised to increase its applicability to samples of natural language. A third rated dimension (Imagery) was added, and normative scores were obtained for natural English. Evidence supports the reliability and validity of ratings. The revised Dictionary, which contains ratings for characteristic words of natural language, is a portable tool that can be applied in almost any situation involving language (Whissell 2009).

The NRC Emotion Lexicon (EmoLex) is a list of English words and their associations with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and two sentiments (negative and positive) (Mohammad & Turney 2013). The annotations were manually done by crowd sourcing. Despite some cultural differences, it has been shown that a majority of affective norms are stable across languages. Given that three basic dimensions, namely, valence,

arousal and dominance are commonly used by researchers, we decided to use the Spanish version of ANEW dictionary merged with the Spanish version of EmoLex. Our approach to manage word lists is inspired by the effort of Andrew et al. (2011), so we can collect affective indicators for words.

Fuzzy Logic Applied to Sentiment Analysis

Regarding fuzzy logic supporting sentiment analysis, several studies have been carried out, however not all in virtual learning. For instance, Moharrer et al. (2015) proposed a novel two-phase methodology based on interval type-2 fuzzy sets (T2FSs) to model the human perceptions of the linguistic terms used to describe the online services satisfaction. The analysis is carried out by using well-established metrics and results from the Social Sciences context.

Andreevskaia & Bergler (2006) focus on the WordNet dictionary so that they can acquire some awareness from text mining based on the sentiment analysis approach. Ali et al. (2016) focused on sentiment analysis in the academic field and were able to capture average behaviors shown by words, based on regular statistics analysis. They managed to build a fuzzy logic scheme, aimed at producing a qualitative description for words, turning quantitative magnitudes into literal terms bound to qualitative perceptions, such as good, bad, etc. Another application field for sentiment analysis is the stock market, where one can build trading scenarios based on opinion mining from information extracted from technical analysis and trends shown by stock trading systems and the market. This is discussed in Wu et al. (2014), where features, predictions and trading are dealt using intelligent support.

Our proposal aims to produce a scheme for word treatment, similar to Andreevskaia & Bergler (2006). Nevertheless, our study focuses on the ANEW dictionary and fuzzy qualifiers bound to amounts such as few, regular, and many. These qualifiers will be then crossed over throughout specific inference rules.

These inference rules can explain the amounts achieved by accumulating numeric data from indicators and express them in qualitative terms. Hence, high level emotions could be inferred from plain numbers. Inspired by Kuncheva (2000), our proposal is a fuzzy classifier, more precisely, a statistical classifier, which provides a priori qualitative assessment to the amounts assigned by

ANEW to every word. The advantage of the statistical approach is to reduce the classical *pollution* problem of training and analyzing the scenario using the same dataset. Affective dictionaries usually have a limited number of words. Our statistical classifier uses centrality and dispersion measures calculated from the ANEW analysis dimensions. These measures are used to build the fuzzy classifier, as explained later in this paper.

3 A Fuzzy Based Classification Model For Inferring Affective States

Fuzzy Sets and Fuzzy Rules

Before describing our conceptual fuzzy-based model for inferring emotional states, we briefly present the process for setting the different fuzzy sets as well as the way we match our emotional thesaurus and we build our fuzzy rules system.

Regarding our fuzzy sets, as we explained in Arguedas et al. (2016), we developed our own fuzzy classifier. In order to calculate the curves of each group corresponding to each magnitude (Pleasure, Arousal and Dominance), we run our tool three times, one for each dimension. At each run we gave the tool the mean of the values for that channel, its standard deviation, a file with the complete group of values for that dimension and a file with the qualitative values for that channel. In fact we used the same qualitative values for the three channels but we configured the tool to establish different qualitative values for each one of the channels. From these results, we built the curves of the three groups related to each magnitude (Pleasure, Arousal and Dominance). The final outcome of this process has been a text file where each line contains a word from the dictionary, its term in English, its term in Spanish, the values for Pleasure, Arousal and Dominance, and their corresponding numerical values. This process is described in detail in (Arguedas et al. 2016).

The group of various rules constitutes a rule base or knowledge base. Our model contains 24 rules that result from combining eight qualitative values obtained for every emotional axis of the Plutchick's model as primary, secondary and tertiary dyads. These rules are propositions that allow us to express the available knowledge about the relationship between antecedents and consequents and make affirmations of the *If-Then* type. For more details see Arguedas et al. (2016).

Our conceptual fuzzy-based model for inferring emotional states

This section presents our conceptual model which is built by a set of components depicted in Figure 1. These components process the students' texts through various steps which are explained below in more detail. In fact, our conceptual fuzzy-based model for inferring emotional states is described in three layers as shown in Figure 1.

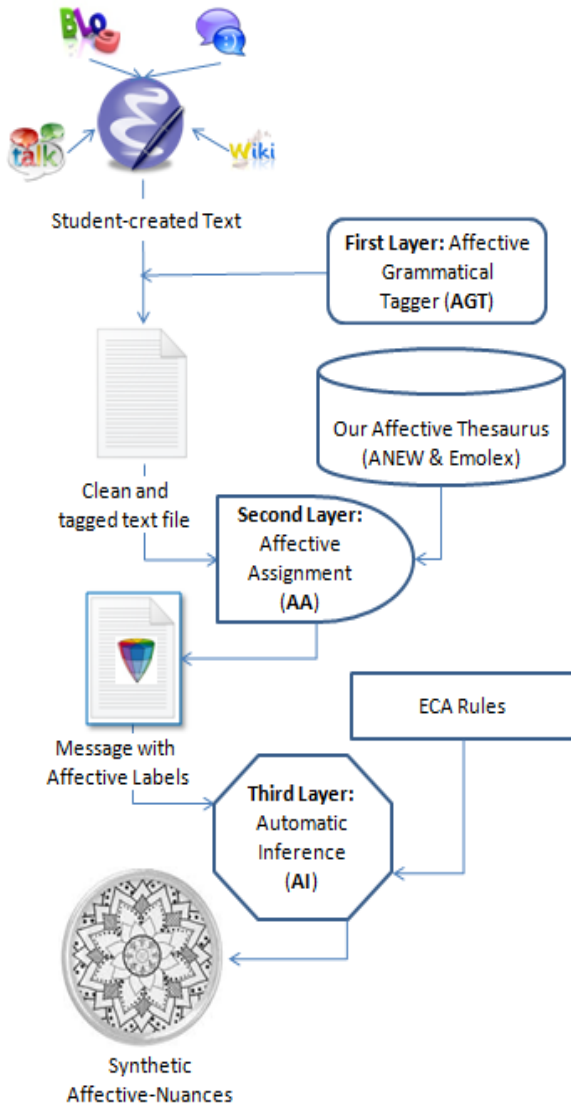


Figure 1. A conceptual fuzzy-based model for inferring emotional states

First Layer: The Affective Grammatical Tagger (AGT) takes the text of discourse created by one or more

students and performs several processes: (1) The AGT checks the messages word by word in order to solve typing problems that could have been in the text. (2) The AGT verifies that the words are in the dictionary. During this process, the AGT analyzes the text received as input. The output of this component will be two files, one of them with a clean and tagged text and the other with codes of "Complemented Language" (Etchevers 2006) that contains emoticons, onomatopoeia (haha, mua, m-mhmm), repetition of words (exceeeeeelent), etc., as shown in Figure 2.

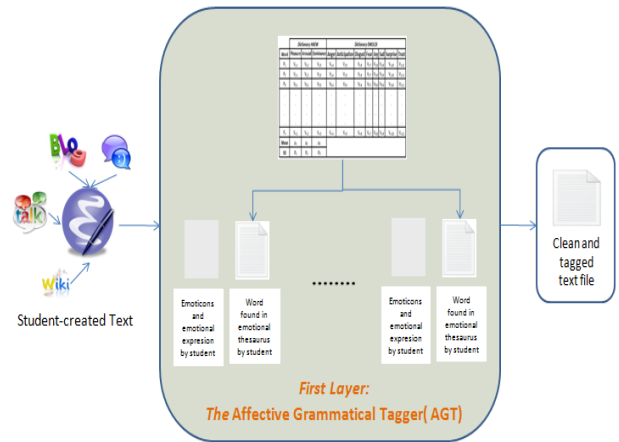


Figure 2. First Layer: The Affective Grammatical Tagger (AGT)

Second layer: The Affective Assignment (AA) takes each of the text files obtained previously, associates each word of them with the terms contained in our thesaurus and assigns to every word the dimensional and categorical emotional load.

The dimensional affective load is obtained using the information provided by ANEW dictionary, while the categorical affective load, as well as the valence (positive and negative) for each of the words is obtained using the information provided by Emolex.

Once all the text has been processed, the mean of the obtained value for each characteristic is then calculated and the obtained mean values for Pleasure, Arousal and Dominance are fuzzified.

Finally, a new file is generated as shown in Figure 3. This file contains a line with the obtained values for each of the authors of the discourse that is being analyzed every time.

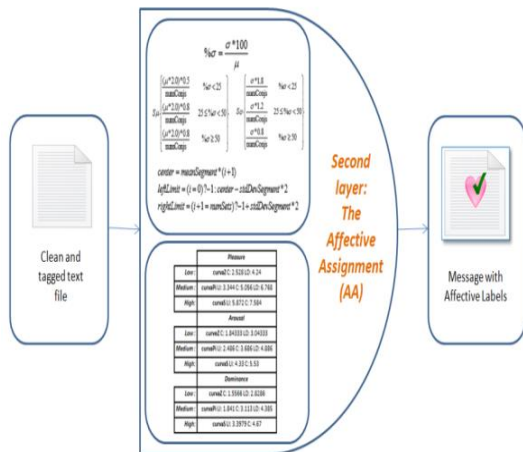


Figure 3. Second Layer: The Affective Assignment (AA)

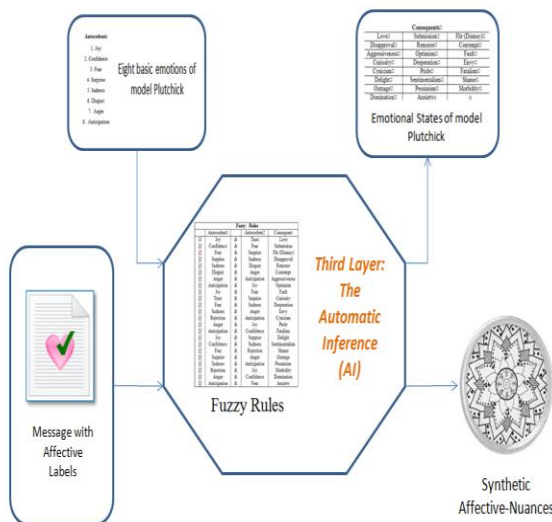


Figure 4. Third Layer: The Automatic Inference (AI)

Third layer: In this step, the **Automatic Inference (AI)** component takes the message with affective labels file as input. This file, that contains the average affective information of each author both at dimensional and categorical level, is used to trigger our fuzzy rules system. The outcome of fuzzy rules system is to obtain the emotional states of each author during his/her learning process.

The output of this component is the synthetic opinion of the original message from the point of view of the emotional and affective state of the authors of the analyzed discourse, as shown in Figure 4.

4 Implementation and Extension of the Model

Our fuzzy based classification model (FBCM)¹ is built out of a set of tools implemented in Java. These tools process the students-created texts following the phases defined in the conceptual model.

Once the process has finished, we obtain the synthetic opinion of the original message that depicts author's emotional and affective state, which is then presented to the Affective Pedagogical Tutor (APT).

The APT can be human or virtual. In this work, we implemented and used a virtual APT for experimentation. It is a client-server Web application and constitutes an important extension of the model.

This application is installed in the computer of each student, connects to the server and displays the APT on the left side of the screen. In fact, we created an integrated learning environment that integrates the APT, the Moodle LMS embedded on the right side of the screen, a text edit box that allows the student to contact the APT textually, as well as the FBCM model that processes the emotional information.

The APT can be represented by two characters (masculine/feminine) so that each student can choose the character that is more comfortable for him/her. As such, the APT is characterized by a specific voice, the emotional expressions that can display and the dialogues that can be involved.

Students carry out the activity designed by the human teacher in the Moodle LMS. The interactive e-learning course is broadcasted to every user with the same contents. Depending on the student's academic and emotional evolution throughout the course, the human teacher can provide additional personalized material and exercises.

The student works on the LMS, carries out his/her tasks, collaborates with peers whereas at the same time he/she can interact with the APT in a textual manner through the edit box located at the bottom of the screen.

¹ The fuzzy based classification model has been implemented and patented by Marta Arguedas with name FUZZYEMOSYS, on March 16, 2016 under registration number TXu 1-997-006.

The Affective Pedagogical Tutor responds to student with audible and gestural signals that were scheduled in advance, while providing the student the information that he/she previously requested, as shown in Figure 5.



Figure 5. Gestural signals that were scheduled in advance in the APT.

When a student completes a task, the APT accesses the student-created texts and forwards them to be processed by the FBCM model. The FBCM stores the extracted emotional information in the web server.

Finally, the APT can easily access every emotional log (sorted by student name, course name and task) through the Moodle platform, take it as input and create adequate affective feedback.

The framework used offers some predefined virtual agents to be used (if needed) as virtual tutors in the learning environment. Each available virtual agent has its own set of facial expressions and body animations that can be used as needed.

The goal of these expressions and animations is to make the virtual agents able to express different emotional states, making their representation more believable. In order to increase the lifelikeness of the communication between the student and the virtual agent, bidirectional natural language conversation has been enabled.

The virtual agent is able to speak to the student and to understand what the student is saying. The student-virtual agent communication can be established via text, thanks to an integrated text box in which the student can chat, so text messages can be clearly displayed on the screen. Thanks to the use of an AIML²-based conversational engine and a fusion rule engine that processes questions in natural language, the student is allowed to use text and express him/herself in natural language as well as to ask questions to an Ontology repository (like DBPEDIA³) to obtain the answer.

5 Data Sets from a Real Learning Context

Learning Context

Nowadays computer-supported collaborative learning (CSCL) environments are viewed as an important electronic learning medium for distance education. Working together, while accomplishing a task is seen as a characteristic of a powerful learning environment, aiming at active construction of knowledge. Through a process of interaction and negotiation students have an active and constructive role in the learning process (Dewiyanti 2007). Taking emotions into account, we need to provide teachers with different methods and tools to let them understand and analyze the emotional phenomenon and how it evolves over time. Our approach

² AIML: Artificial Intelligence Markup Language (www.alicebot.org/aiml.html)

³ DBpedia is a crowd-sourced community effort to extract structured information from Wikipedia and make this information available on the Web (<http://wiki.dbpedia.org/about>)

lies on an emotion analysis model, which has been widely described in Arguedas & Daradoumis (2013) (Arguedas, Daradoumis & Xhafa, 2016a; 2016b). This model is based on the Activity Theory (AT) (Engeström et al. 1999) and describes a scenario where participants (teacher and students) work together and interact with specific objects to carry out goal-oriented activities.

Within this scenario we initially developed a discourse analysis method to analyze collaborative learning activities (that included written text and dialogues) in a non-intrusive way in order to capture process and identify the emotional information extracted from each student individually as well as from the students' group as a whole. This information was presented to the human tutor and provided him/her with the necessary emotion awareness with regard to the way students' emotions appear and evolve over time. This enabled the tutor to offer students cognitive and affective feedback. Apart from identifying emotional states and behavior, the result of this approach has been the graphical representation of the students' emotions that took place during these activities.

Research Hypothesis and Goals

Goal: The main goal of this work has been to extend the above approach and build a new model that provides an improved and more calibrated measurement and representation of students' emotions that can be used as input to a virtual tutor who can create adequate affective feedback.

At this stage of our research, our aim has been to analyze the combined effects of the emotion awareness produced by our new model and of basic online tutorial affective feedback on students' learning performance in long-term blended collaborative learning practices.

Hypothesis: "Providing enhanced and more calibrated emotion awareness as well as automatic online tutorial affective feedback to learners improves their learning outcomes on time."

Research Question

Is there any significant correlation between students' calibrated emotion awareness, the related virtual tutor affective feedback and students' learning outcome?

Independent Variable:

X = emotion awareness & affective feedback

Dependent Variables:

J = learning outcome

Participants and Procedure

Participants were a sample of 48 fourth-year high school students attending the subject "Web Design". Among students, 24 were girls (50%) and 24 were boys (50%) We divided students in 16 groups of four and we chose 8 of them as the experimental group and the rest as the control group. The experiment was conducted for five weeks with a total of 15 sessions.

The procedure we followed was to design a scenario that included a collaborative learning activity which was implemented following the Problem-Based Learning method and the Jigsaw collaborative strategy. The topic of the activity was "How to design a web page" and was carried out in the Moodle environment.

The activity designed by the human teacher was arranged in several synchronous and asynchronous tasks that included a website design, a forum debate and a chat, where students were encouraged to participate actively in building their knowledge. In this way, the teacher's role was reduced to guide and give support to the learning activity, by providing appropriate advice or help when needed. Based on the Jigsaw collaborative strategy, the learning activity was divided in ten stages which in turn were grouped around five tasks to facilitate their implementation. For each task, the teacher provided all the necessary resources (documents and tools).

Research Instruments

At the end of each task of the activity, we used our fuzzy based classification model for processing and identifying the emotional information extracted from texts created by students in the chat and forum debate.

On the one hand, the data collected was used to trigger fuzzy rules according to Plutchik's model (Plutchik 2001) and obtain the emotional states that were experienced by students during the realization of the respective tasks.

On the other hand, we used the data gathered to represent graphically the arousal of text, according to the ANEW dictionary as well as the valence and the different types of emotion detected during the analysis, according to the NCR Emotion Lexicon. Regarding the statistical techniques employed in the analysis of the gathered data, we used descriptive statistics, calculating relative

frequencies (%), as well as graphics to represent reality objectively. We also used bivariate correlation and analysis of variance to find relationships between the variables under study for the research question of our study.

6 Model testing

As mentioned before, an experiment was carried out with high school students for testing our tool. We considered the texts produced by the students in various forum debates and chats.

These texts along with the thesaurus built before served as an input for our tool. The output of FBCM consists of a file that contains a line for each of the authors of the discourse with the obtained values for the dimensional and categorical characteristics, as well as the emotional states that result from the rules application to each set of categorical characteristics we obtained.

Data Collection

At this point, in order to carry out the statistical analysis of the data resulting from our tool and the student grades, we defined the following indicators, as shown in Table 1:

(a) Dimensional Characteristics (DI) include indicators that concern fuzzy values of three basic dimensions through which the entire range of human emotions can be organized: Pleasure, Arousal and Dominance or Control,

(b) Categorical Characteristics (CI) include indicators that concern positive and negative emotions based on Plutchick model and

(c) Emotional States of Fuzzy Rules (CFR) include indicators that concern emotional states based on the results from applying our fuzzy rules.

Table 1. Indicators and their tags used in statistical calculations.

Dimensional Characteristics (DI)	
Tag	AXES/ Indicators
DI.1	Pleasure
DI.2	Arousal
DI.3	Dominance

Categorical Characteristics (CI)	
Tag	AXES/Indicators
CI.1	Joy
CI.2	Confidence
CI.3	Fear
CI.4	Surprise
CI.5	Sadness
CI.6	Disgust
CI.7	Anger
CI.8	Anticipation

Emotional States of Fuzzy Rules (CFR)	
Tag	AXES/Indicators
CFR.1	Love
CFR.2	Disapproval
CFR.3	Aggressiveness
CFR.4	Curiosity
CFR.5	Cynicism
CFR.6	Delight
CFR.7	Outrage
CFR.8	Domination
CFR.9	Submission
CFR.10	Remorse
CFR.11	Optimism
CFR.12	Desperation
CFR.13	Pride
CFR.14	Sentimentalism
CFR.15	Pessimism
CFR.16	Anxiety
CFR.17	Flit (Dismay)
CFR.18	Contempt
CFR.19	Fault
CFR.20	Envy
CFR.21	Fatalism
CFR.22	Shame
CFR.23	Morbidity

Emotion Awareness

In order to provide emotion awareness among participants in the experimental group, we applied our FBCM at all conversations that took place in the group during the learning activity.

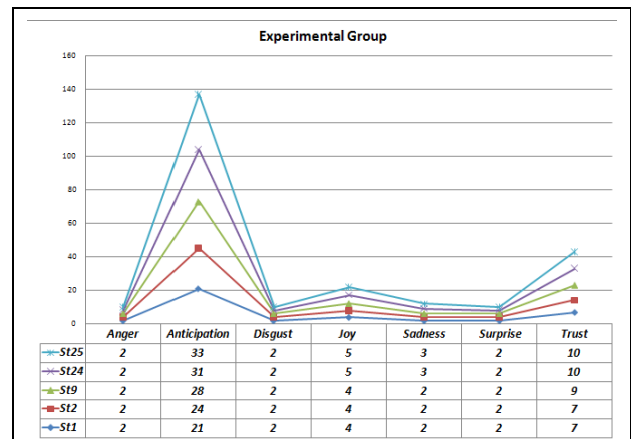
The outcome of our model was a graphical representation of word clouds attributed to each student. This graphical representation was provided to both APT and students of the experimental group, at the end of each task, as shown in Figure 7.



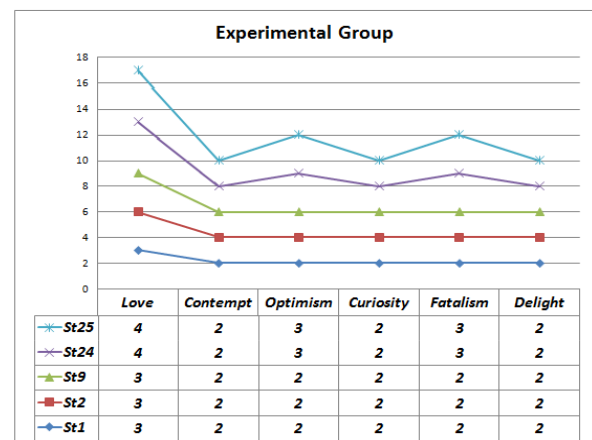
Figure 7.- Example of Word Clouds provided emotion awareness to each participant in the experimental group

In this way, the APT was aware of students' emotions during their interactions in the virtual learning space (chat and forum). In fact, the APT could observe not only which emotions and emotional states of its students were present but also, and most importantly, the level of pleasure, arousal and dominance of these emotions. Figure 8 shows the evolution of dimensional

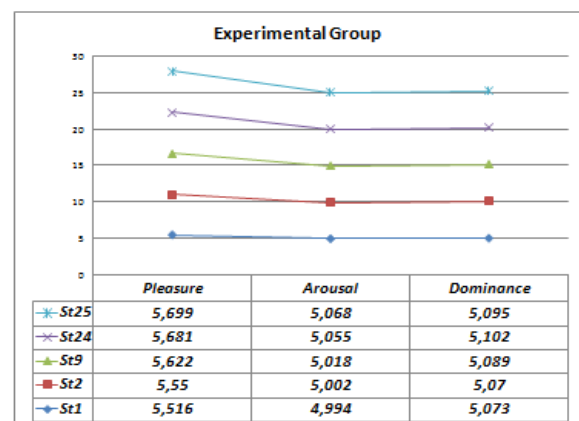
and categorical emotions as well as the emotional states of one team of 5 students in the EG.



(a) Categorical emotions shown by a 5-student team



(b) Emotional states shown by a 5-student team



(c) Dimensional emotions shown by a 5-student team

Figure 8.- Graphical Representation of emotion awareness of one team of 5 students in the EG

In this way, the APT could intervene on time. In addition, students were aware of the emotions, emotional states and the level of pleasure, arousal and dominance that they experienced during the accomplishment of the task. In contrast, the students of the control group were not supported by this facility and carried out their activity in a conventional way.

Presentation of the results.

i. The Cronbach's alpha coefficient.

To ensure the reliability of data collection, we applied the Cronbach's alpha coefficient to both Experimental Group (EG) and Control Group (CG). Cronbach's alpha is considered to be a coefficient of reliability (or consistency).

A reliability coefficient of .70 or higher is considered "acceptable" in most social science research situations.

All values of Cronbach's alpha in Table 2 are higher than .70, which reinforces the reliability of our indicators.

Table 2. The Cronbach's alpha coefficient of DI, CI and CFR, in EG and CG.

THE CRONBACH'S ALPHA (EG) ((DI), (CI) and (CFR))
,769
THE CRONBACH'S ALPHA (CG) ((DI), (CI) and (CFR))
,900

ii Descriptive statistics

The skewness and kurtosis of each variable were examined to check for multivariate normality.

The absolute values of skewness and the absolute values of kurtosis did not exceed a univariate skewness of 2.0 and a univariate of kurtosis of 7.0.

In our case, it was assumed that there was no critical problem regarding multivariate normality in EG.

However, in CG the only cases where the values of skewness and kurtosis did not exceed a univariate skewness of 2.0 and a univariate of kurtosis of 7.0 occurred for the items of DI.1, DI.2 and DI.3 in the control group, as shown in Table 3.

Table 3. The descriptive statistics of DI, CI and CFR, in EG and CG.

	The results of descriptive statistics (EG)					
	Min	Max.	Mean	SD	Skewness	Kurtosis
DI.1	5,27	5,70	5,4876	,12883	,327	-,628
DI.2	4,92	5,07	4,9930	,04047	,688	,046
DI.3	4,93	5,10	5,0454	,05081	-1,182	,661
CI.7	2	2	2,00	,000	.	.
CI.8	16	33	22,55	5,343	1,189	,064
CI.6	2	2	2,00	,000	.	.
CI.1	4	5	4,18	,395	1,773	1,250
CI.5	1	3	1,50	,802	1,220	-,202
CI.4	2	2	2,00	,000	.	.
CI.2	5	10	7,36	1,529	,548	-,067
CFR.1	3	4	3,18	,395	1,773	1,250
CFR.18	2	2	2,00	,000	.	.
CFR.11	2	3	2,18	,395	1,773	1,250
CFR.4	2	2	2,00	,000	.	.
CFR.21	2	3	2,18	,395	1,773	1,250
CFR.6	2	2	2,00	,000	.	.

	The results of descriptive statistics (CG)					
	Min	Max.	Mean	SD	Skewness	Kurtosis
DI.1	5,59	5,75	5,6420	,04051	1,090	,391
DI.2	5,07	5,11	5,0958	,00918	-1,089	1,344
DI.3	5,04	5,12	5,0668	,02520	,878	-,593
CI.7	4	12	6,09	1,411	3,339	15,517
CI.8	35	84	40,70	9,484	4,720	22,519
CI.6	4	12	6,09	1,411	3,339	15,517
CI.1	5	16	7,48	2,129	2,988	12,095
CI.5	3	8	3,91	,996	3,210	13,730
CI.4	2	6	3,09	,668	3,906	18,526
CI.2	10	24	12,09	2,661	4,414	20,563
CFR.1	4	12	5,78	1,476	3,570	15,534
CFR.18	4	12	6,09	1,411	3,339	15,517
CFR.11	3	8	4,13	,869	4,284	20,216
CFR.4	2	6	3,09	,668	3,906	18,526
CFR.21	3	8	4,13	,869	4,284	20,216
CFR.6	2	6	3,09	,668	3,906	18,526

iii *Pearson's Correlations.*

Finally, we present the correlations between variables DI & Mark, CI & Mark and CFR & Mark that were found in the experimental group in Tables 4. Mark is directly related to students' learning outcome.

First, a significant positive correlation was found between DI (dimensional characteristics) and Mark in the Experimental Group. In particular, we found higher correlations between Mark, and Pleasure (DI.1) ($r = .760$, $p < .01$), Arousal (DI.2) ($r = .755$, $p < .05$) and Dominance (DI.3) ($r = .556$, $p < .01$)

Second, a significant positive correlation was found between CI (categorical characteristics) and Mark. In particular, we found higher correlations between Mark and Joy (CI.1) ($r = .704$, $p < .01$), Sadness (CI.5) ($r = .895$, $p < .01$), Trust (CI.2) ($r = .785$, $p < .01$) and Anticipation (CI.8) ($r = .844$, $p < .01$).

Third, a significant positive correlation was found between CFR (Emotional States of Fuzzy Rules) and Mark. In particular, we found higher correlations between Mark and Love (CFR.1) ($r = .704$, $p < .01$), Optimism (CFR.11) ($r = .704$, $p < .01$) and Fatalism (CFR.21) ($r = .704$, $p < .01$).

In contrast, in Control Group the only significant correlation found was between Mark and Arousal ($r = .480$, $p < .05$)

Table 4. Pearson's Correlation of DI, CI and CFR with Marks, in EG and CG

Pearson's Correlation - EG					
	Mark		Mark		Mark
DI.1	-.760**	CI.1	-.704**	CFR.1	-.704**
DI.2	-.755**	CI.5	-.895**	CFR.11	-.704**
DI.3	-.556**	CI.2	-.785**	CFR.21	-.704**
		CI.8	-.844**		
* Correlation is significant at the 0.05 level (2-tailed)					
** Correlation is significant at the 0.01 level (2-tailed)					

Pearson's Correlation - CG					
	Mark		Mark		Mark
DI.1	-.216	CI.1	.007	CFR.1	-.021
DI.2	.480*	CI.5	.119	CFR.11	.048
DI.3	-.229	CI.2	-.003	CFR.21	.048
		CI.8	.011		
* Correlation is significant at the 0.05 level (2-tailed)					
** Correlation is significant at the 0.01 level (2-tailed)					

7 Discussion

Based on the results we obtained in the above Section 6, we proceed to discuss and provide a response to the research question we set in this work:

RQ: Is there any significant correlation between students' calibrated emotion awareness, the related virtual tutor affective feedback and students' learning outcome?

Our results showed that this question had a positive answer in all aspects in the Experimental Group which was had been endowed with emotion awareness and emotional feedback facilities. Table 3 shows that dimensional characteristics (such as Pleasure, Arousal and Dominance), categorical characteristics (such as joy, sadness, confidence/trust and anticipation) and emotional states (such as Love, Optimism and Fatalism) had significant positive effects on a better learning performance in the EG.

In particular, our experiment showed that dimensional characteristics had a higher positive impact in the outcomes of the learning activity of students in the EG, whereas in the Control Group only Arousal had an implicit positive relation with the students' obtained outcome.

Moreover, categorical characteristics (like joy, sadness, trust and anticipation) had significant positive effects on the activity performance for the EG students. No significant correlation was reported for CG students.

Finally, once our ECA Fuzzy Rules were triggered, the monitoring of emotional states such as Love, Optimism and Fatalism by the Affective Pedagogical Tutor (APT) had significant positive effects on the performance of EG students. In contrast, no significant correlation was reported for CG students.

8 Conclusions and Future Work.

In this work we have presented a model for monitoring students' emotions using fuzzy logic in an e-learning environment, with the aim to provide both emotion awareness and affective feedback to students through an online Affective Pedagogical Tutor (APT). To test and validate our model we run an experiment in a real e-collaborative learning situation.

The results of the experiment showed that our model was highly effective in the Experimental Group (EG) of students who were supplied with both emotion awareness and affective feedback. In this group, the explicit graphical representation of dimensional and categorical emotions after every task proved to produce much better learning results than those in the Control Group (CG). In fact, the Word Clouds produced helped students be conscious of their emotions, overcome possible situations of sadness, enhance anticipation and trust and accomplish their task successfully. Regarding the function and influence of APT, the results showed that its affective feedback had really a very positive effect on EG students. CG students really missed this opportunity. In fact, the support offered by APT, through emotional expressions and advice, helped EG students both overcome emotional states such as fatalism and increase their optimism, which led them to carry out their activities successfully.

Future work is focused on improving our fuzzy rules system in order to accomplish three important goals in virtual learning environments: increase its accuracy, include more emotional states in its monitoring capabilities, and process and analyze the students' emotional states in real time.

Acknowledgements The authors want to thank the Technological Institute of Aragon for allowing us to employ the images and avatars used in this work.

Compliance with Ethical Standards

Conflict of interest: The authors declare that they have no conflict of interest. In addition, all procedures

performed in our experiment that involved human participants (a human tutor and students) were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

References

- Ali, Farman, Kyung-Sup Kwak, and Yong-Gi Kim (2016) Opinion mining based on fuzzy domain ontology and Support Vector Machine: a proposal to automate online review classification. *Applied Soft Computing*, 47, October 2016, pp. 235–250.
- Andreevskaia, A. & Bergler, S. "Mining WordNet for a fuzzy sentiment: Sentiment tag extraction from WordNet glosses," in *Proceedings of the European Chapter of the Association for Computational Linguistics (EACL)*, 2006.
- Andrew L. Maas, Raymond E. Daly, Peter T. Pham, Dan Huang, Andrew Y. Ng, and Christopher Potts (2011) Learning word vectors for sentiment analysis. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1 (HLT '11)*, Vol. 1. Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 142-150.
- Arguedas, M., Casillas, L., Xhafa, F., Daradoumis, T., Peña, A. and Caballé, S. (2016). A Fuzzy-based Approach for Classifying Students' Emotional States in Online Collaborative Work. In *Proceedings of the 10th International Conference on Complex, Intelligent and Software Intensive Systems (CISIS-2016)*, Fukuoka, Japan, July 6-8, 2016, IEEE Computer Society, Los Alamitos, CA, USA. In press.
- Arguedas, M., Daradoumis, T. and Xhafa, F. (2014). Towards an Emotion Labeling Model to Detect Emotions in Educational Discourse. In *proceedings of the 8th International Conference on Complex, Intelligent and Software Intensive Systems (CISIS-2014)*, Birmingham, UK, July 2-4, 2014, IEEE Computer Society, Los Alamitos, CA, USA. DOI 10.1109/CISIS.2014.36, ISBN: 978-1-4799-4325-8, pp. 72-78.
- Arguedas, M., Daradoumis, T. and Xhafa, F. (2016a). Analyzing the effects of emotion management on time and self-management in computer-based learning. *Computers in Human Behavior (CHB)*, 63, 517-529.
- Arguedas, M., Daradoumis, T. and Xhafa, F. (2016b). Analyzing how emotion awareness influences students' motivation, engagement, self-regulation and learning outcome. *Educational Technology & Society*, Special Issue on "Intelligent and Affective Learning Environments: New Trends and Challenges", 19 (2), 87-103.
- Baccianella, S., Esuli, A., & Sebastiani, F. (2010, May). SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In *LREC (Vol. 10, pp. 2200-2204)*.

- Becker, C., Kopp, S., & Wachsmuth, I. (2007). Why emotions should be integrated into conversational agents. *Conversational informatics: an engineering approach*, 49-68.
- Bradley, M. M., & Lang, P. J. (1999). Affective norms for English words (ANEW): Instruction manual and affective ratings (pp. 1-45). Technical report C-1, the center for research in psychophysiology, University of Florida.
- Calvo, R. A. & D'Mello, S. K. (2010). Affect Detection: An Interdisciplinary Review of Models, Methods, and their Applications. *IEEE Transactions on Affective Computing*, 1(1), pp.18-37.
- Daradoumis, T., Arguedas, M. and Xhafa, F. (2013). Building Intelligent Emotion Awareness for Improving Collaborative e-Learning. In proceedings of the 5th IEEE International Conference on Intelligent Networking and Collaborative Systems (INCoS 2013), September 9-11, 2013, Xi'an, China, IEEE Computer Society, Los Alamitos, CA, USA. DOI: 10.1109/INCoS.2013.49, ISBN: 978-0-7695-4988-0. pp. 281-288.
- Dewiyanti, S., Brand-Gruwel, S., Jochems, W., & Broers, N. J. (2007). Students' experiences with collaborative learning in asynchronous computer-supported collaborative learning environments. *Computers in Human Behavior*, 23(1), 496-514.
- Ekman, P. & Friesen, W. V. (1971). Constants across culture in the face and emotion. *Journal of Personality and Social Psychology*, 17(2), pp. 124-129.
- Esuli, A., & Sebastiani, F. (2006, May). Sentiwordnet: A publicly available lexical resource for opinion mining. In *Proceedings of LREC* (Vol. 6, pp. 417-422)
- Etchevers, N. (2006). Los nuevos códigos de la comunicación emocional utilizados en Internet. *Revista Electrónica Teoría de la Educación: Educación y Cultura en la Sociedad de la Información*. Monográfico sobre el «Estudio de los comportamientos emocionales en la red» Vol. 7, nº 2 (pp. 92-106). Universidad de Salamanca
- Feidakis, M., Daradoumis, T. & Caballé, S. (2013). Building emotion-aware features in Computer Supported Collaborative Learning (CSCL) systems. *Alpine Rendez-Vous (ARV) Workshop on Tools and Technologies for Emotion Awareness in ComputerMediated Collaboration and Learning (ARV 2013)*. Villard-de-Lans, Vercors, France. January 28 – February 1, 2013. TELEARC & EATEL associations.
- Frasson, C. & Chalfoun, P. (2010). Managing Learner's Affective States in Intelligent Tutoring Systems. In Nkambou, R., Mizoguchi, R. & Bourdeau, J. (Eds.) *Advances in Intelligent Tutoring Systems* (pp. 339-358). Springer-Verlag, Berlin-Heidelberg.
- Goetz, T., Pekrun, R., Hall, N., & Haag, L. (2006). Academic emotions from a social cognitive perspective: antecedents and domain specificity of students' affect in the context of Latin instruction. *The British Journal of Educational Psychology*, 76 (Pt 2), 289-308.
- Gonçalves, V. P., Giancristofaro, G. T., Geraldo Filho, P. R., Johnson, T., Carvalho, V., Pessin, G., & Ueyama, J. (2016). Assessing users' emotion at interaction time: a multimodal approach with multiple sensors. *Soft Computing*, 1-15.
- Ibarrola, B. (2000). L'educació emocional. *Infància*, 116, 5-8.
- Isen, A.M. (2000). Positive Affect and Decision Making (ed) *Handbook of Emotions*. Guilford, New York *Handbook of Emotions* (2nd ed., pp. 417-435). New York: Guilford
- Kort, B. & Reilly, R. (2002). Analytical Models of Emotions, Learning and Relationships: Towards an affectsensitive cognitive machine. *Proceedings of the International Conference on Virtual Worlds and Simulation (VWSim 2002)*, San Antonio, Texas
- Kuncheva, Ludmila. *Fuzzy classifier design*. Vol. 49. Springer Science & Business Media, 2000.
- Li, J., Fong, S., Zhuang, Y., & Khoury, R. (2015). Hierarchical classification in text mining for sentiment analysis of online news. *Soft Computing*, 1-10.
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis lectures on human language technologies*, 5(1), 1-167.
- Mehrabian, A. (1972). *Nonverbal communication*. Transaction Publishers.
- Mehrabian, A., & O'Reilly, E. (1980). Analysis of personality measures in terms of basic dimensions of temperament. *Journal of Personality and Social Psychology*, 38(3), 492..
- Mohammad, S. M., & Turney, P. D. (2013). *NRC Emotion Lexicon*. NRC Technical Report.
- Moharrer, M., Tahayori, H., Livi, L., Sadeghian, A., & Rizzi, A. (2015). Interval type-2 fuzzy sets to model linguistic label perception in online services satisfaction. *Soft Computing*, 19(1), 237-250.
- Ortony, A., Clore, G. & Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge: Cambridge University Press.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. 1957: *The Measurement of Meaning*. Urbana, IL: University of Illinois Press.
- Pekrun, R. (1992). The Impact of Emotions on Learning and Achievement: Towards a Theory of Cognitive/Motivational Mediators. *Applied Psychology: An International Review*, 41(4), pp. 359-376.
- Plutchik, R. (2001). The nature of emotions. *American Scientist*, 89, 34-350.
- Redondo, J., Fraga, I., Padrón, I., & Comesaña, M. (2007). The Spanish adaptation of ANEW (affective norms for English words). *Behavior research methods*, 39(3), 600-605.
- Russell, J. A. (1983). Pancultural aspects of human conceptual organization of emotion. *Journal of Personality and Social Psychology*, 45, 1281-1288.
- Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, 44(4), 693-727.
- Scherer, K.R., Shuman, V., Fontaine, J.R.J., & Soriano, C. (2013). The GRID meets the Wheel: Assessing emotional feeling via selfreport. In Johnny R.J. Fontaine, Klaus R. Scherer & C. Soriano (Eds.), *Components of Emotional Meaning: A sourcebook* (pp. 281-298). Oxford: Oxford University Press

Strapparava, C., & Valitutti, A. (2004, May). WordNet Affect: an Affective Extension of WordNet. In *LREC* (Vol. 4, pp. 1083-1086).

Tao, Jianhua; Tan, Tieniu (2005). Affective Computing and Intelligent Interaction. *Affective Computing: A Review*. LNCS 3784. Springer. pp. 981– 995. doi:10.1007/11573548.

Whissell, C. (2009). Using the revised dictionary of affect in language to quantify the emotional undertones of

samples of natural language 1, 2. *Psychological reports*, 105(2), 509-521.

Wu, J. L., Yu, L. C., & Chang, P. C. (2014). An intelligent stock trading system using comprehensive features. *Applied Soft Computing*, 23, October 2014, pp. 39-50.

Wundt, W. M. (1896). *Compendio de psicología*. La España Moderna.