KNOWLEDGE ENGINEERING FOR AFFECTIVE BI-MODAL HUMAN-COMPUTER INTERACTION

Efthymios Alepis, Maria Virvou
Department of Informatics, University of Piraeus, 80 Karaoli & Dimitriou St., 18534, Piraeus, Greece

Katerina Kabassi
Department of Ecology and the Environment, Technological Educational Institute of the Ionian Islands
2 Kalvou Sq., 29100 Zakynthos, Greece

Keywords: Human-Machine Interface, e-learning, knowledge engineering, multi-criteria decision making theories.

Abstract: This paper presents knowledge engineering for a system that incorporates user stereotypes as well as a multi criteria decision making theory for affective interaction. The system bases its inferences about students' emotions on user input evidence from the keyboard and the microphone. Evidence form these two modes is combined by a user modelling component underlying the user interface. The user modelling component reasons about users' actions and voice input and makes inferences in relation to their possible emotional states. The mechanism that integrates the inferences form the two modes has been based on the results of two empirical studies that were conducted in the context of requirements analysis of the system. The evaluation of the developed system showed significant improvements in the recognition of the emotional states of users.

1 INTRODUCTION

The unprecedented growth in HCI has led to a redefinition of requirements for effective user interfaces. A key component for these requirements is the ability of systems to address affect (Hudlicka, 2003). This is especially the case for computer-based educational applications that are targeted to students who are in the process of learning. Learning is a complex cognitive process and it is argued that how people feel may play an important role on their cognitive processes as well (Goleman, 1995). At the same time, many researchers acknowledge that affect has been overlooked by the computer community in general (Picard & Klein, 2002).

Picard (Picard, 2003) argues that people's expression of emotion is so idiosyncratic and variable, that there is little hope of accurately recognising an individual's emotional state from the available data. Therefore, many researchers have pointed out that there is a need of combining evidence from many modes of interaction so that a computer system can generate as valid hypotheses as possible about users' emotions (e.g. Oviatt, 2003, Pantic & Rothkrantz, 2003). However, for the time being, very little research has been reported in the literature towards this direction.

In this paper, we present the knowledge engineering process for combining two modes of interaction, namely keyboard and microphone, for the development of an affective educational application. The educational application is called Edu-Affe-Mikey and is an affective educational software application targeted to first-year medical students.

The main characteristic of the system is that it combines evidence from the two modes mentioned above in order to identify the users' emotions. The results of the two modes are combined through a multi-criteria decision making method. More specifically, the system uses Simple Additive Weighting (SAW) (Fishburn 1967, Hwang & Yoon 1981) for evaluating different emotions, taking into account the input of the two different modes and selects the one that seems more likely to have been felt by the user. In this respect, emotion recognition is based on several criteria that a human tutor would

Alepis E., Virvou M. and Kabassi K. (2007).
KNOWLEDGE ENGINEERING FOR AFFECTIVE BI-MODAL HUMAN-COMPUTER INTERACTION.
In *Proceedings of the Second International Conference on Signal Processing and Multimedia Applications*, pages 218-225
DOI: 10.5220/0002138302180225

have used in order to perform emotion recognition of his/her students during the teaching course.

The values of the criteria used in our novel approach that is described in this paper, are acquired by user stereotypes. For this purpose, user stereotypes have been constructed with respect to the different emotional states of users that these users are likely to have experienced in typical situations during the educational process and their interaction with the educational software. Examples of such situations are when a student makes an error while answering an exam question or when a user reads about a new topic within the educational application etc.

The user stereotypes have been resulted from an empirical study that we conducted among 50 users. The empirical study aimed at finding out common user reactions of the target group of the application that express user feelings while they interact with educational software.

The main body of this paper is organized as follows: In section 2 we present and discuss related work. In the next section we describe the overall educational application and in sections 4 and 5 we present briefly the experimental studies for requirements analysis. Section 6 describes the application of the multi-criteria decision making method in the context of the educational application. In section 7 we present and discuss the results of the evaluation of the multi-criteria model. Finally, in section 8 we give the conclusions drawn from this work.

2 RELATED WORK

2.1 Stereotypes

Stereotypes constitute a popular user modeling technique for drawing inferences about users belonging to different groups and were first introduced by Rich (Rich, 1983). Stereotype-based reasoning takes an initial impression of the user and uses this to build a user model based on default assumptions (Kay, 2000). Therefore, Kobsa et al. (Kobsa et al, 2001) describe a stereotype as consisting of two main components: A set of activation conditions (triggers) for applying the stereotype to a user and a body, which contains information that is typically true of users to whom the stereotype applies. The information that a stereotype contains is further used by a system in order to personalise interaction.

The need of incorporating stereotypes concerning users' characteristics in modern multi-modal application interfaces is important; individuals can more effectively understand universal emotions expressed by members of a cultural group to which they have greater exposure (Elfenbein & Ambady, 2003).

The importance of user stereotypes is acknowledged by other researchers as well in the area of emotion recognition. For example, in (Moriyama & Ozawa 2001) is suggested that the incorporation of stereotypes in emotion-recognition systems improves the systems' accuracy. Despite this importance, in 2001 there were only a few studies based on emotion stereotypes but the interest in such approaches was rapidly growing (Moriyama et al, 2001).

2.2 Simple Additive Weighting

The Simple Additive Weighting (SAW) (Fishburn, 1967, Hwang & Yoon, 1981) method is among the best known and most widely used decision making method. SAW consists of two basic steps:

1. **Scale the values of the *n* criteria to make them comparable.** There are cases where the values of some criteria take their values in [0,1] whereas there are others that take their values in [0,1000]. Such values are not easily comparable. A solution to this problem is given by transforming the values of criteria in such a way that they are in the same interval. If the values of the criteria are already scaled up this step is omitted.
2. **Sum up the values of the *n* criteria for each alternative.** As soon as the weights and the values of the *n* criteria have been defined, the value of a multi-criteria function is calculated for each alternative as a linear combination of the values of the *n* criteria.

The SAW approach consists of translating a decision problem into the optimisation of some multi-criteria utility function U defined on A. The decision maker estimates the value of function $U(X_j)$ for every alternative X_j and selects the one with the highest value. The multi-criteria utility function U can be calculated in the SAW method as a linear combination of the values of the *n* criteria:

$$U(X_j) = \sum_{i=1}^{n} w_i x_{ij} \qquad (1)$$

where X_j is one alternative and x_{ij} is the value of the *i* criterion for the X_j alternative.

3 REQUIREMENTS ANALYSIS

Requirements specification and analysis resulted from two different empirical studies. The first empirical study participated 50 potential users of the educational system and it revealed the basic requirements for affective bi-modal interaction. The second empirical study, on the other hand, participated 16 expert users and the information collected were used for defining the criteria for determining the emotional states of users. These criteria would be used in the next phases of the software life-cycle for applying the multi-criteria decision making model.

3.1 Determining Requirements for Affective Bi-modal Interaction

In order to find out how users express their emotions through a bi-modal interface that combines voice recognition and input from keyboard we have conducted an empirical study. This empirical study involved 50 users (male and female), of the age range 17-19 and at the novice level of computer experience. The particular users were selected because such a profile describes the majority of first year medical students in a Greek university which the educational application is targeted to. They are usually between the age of 17 and 19 and usually have only limited computing experience, since the background knowledge required for medical studies does not include advanced computer skills.

In the first phase of the empirical study these users were given questionnaires concerning their emotional reactions to several situations of computer use in terms of their actions using the keyboard and what they say. Participants were asked to determine what their possible reactions would be when they are at certain emotional states during their interaction. Our aim was to recognise the possible changes in the users' behaviour and then to associate these changes with emotional states like anger, happiness, boredom, etc.

After collecting and processing the information of the empirical study we came up with results that led to the design of the affective module of the educational application. For this purpose, some common positive and negative feelings were identified.

The results of the empirical study were also used for designing the user stereotypes. In our study user stereotypes where built first by categorizing users by their age, their educational level and by their computer knowledge level. The reason why this was done was that people's behaviour while doing something may be affected by several factors concerning their personality, age, experience, etc. Indeed, the empirical study revealed many cases of differences among users. For example, experienced computer users may be less frustrated than novice users. Younger computer users are usually more expressive than older users while interacting with an animated agent and we may expect to have more data from audio mode than by the use of a keyboard. The same case is when a user is less experienced in using a computer than a user with a high computer knowledge level. In all these cases stereotypes were constructed to indicate which specific characteristics in a user's behaviour should be taken more to account in order make more accurate assumptions about the users' emotional state.

The empirical study also revealed that the users would also appreciate if the system adapted its interaction to the users' e-motional state. Therefore, the system could use the evidence of the emotional state of a user collected by a bi-modal interface in order to re-feed the system, adapt the agent's behaviour to the particular user interacting with the system and as a result make the system more accurate and friendly.

3.2 Determining Multiple Criteria

Decision making theories provide precise mathematical methods for combining criteria in order to make decisions but do not define the criteria. Therefore, in order to locate the criteria that human experts take into account while providing individualised advice, we conducted a second empirical study.

The empirical study should involve a satisfactory number of human experts, who will act as the human decision makers and are reviewed about the criteria that they take into account when providing individualised advice. Therefore, in the experiment conducted for the application of the multi-criteria theory in the e-learning system, 16 human experts were selected in order to participate in the empirical study. All the human experts possessed a first and/or higher degree in Computer Science.

The participants of the empirical study were asked which input action from the keyboard and the microphone would help them find out what the emotions of the users were. From the input actions that appeared in the experiment, only those proposed by the majority of the human experts were selected. In particular considering the keyboard we have: a) user types normally b) user types quickly (speed higher than the usual speed of the particular user) c) user types slowly (speed lower than the usual speed of the particular user) d) user uses the backspace key often e) user hits unrelated keys on the keyboard f) user does not use the keyboard.

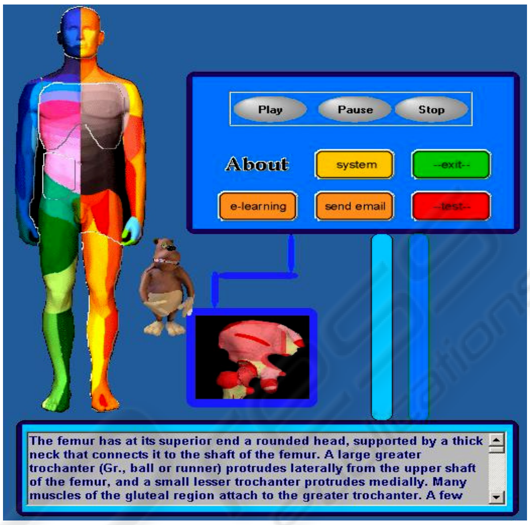
Considering the users' basic input actions through the microphone we have 7 cases: a) user speaks using strong language b) users uses exclamations c) user speaks with a high voice volume (higher than the average recorded level) d) user speaks with a low voice volume (low than the average recorded level) e) user speaks in a normal voice volume f) user speaks words from a specific list of words showing an emotion g) user does not say anything.

Concerning the combination of the two modes in terms of emotion recognition we came to the conclusion that the two modes are complementary to each other to a high extent. In many cases the human experts stated that they can generate a hypothesis about the emotional state of the user with a higher degree of certainty if they take into account evidence from the combination of the two modes rather than one mode. Happiness has positive effects and anger and boredom have negative effects that may be measured and processed properly in order to give information used for a human-computer affective interaction. For example, when the rate of typing backspace of a user increases, this may mean that the user makes more mistakes due to a negative feeling. However this hypothesis can be reinforced by evidence from speech if the user says something bad that expresses negative feelings.

4 OVERVIEW OF THE SYSTEM

In this section, we describe the overall functionality and emotion recognition features of our system, Edu-Affe-Mikey. The architecture of Edu-Affe-Mikey consists of the main educational application with the presentation of theory and tests, a programmable human-like animated agent, a monitoring user modelling component and a database.

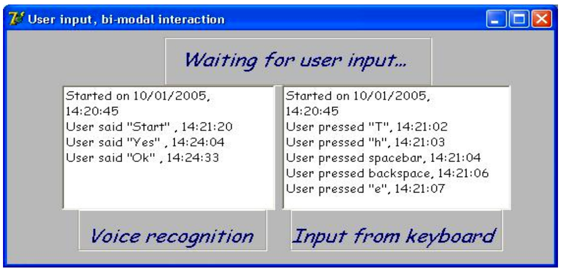
While using the educational application from a desktop computer, students are being taught a particular medical course. The information is given in text form while at the same time the animated agent reads it out loud using a speech engine. The student can choose a specific part of the human body and all the available information is retrieved from the systems' database. In particular, the main application is installed either on a public computer where all students have access, or alternatively each student may have a copy on his/her own personal computer. An example of using the main application is illustrated in figure 1. The animated agent is present in these modes to make the interaction more human-like.



The femur has at its superior end a rounded head, supported by a thick neck that connects it to the shaft of the femur. A large greater trochanter (Gr., ball or runner) protrudes laterally from the upper shaft of the femur, and a small lesser trochanter protrudes medially. Many muscles of the gluteal region attach to the greater trochanter. A few

Figure 1: A screen-shot of theory presentation in Edu-Affe-Mikey educational application.

While the users interact with the main educational application and for the needs of emotion recognition a monitoring component records the actions of users from the keyboard and the microphone. These actions are then processed in conjunction with the multi-criteria model and interpreted in terms of emotions. The basic function of the monitoring component is to capture all the data inserted by the user either orally or by using the keyboard and the mouse of the computer. The data is recorded to a database and the results are returned to the basic application the user interacts with. Figure 2 illustrates the "monitoring" component that records the user's input and the exact time of each event.



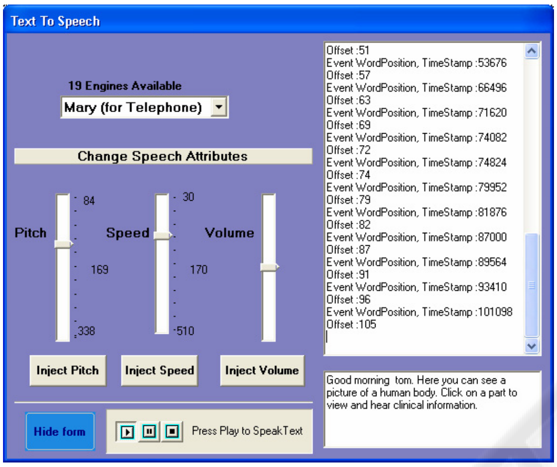
```
Started on 10/01/2005,
14:20:45
User said "Start" , 14:21:20
User said "Yes" , 14:24:04
User said "Ok" , 14:24:33
```

```
Started on 10/01/2005,
14:20:45
User pressed "T", 14:21:02
User pressed "h", 14:21:03
User pressed spacebar, 14:21:04
User pressed backspace, 14:21:06
User pressed "e", 14:21:07
```

Figure 2: Snapshot of operation of the user modeling component.

Instructors have also the ability to manipulate the agents' behaviour with regard to the agents' on screen movements and gestures, as well as speech attributes such as speed, volume and pitch. Instructors may programmatically interfere to the agent's behaviour and the agent's reactions regarding the agents' approval or disapproval of a user's specific actions. This adaptation aims at enhancing the "affectiveness" of the whole interaction. Therefore, the system is enriched with an agent capable to express emotions and, as a result, enforces the user's temper to interact with more noticeable evidence in his/her behaviour.

Figure 3 illustrates a form where an instructor may change speech attributes.

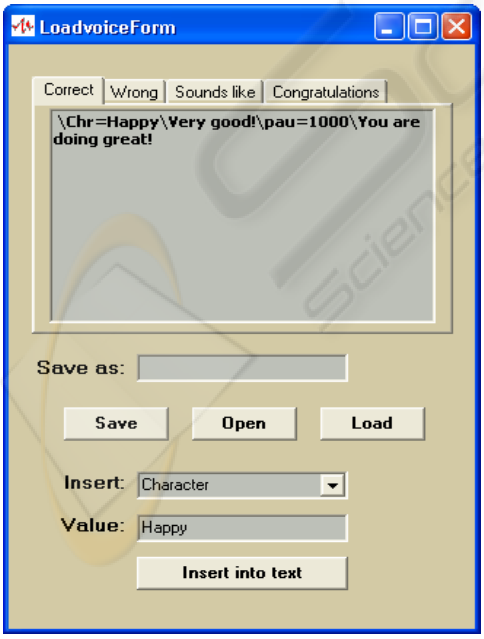


Mary (for Telephone)

```
Offset :51
Event WordPosition, TimeStamp :53676
Offset :57
Event WordPosition, TimeStamp :66496
Offset :63
Event WordPosition, TimeStamp :71620
Offset :69
Event WordPosition, TimeStamp :74082
Offset :72
Event WordPosition, TimeStamp :74824
Offset :74
Event WordPosition, TimeStamp :79952
Offset :79
Event WordPosition, TimeStamp :81876
Offset :82
Event WordPosition, TimeStamp :87000
Offset :87
Event WordPosition, TimeStamp :89564
Offset :91
Event WordPosition, TimeStamp :93410
Offset :96
Event WordPosition, TimeStamp :101098
Offset :105
```

Good morning tom. Here you can see a picture of a human body. Click on a part to view and hear clinical information.

Figure 3: Setting parameters for the voice of the tutoring character.



\Chr=Happy\Very good!\pau=1000\You are doing great!

Insert: Character

Value: Happy

Figure 4: Programming the behaviour of animated agents depending on particular students' actions.

Within this context the instructor may create and store for future use many kinds of voice tones such as happy tone, angry tone, whisper and many others depending on the need of a specific affective agent-user interaction. In some cases a user's actions may be rewarded with a positive message by the agent accompanied by a smile and a happy tone in the agent's voice, while in other cases a more austere behaviour may be desirable for educational needs. Figure 4 illustrates how an instructor may set possible actions for the agent in specific interactive situations while a user takes a test.

5 APPLICATION OF THE MULTI-CRITERIA MODEL

The input actions that were identified by the human experts during the second experimental study during requirements specification and analysis provided information for the emotional states that may occur while a user interacts with an educational system. These input actions are considered as criteria for evaluating all different emotions and selecting the one that seems more prevailing. More specifically, each emotion is evaluated first using only the criteria (input actions) from the keyboard and then only the criteria (input actions) from the microphone. In cases where both modals (keyboard and microphone) indicate the same emotion then the probability that this emotion has occurred is increased significantly. Otherwise, the mean of the values that have occurred by the evaluation of each emotion is calculated and the one with the higher mean is selected.

For the evaluation of each alternative emotion the system uses SAW for a particular category of users. This particular category comprises of the young (under the age of 19) and novice users (in computer skills). The likelihood for a specific emotion (happiness, sadness, anger, surprise, neutral and disgust) to have occurred by a specific action is calculated using the formula below:

$$\boxed{\frac{em_{1e_1 1} + em_{1e_1 2}}{2}}$$

$$\boxed{em_{1e_1 1} = w_{1e_1 k1} k_1 + w_{1e_1 k2} k_2 + w_{1e_1 k3} k_3 + w_{1e_1 k4} k_4}$$

$$\boxed{+ w_{1e_1 k5} k_5 + w_{1e_1 k6} k_6} \quad \text{(Formula 1)}$$

$$\boxed{em_{1e_1 2} = w_{1e_1 m1} m_1 + w_{1e_1 m2} m_2 + w_{1e_1 m3} m_3 + w_{1e_1 m4} m_4}$$

$$\boxed{+ w_{1e_1 m5} m_5 + w_{1e_1 m6} m_6 + w_{1e_1 m7} m_7} \quad \text{(Formula 2)}$$

$em_{1e_1 1}$ is the probability that an emotion has occurred based on the keyboard actions and $em_{1e_1 2}$ is the probability that refers to an emotional state using the users' input from the microphone. These probabilities result from the application of the decision making model of SAW and are presented in formulae 1 and 2 respectively. $em_{1e_1 1}$ and $em_{1e_1 2}$ take their values in [0,1].

In formula 1 the *k*'s from *k1* to *k6* refer to the six basic input actions that correspond to the keyboard. In formula 2 the *m*'s from *m1* to *m7* refer to the seven basic input actions that correspond to the microphone. These variables are Boolean. In each moment the system takes data from the bi-modal interface and translates them in terms of keyboard and microphone actions. If an action has occurred the corresponding criterion takes the value 1, otherwise its value is set to 0. The *w*'s represent the weights. These weights correspond to a specific emotion and to a specific input action and are acquired by the stereotype database. More specifically, the weights are acquired by the stereotypes about the emotions.

In order to identify the emotion of the user interacting with the system, the mean of the values that have occurred using formulae 1 and 2 for that emotion is estimated. The system compares the values from all the different emotions and determines whether an emotion is taking effect during the interaction. As an example we give the two formulae with their weights for the two modes of interaction that correspond to the emotion of happiness when a user (under the age of 19) gives the correct answer in a test of our educational application. In case of $em_{1e_1 1}$ considering the keyboard we have:

$$\boxed{em_{1e_1 1} = 0.4k_1 + 0.4k_2 + 0.1k_3 + 0.05k_4 + 0.05k_5 + 0k_6}$$

In this formula, which corresponds to the emotion of happiness, we can observe that the higher weight values correspond to the normal and quickly way of typing. Slow typing, often use of the backspace key and use of unrelated keys are actions with lower values of stereotypic weights. Absence of typing is unlikely to take place. Concerning the second mode (microphone) we have:

$$\boxed{em_{1e_1 2} = 0.06m_1 + 0.18m_2 + 0.15m_3 + 0.02m_4 + 0.14m_5 + 0.3m_6 + 0.15m_7}$$

In the second formula, which also corresponds to the emotion of happiness, we can see that the highest weight corresponds to *m6* which refers to the 'speaking of a word from a specific list of words showing an emotion' action. The empirical study gave us strong evidence for a specific list of words. In the case of words that express happiness, these words are more likely to occur in a situation where a novice young user gives a correct answer to the system. Quite high are also the weights for variables *m2* and *m3* that correspond to the use of exclamations by the user and to the raising of the user's voice volume. In our example the user may do something orally or by using the keyboard or by a combination of the two modes. The absence or presence of an action in both modes will give the Boolean values to the variables *k1…k6* and *m1…m7*.

A possible situation where a user would use both the keyboard and the microphone could be the following: The specific user knows the correct answer and types in a speed higher than the normal speed of writing. The system confirms that the answer is correct and the user says a word like 'bravo' that is included in the specific list of the system for the emotion of happiness. The user also speaks in a higher voice volume. In that case the variables k1, m3 and m6 take the value 1 and all the others are zeroed. The above formulas then give us

$$em_{1e_1 1} = 0.4 * 1 = 0.4 \quad \text{and}$$

$$em_{1e_1 2} = 0.15 * 1 + 0.3 * 1 = 0.45 .$$

In the same way the system then calculates the corresponding values for all the other emotions using other formulae. For each basic action in the educational application and for each emotion the corresponding formula have different weights deriving from the stereotypical analysis of the empirical study. In our example in the final comparison of the values for the six basic emotions the system will accept the emotion of happiness as the most probable to occur.

6 EVALUATION OF THE MULTI-CRITERIA MODEL

In section 5 we have described how the system incorporates the multi-criteria decision making theory SAW and uses stereotypic models derived from empirical studies in order to make a multi-criteria decision about the emotions that occur during the educational human-computer interaction. Each mode uses user stereotypes with specific weights for each input action and produces values for each one of the six basic emotions in our study. Correspondingly, each mode produces hypotheses for the six basic emotions and classifies them by their probabilities of occurrence. The final conclusion on the user's emotion is based on the conjunction of evidence from the two modes using SAW.

The 50 medical students that were involved in the first phase of the empirical study in section 3.1 were also used in the second phase of the empirical study for the evaluation of the multi-criteria emotion recognition system. In this section we present and compare results of successful emotion recognition in audio mode, keyboard mode and the two modes combined. For the purposes of our study the whole interaction of all users with the educational application was video recorded. Then the videos collected were presented to the users that participated the experiment in order to perform emotion recognition for themselves with regard to the six emotional states, namely happiness, sadness, surprise, anger, disgust and the neutral emotional state. The students as observers were asked to justify the recognition of an emotion by indicating the criteria that s/he had used in terms of the audio mode and keyboard actions. Whenever a participant recognized an emotional state, the emotion was marked and stored as data in the system's database. Finally, after the completion of the empirical study, the data were compared with the systems' corresponding hypothesis in each case an emotion was detected.

Table 1 illustrates the percentages of successful emotion recognition of each mode after the incorporation of stereotypic weights and the combination through the multi-criteria approach.

Provided the correct corresponding emotions for each situation and by each user we were able to come up with conclusions about the efficacy of our systems' emotion recognition ability. Indeed, the results presented in table 1 indicate that the incorporation of user stereotypes as well as the application of the multi-criteria model lead our system to noticeable improvements in its ability to recognize emotional states of users successfully.

Table 1: Recognition of emotions using stereotypes and SAW theory.

Using Stereotypes and SAW	
Emotions	Multi-criteria bi-modal recognition
Neutral	46%
Happiness	64%
Sadness	70%
Surprise	45%
Anger	70%
Disgust	58%

7 CONCLUSIONS

In this paper we have described an affective educational application that recognizes students' emotions based on their words and actions that are identified by the microphone and the keyboard, respectively. The system uses an innovative approach that combines evidence from the two modes of interaction based on user stereotypes and a multi-criteria decision making theory.

For requirements analysis and the effective application of the particular approach two different experimental studies have been conducted. The experimental studies involved real end users as well as human experts. In this way the application of the multi-criteria model in the design of the system was more accurate as it was based on facts from real users' reasoning process.

Finally, the approach for emotion recognition was evaluated. More specifically, some users interacted with the educational application and their interaction was video recorded. The videos were then presented to the same users, who were asked to comment on their emotion. The emotions the users identified were compared to the emotions identified by the system. This comparison revealed that the system could adequately identify the users' emotion. However, its hypotheses were more accurate when there was a combination of the evidence from two different modes using the multi-criteria decision making theory.

In future work we plan to improve our system by the incorporation of stereotypes concerning users of several ages, educational backgrounds and computer knowledge levels. Moreover, there is ongoing research work in progress that exploits a third mode of interaction, visual this time (Stathopoulou & Tsihrintzis, 2005), to add information to the system's database and complement the inferences of the user modelling component about users' emotions. The third mode is going to be integrated to our system by adding cameras and also providing the appropriate software, as for a future work.

ACKNOWLEDGEMENTS

Support for this work was provided by the General Secretariat of Research and Technology, Greece, under the auspices of the PENED-2003 program.

REFERENCES

Elfenbein, H.A., Ambady, N., 2003. When Familiarity Breeds Accuracy. Cultural Exposure and Facial Emotion Recognition, Journal of Personality and Social Psychology, Vol. 85, No. 2, pp. 276–290.

Fishburn, P.C., 1967. Additive Utilities with Incomplete Product Set: Applications to Priorities and Assignments, Operations Research.

Goleman, D., 1995. Emotional Intelligence, Bantam Books, New York .

Hudlicka, E., 2003. To feel or not to feel: The role of affect in human-computer interaction. International Journal of Human-Computer Studies, Elsevier Science, London, pp. 1-32.

Hwang, C.L., Yoon, K., 1981. Multiple Attribute Decision Making: Methods and Applications. Lecture Notes in Economics and Mathematical Systems 186, Springer, Berlin/Heidelberg/New York.

Kay, J., 2000. Stereotypes, student models and scrutability. In: G. Gauthier, C. Frasson and K. VanLehn (eds.): Proceedings of the *Fifth International Conference on Intelligent Tutoring Systems*, Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, Vol. 1839, pp. 19-30.

Kobsa, A., Koenemann, J., Pohl, W., 2001. Personalized hypermedia presentation techniques for improving online customer relationships. The Knowledge Engineering Review, vol. 16, pp. 111-115.

Moriyama, T., Ozawa, S., 2001. Measurement of Human Vocal Emotion Using Fuzzy Control. Systems and Computers in Japan, Vol. 32, No. 4.

Moriyama, T., Saito, H., Ozawa, S., 2001. Evaluation of the Relation between Emotional Concepts and Emotional Parameters in Speech. Systems and Computers in Japan, Vol. 32, No. 3.

Oviatt, S., 2003. User-modeling and evaluation of multimodal interfaces. Proceedings of the IEEE, Institute of Electrical and Electronics Engineers, pp. 1457-1468.

Pantic, M., Rothkrantz, L.J.M., 2003. Toward an affect-sensitive multimodal human-cumputer interaction. Vol. 91, Proceedings of the IEEE, Institute of Electrical and Electronics Engineers, pp. 1370-1390.

Picard, R.W., 2003. Affective Computing: Challenges. Int. Journal of Human-Computer Studies, Vol. 59, Issues 1-2, pp. 55-64.

Picard, R.W., Klein, J., 2002. Computers that recognise and respond to user emotion: theoretical and practical implications. Interacting with Computers 14, pp. 141-169.

Rich, E., 1983. Users are individuals: individualizing user models. International Journal of Man-Machine Studies 18, pp. 199-214.

Stathopoulou, I.O., Tsihrintzis, G.A., 2005. Detection and Expression Classification System for Face Images (FADECS), *IEEE Workshop on Signal Processing Systems*, Athens, Greece.