Preface

Tremendous advances in intelligent paradigms, such as artificial neural systems, machine learning, evolutionary computing, web-based systems and intelligent agent systems, have immensely contributed in the field of interactive systems. The personalization of intelligent interactive systems has become reality. Researchers are working on fusing adaptivity and learning in interactive systems. Intelligent recommender systems, intelligent tutoring systems, intelligent web-shops, intelligent interactive TV, intelligent mobile services, affective systems and narrative environments constitute typical intelligent interactive systems. However, the list of intelligent interactive systems is not exhausted to the aforementioned applications, as research and developments in the area expand to many more applications.

The main aim of this book is to report a sample of the most recent advances in the field of intelligent interactive systems in knowledge-based environments. This book consists of ten chapters reflecting the theoretical foundations as well as typical applications in the area of intelligent interactive systems.

We wish to express our gratitude to the authors and reviewers for their wonderful contributions. Thanks are due to Springer-Verlag for their editorial support. We would also like to express our sincere thanks to Ms Sridevi Ravi for her wonderful editorial support. We believe that this book would help in creating interest among researchers and practitioners towards realizing human-like machines. This book would prove useful to the researchers, professors, research students and practitioners as it reports novel research work on challenging topics in the area of intelligent interactive systems. Moreover, special emphasis has been put on highlighting issues concerning the development process of such complex interactive systems, thus revisiting the difficult issue of knowledge engineering of such systems. In this way, the book aims at providing the readers with a better understanding of how intelligent interactive systems can be successfully implemented to incorporate recent trends and advances in theory and applications of intelligent systems.

Greece	*Maria Virvou*
Australia	*Lakhmi C. Jain*

2

Requirements Analysis and Design of an Affective Bi-Modal Intelligent Tutoring System: The Case of Keyboard and Microphone

Efthymios Alepis[1], Maria Virvou[1], and Katerina Kabassi[2]

[1] Department of Informatics, University of Piraeus, 80 Karaoli and Dimitriou St., 18534 Piraeus, Greece, talepis@unipi.gr, mvirvou@unipi.gr
[2] Department of Ecology and the Environment, Technological Educational Institute of the Ionian Islands, 2 Kalvou Sq., 29100 Zakynthos, Greece, kkabassi@teiion.gr

Summary. This chapter presents an affective bi-modal Intelligent Tutoring System (ITS) with emphasis on the early stages of its creation. Affective ITSs are expected to provide a more human-like interaction between students and educational software. The ITS is named Edu-Affe-Mikey and its tutoring domain is Medicine. The two modes of interaction presented in this chapter concern the keyboard and the microphone. Emotions of students are recognised by each modality separately and then, evidence from the two modalities is combined through a decision making theory. After emotion recognition has been performed Edu-Affe-Mikey adapts dynamically its tutoring behaviour to an appropriate emotion of an animated tutoring agent. In this respect an affective and adaptive interaction is achieved in the interactions of the student with the ITS, by performing both affect recognition and affect generation.

2.1 Introduction

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 [9]. Indeed, as Coles [3] points out poor learning can produce negative emotions; negative emotions can impair learning; and positive emotions can contribute to learning achievement and vice versa. Therefore, a way of improving the learning process is recognising the users' emotions by observing them during their engagement with the educational software and then adapting its interaction to their emotional state.

Ahn and Picard [1] point out that affective biases from affective anticipatory rewards can be applied for improving the speed of learning, regulating the trade-of between exploration and exploitation in learning more efficiently. In Kim et al. [13] as well as in Burleson [2] it is suggested that pedagogical

E. Alepis et al.: *Requirements Analysis and Design of an Affective Bi-Modal Intelligent Tutoring System: The Case of Keyboard and Microphone*, Studies in Computational Intelligence (SCI) **104**, 9–24 (2008)
www.springerlink.com

agents with emotional interaction with learners should be used in learning environments. The findings of these studies imply that the emotional states of learning companions can be utilised to optimise students' motivation, learning and perseverance. For this purpose different approaches have been proposed in the literature. For example, Kapoor and Picard [12] propose a multi-sensor affect recognition system for children trying to solve an educational puzzle on the computer.

However, a problem that arises in such approaches in the way that these interaction modalities are combined in order to improve the accuracy and overall performance of emotion recognition in an Intelligent Tutoring System. In fact, the mathematical tools and theories that have been used for affect recognition can drive a classification of affect recognisers. Such a classification has been made in [14]. Liao and his colleagues, have classified affect recognisers into two groups on the basis of the mathematical tools that these recognisers have used: (1) the first group using traditional classification methods in pattern recognition, including rule-based systems, discriminate analysis, fuzzy rules, case-based and instance-based learning, linear and nonlinear regression, neural networks, Bayesian learning and other learning techniques. (2) The second group of approaches using Hidden Markov Models, Bayesian networks etc. Indeed, a recent piece of research uses the above approaches for the integration of audio-visual evidence [24]. Specifically, for person-dependent recognition, they apply the voting method to combine the frame-based classification results from both audio and visual channels. For person-independent test, they apply multistream hidden Markov models (HMM) to combine the information from multiple component streams. In contrast with these approaches we have used a decision making method in order to make decisions about which modalities should be taken to account while trying to recognise different emotional states. For this purpose two empirical studies were conducted and the results of these studies were used as criteria for the proposed decision making model.

A rather promising approach that has not attracted adequate attention despite of the potential benefits of its application is multi-criteria decision making theories. The main advantages of this approach derive from the fact that user-computer interaction is, by nature, multi-criteria-based. However, multi-criteria decision making requires several development steps for their application. More specifically, there is a need for experimental studies for selecting the criteria, estimating their weights of importance, test the models effectiveness, etc. Therefore, special emphasis should be given during requirements analysis and the design of an Intelligent Tutoring System (ITS). Indeed, as Virvou [23] point out, both students and tutors have to be involved in many phases of the software life-cycle in order to ensure that the software is really useful to students.

In view of the above, in this chapter, we discuss the requirements analysis and design of an ITS called Edu-Affe-Mikey. The proposed ITS is targeted to first-year medical students and its main characteristic is that it can adapt its interaction to each user's emotional state. For this purpose, the system uses a multi-criteria decision making method called Simple Additive Weighting (SAW) [8, 10] for combining two modes of interaction, namely keyboard and microphone. More specifically, SAW is used 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.

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 system was more accurate as it was based on facts from real users' reasoning process. The main aim of the first study was to capture videos with real user's interaction with the system and as a result finding out how users express their emotions while interacting with educational software. The second empirical study involved human experts. These experts were asked to define the criteria that they usually use to perform emotion recognition of his/her students during the teaching course as well as their weights of importance.

The main body of this chapter is organised as follows: In Sect. 2.2 we present the multi-criteria decision making method called SAW. In Sect. 2.3 we present the first empirical study during requirements specification and analysis. In Sect. 2.4 we present the second empirical study for defining the criteria that are taken into account while performing emotion recognition. In Sect. 2.5 we give an overall description of the system and in Sect. 2.6 we give information of how the decision making method has been applied in the system for combining evidence from two different modes and select the user's emotion. Finally, in Sect. 2.7 we give the conclusions drawn by this work and discuss ongoing work.

2.2 SAW: A Simple Decision Making Method

A multi-criteria decision problem is a situation in which, having defined a set A of actions and a consistent family F of n criteria $g_1,\ g_2, \dots,\ g_n$ $(n \geq 3)$ on A, one wishes to rank the actions of A from best to worst and determine a subset of actions considered to be the best with respect to F [22].

When the DM must compare two actions a and b, there are three cases that can describe the outcome of the comparison: the DM prefers a to b, the DM is indifferent between the two or the two actions are incompatible.

The traditional approach is to translate a decision problem into the optimisation of some function g defined on A. If $g(a) > g(b)$ then the DM prefers a to b, whereas if $g(a) = g(b)$ then the DM is indifferent between the two.

For the calculation of the function g many different decision making theories are introduced in the literacy. The Simple Additive Weighting (SAW) [8,10] 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:

$$U(X_j) = \sum_{i=1}^{n} w_i x_{ij} \tag{2.1}$$

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

2.3 Requirements Analysis for Affective Bi-Modal Interaction

Requirement specification and analysis in the affective bi-modal intelligent tutoring system resulted from an empirical study. The main aim of this study was to find out how users express their emotions through a bi-modal interface that combines voice recognition and input from keyboard. 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. These data were used for identifying the criteria that are taken into account when evaluating an emotion and a database was built for acquiring the weights of importance of these criteria.

The empirical study also revealed that the users would also appreciate if the system adapted its interaction to the users' emotional 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.

One conclusion concerning the combination of the two modes in terms of emotion recognition is that the two modes are complementary to each other to a high extent. In many cases the system can generate a hypothesis about the emotional state of the user with a higher degree of certainty if it takes into account evidence from the combination of the two modes rather than one mode. Happiness has positive effects and anger or boredom have negative effects that may be measured and processed properly in order to give information that is used during 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.

2.3.1 Affect Perception Based on Speech

A small percentage of the participants say something with anger when they make a spelling mistake. However from the participants who do say something, 74% of them consider themselves having little or moderate (2–6 months of practice) computer knowledge. One very interesting result is that people seem to be more expressive when they have negative feelings (Fig. 2.1), than when they have positive feelings.

Another important conclusion coming up from the study is that when people say something either expressing happiness or anger it is highly supported from the changes of their voice. They may raise the tone of their voice or more probably they may change the pitch of their voice, Fig. 2.2.

Moreover it is interesting to notice that a very high percentage (85%) of young students who are also inexperienced with computers find the oral mode of interaction very useful.

2.3.2 Affect Perception Based on Keyboard Actions

While using the keyboard most of the participants agree that when they are nervous the possibility of making mistakes increases rapidly. This is also the

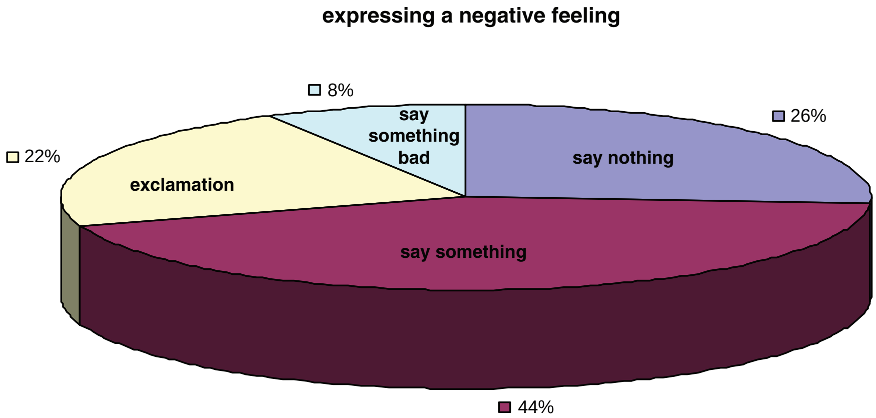

Fig. 2.1. Speech reactions in negative feelings

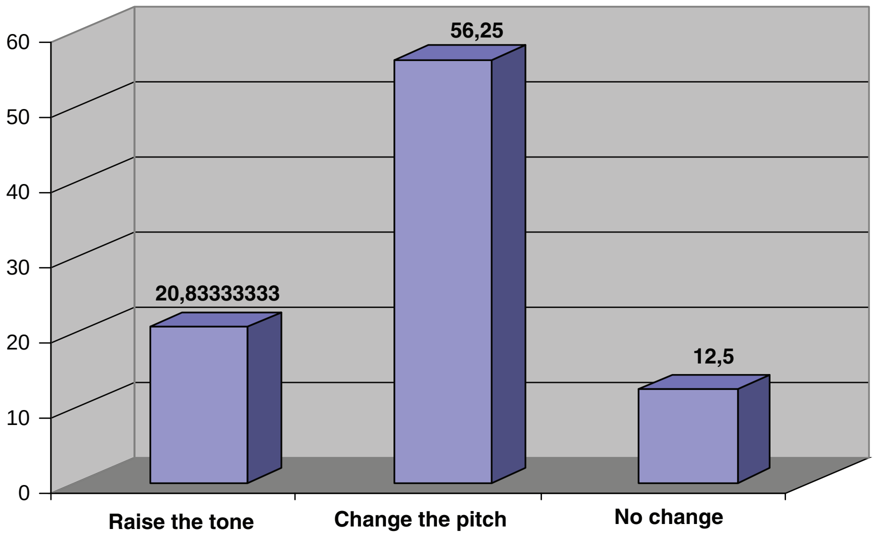

Fig. 2.2. Changes in voice when expressing a feeling (either positive or negative)

case when they have negative feelings. Mistakes in typing are followed by many backspace-key keyboard strokes and concurrent changes in the emotional state of the user in a percentage of 82%. Yet users under 20 years old seem to be more prone to making even more mistakes as a consequence of an initial mistake and lose their concentration while interacting with an application (67%). Students also admit that when they are angry the rate of mistakes increases, the rate of their typing becomes slower 62% (on the contrary, when they are happy they type faster 70%) and the keystrokes on the keyboard become harder (65%). Of course the pressure of the user's fingers on the keyboard is something that can not be measured without the appropriate hardware. Similar effects to the keyboard were reported for the neutral emotional state instead of anger.

2.4 Specification and Analysis of Multiple Criteria

Any decision making method requires prior to its application the specification of some criteria. Therefore, an empirical study was conducted in order to locate the criteria that human experts take into account while performing emotion recognition. 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 seven 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.

2.5 Overview of the System

In this section, the overall functionality and emotion recognition features of our system, Edu-Affe-Mikey is described. 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 Fig. 2.3. The animated agent is present in these modes to make the interaction more human-like.

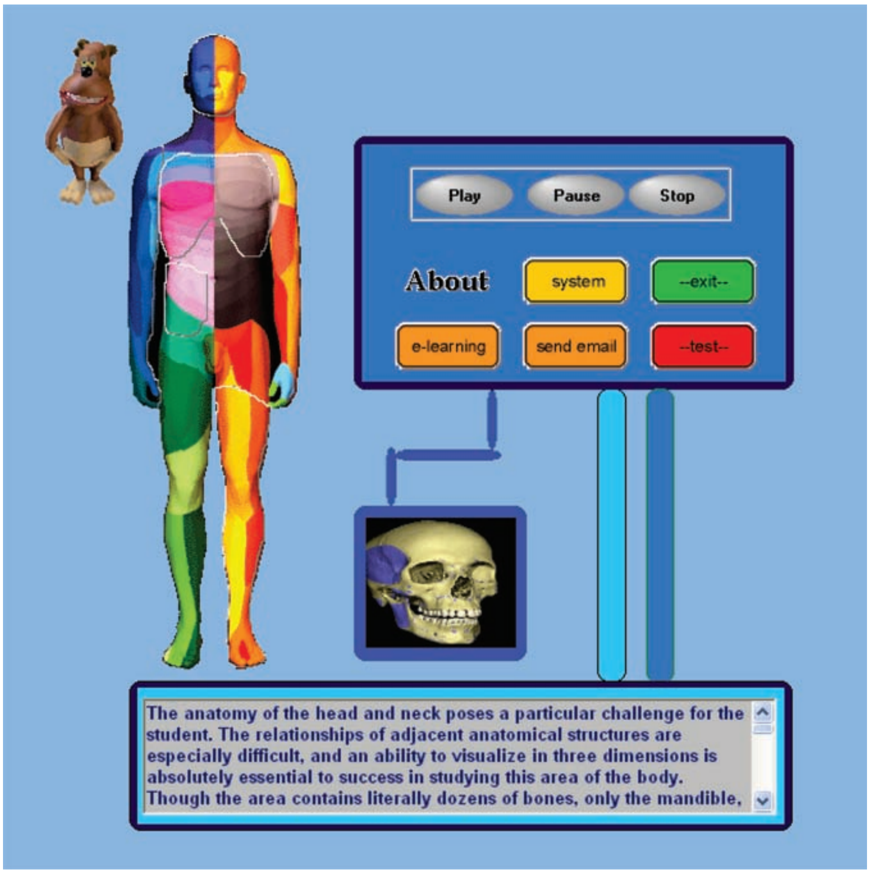

Fig. 2.3. A screen-shot of theory presentation in Edu-Affe-Mikey educational application

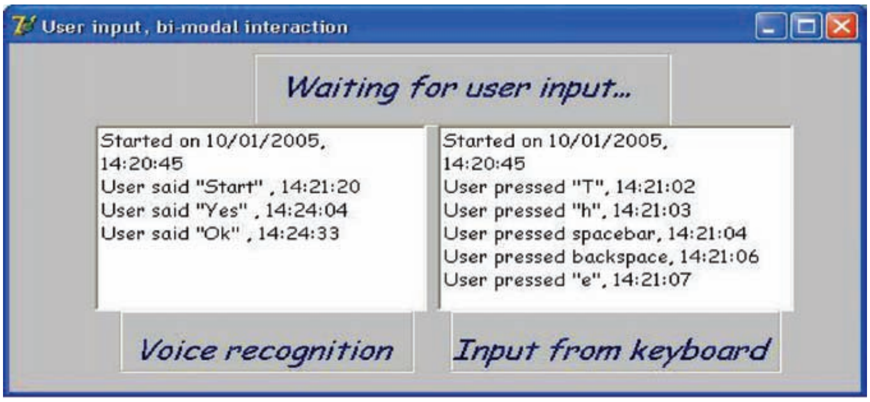

Fig. 2.4. Snapshot of operation of the user modelling component

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 2.4 illustrates the 'monitoring' component that records the user's input from the microphone and the keyboard and the exact time of each event.

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 effectiveness 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 2.5 illustrates a form where an instructor may change speech attributes. Within this context the instructor may create and store for future use (Fig. 2.6) 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 2.7 illustrates how an instructor may set possible actions for the agent in specific interactive situations while a user takes a

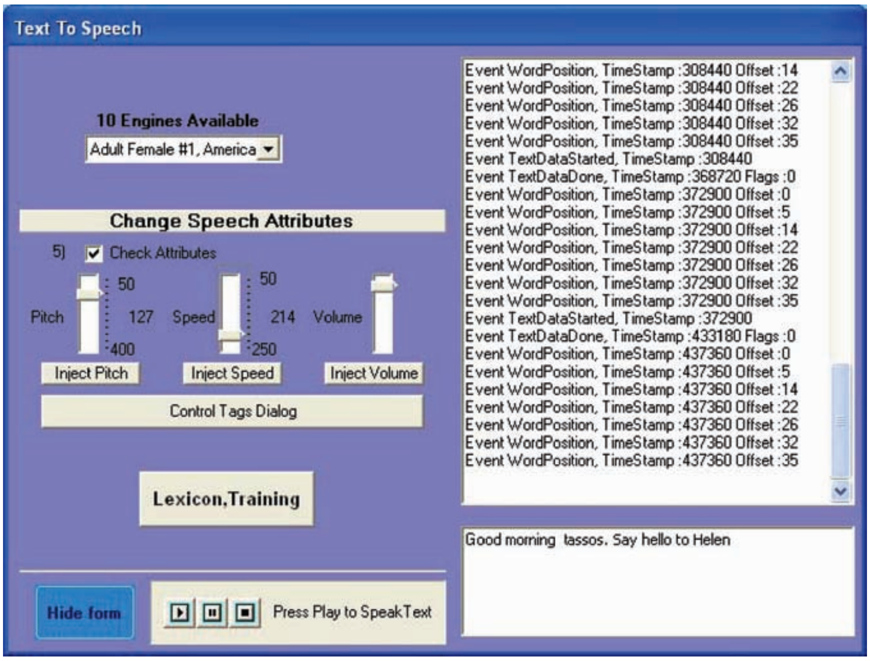

Fig. 2.5. Setting parameters for the voice of the tutoring character

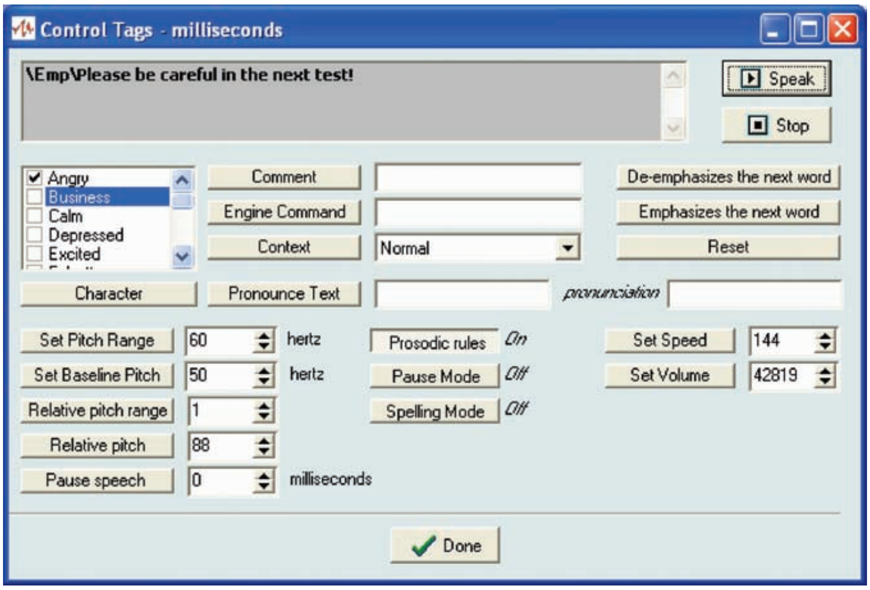

Fig. 2.6. Speech control tags. Default speech attributes for emotional states

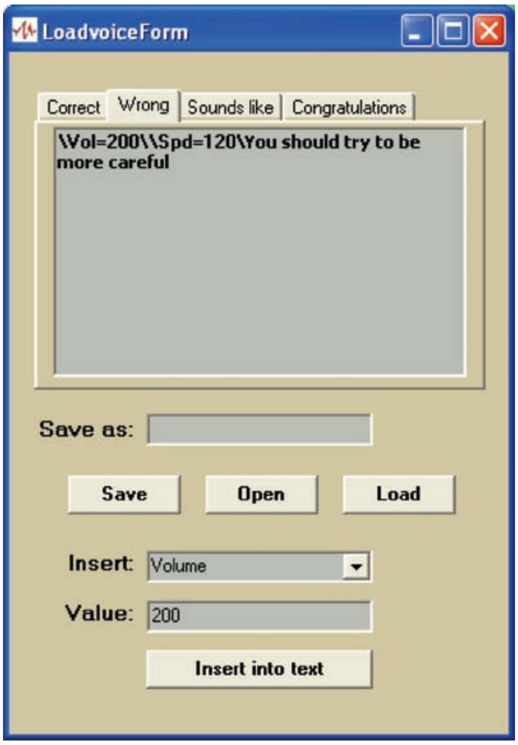

Fig. 2.7. Programming the behaviour of animated agents depending on particular students' actions

test. Instructors have also the potential of authoring their own tests for specific parts of the theory as to adapt the educational process to each individual student (Fig. 2.8). Finally Fig. 2.9 illustrates the interaction between a student and the animated agent while a student is taking a test. The system monitors all user actions and in cases where a student's action and a specified by an instructor event are correlated, the agent performs a pre-programmed behavior (visual and oral) that corresponds to that event.

2.6 Application of the Decision Making Method

For the evaluation of each alternative emotion the system uses as criteria the input actions that relate with the emotional states that may occur while a user interacts with an educational system. These input actions were identified by the human experts during the second experimental study and are considered as criteria for evaluating all different emotions and selecting the one that seems to be more prevailing. More specifically, 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

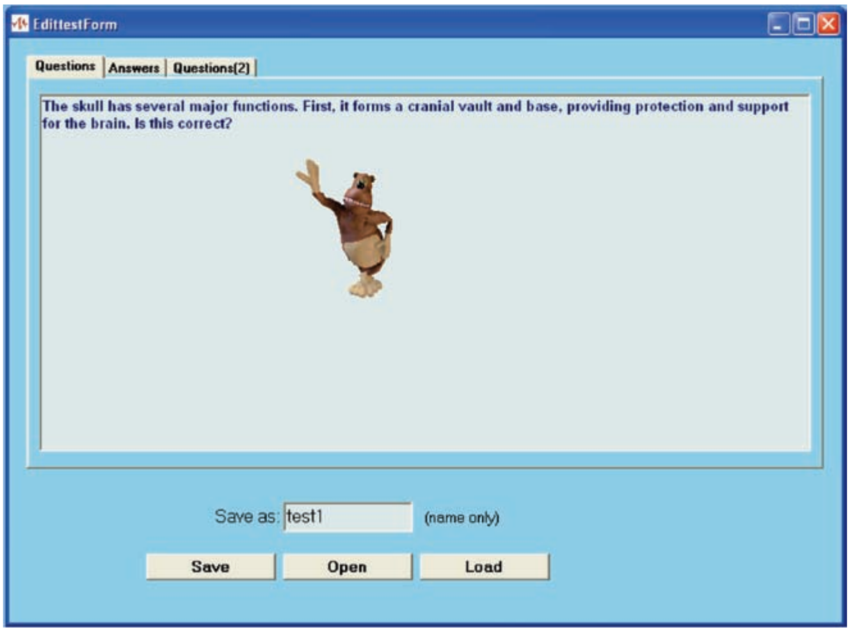

Fig. 2.8. Authoring tests for the educational application

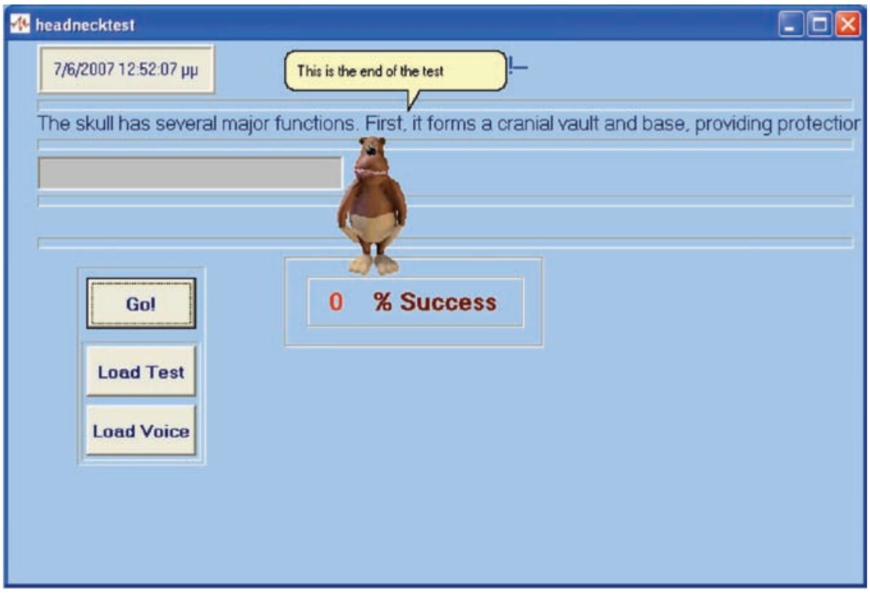

Fig. 2.9. The affective agent interacts with the student, correspondingly to the student's actions

a specific emotion (happiness, sadness, anger, surprise, neutral and disgust) to have occurred by a specific action is calculated using the formula below:

$$\frac{em_{1e_11} + em_{1e_12}}{2}$$

$$em_{1e_11} = w_{1e_1k1}k_1 + w_{1e_1k2}k_2 + w_{1e_1k3}k_3 + w_{1e_1k4}k_4 + w_{1e_1k5}k_5 + w_{1e_1k6}k_6 \quad (2.2)$$

$$em_{1e_12} = w_{1e_1m1}m_1 + w_{1e_1m2}m_2 + w_{1e_1m3}m_3 + w_{1e_1m4}m_4 + w_{1e_1m5}m_5 + w_{1e_1m6}m_6 + w_{1e_1m7}m_7 \quad (2.3)$$

em_{1e_11} is the probability that an emotion has occurred based on the keyboard actions and em_{1e_12} 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 (2.2) and (2.3) respectively. em_{1e_11} and em_{1e_12} take their values in [0,1].

In (2.2) the k's from *k1* to *k6* refer to the six basic input actions that correspond to the keyboard. In (2.3) 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 database.

In order to identify the emotion of the user interacting with the system, the mean of the values that have occurred using (2.2) and (2.3) 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_11} considering the keyboard we have:

$$em_{1e_11} = 0.4k_1 + 0.4k_2 + 0.1k_3 + 0.05k_4 + 0.05k_5 + 0k_6$$

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 weights. Absence of typing is unlikely to take place. Concerning the second mode (microphone) we have:

$$em_{1e_12} = 0.06m_1 + 0.18m_2 + 0.15m_3 + 0.02m_4 + 0.14m_5 + 0.3m_6 + 0.15m_7$$

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 $k1 \ldots k6$ and $m1 \ldots m7$.

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 formulae then give us $em_{1e_1 1} = 0.4{*}1 = 0.4$ 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 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.

2.7 Conclusions and Ongoing Work

In this chapter we have described an affective educational application that recognises 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 using a multi-criteria decision making theory.

The main advantage of the proposed approach is that the whole process is based on experimental studies in which real users participate. Therefore, potential users of the software and human experts have participated in an empirical study. The analysis of the results of the empirical study gave evidence for the design of the affective module of the educational application. As the affective module uses a multi-criteria decision making theory for combining evidence from the two different modes, the results of the empirical study were used for selecting the criteria and estimating their weights of importance.

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 [21], 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. A part of this work is reported in the International Journal of Intelligent Support Technologies, IOS Press, Volume 1, Number 3, 2008.

References

1. Ahn, H., Picard, R.W. Affective-cognitive learning and decision making: A motivational reward framework for affective agents, Lecture Notes in Computer Science, Vol. 3784 LNCS (2005), pp. 866–873
2. Burleson, W. Affective learning companions: strategies for empathetic agents with real-time multimodal affective sensing to foster meta-cognitive and meta-affective approaches to learning, motivation, and perseverance. PhD. Thesis, 2006
3. Coles, Gerald. Literacy, emotions, and the brain. Reading Online, March 1999. http://www.readingonline.org/critical/coles.html
4. Damasio, A.R. Descartes' Error: Emotion, Reason and the Human Brain, Grosset/Putnam, New York (1994)
5. Davidson, R.J., Scherer, K.R., Goldsmith, H.H. Handbook of Affective Sciences, Oxford, New York (2003)
6. Davidson, R.J., Pizzagalli, D., Nitschke, J.B., Kalin, N.H. Parsing the subcomponents of emotion and disorders of emotion: perspectives from affective neuroscience, In Handbook of Affective Sciences, Davidson, R.J., Scherer, K.R., Goldsmith, H.H. (eds.) (2003)
7. Elfenbein, H.A., Ambady, N. When familiarity breeds accuracy. Cultural exposure and facial emotion recognition, Journal of Personality and Social Psychology, 85(2) (2003) 276–290
8. Fishburn, P.C. Additive utilities with incomplete product set: Applications to priorities and assignments, Operations Research 15 (1967) 537–542
9. Goleman, D. Emotional Intelligence, Bantam Books, New York (1995)
10. Hwang, C.L., Yoon, K. Multiple attribute decision making: Methods and applications, Lecture Notes in Economics and Mathematical Systems 186, Springer, Berlin, Heidelberg, New York (1981)
11. Isbister, K., Höök, K. Evaluating affective interactions: Innovative approaches and future direction, Special Issue of the International Journal of Human Computer Studies (2006)
12. Kapoor, A. Picard, R.W. Multimodal affect recognition in learning environments, Proceedings of the 13th annual ACM international conference on Multimedia, pp. 677–682 (2005)

13. Kim, Y., Baylor, A.L., Shen, E. Pedagogical agents as learning companions: The impact of agent emotion and gender, Journal of Computer Assisted Learning, 23(3) (2007) 220–234
14. Liao, W., Zhang, W., Zhu, Z., Ji, Q, Gray, W.D. Toward a decision-theoretic framework for affect recognition and user assistance, International Journal of Human-Computer Studies 64 (2006) 847–873
15. Moriyama, T., Ozawa, S. Measurement of human vocal emotion using fuzzy control, Systems and Computers in Japan 32(4) (2001)
16. Oviatt, S. User-modeling and evaluation of multimodal interfaces, Proceedings of the IEEE, Institute of Electrical and Electronics Engineers (2003) 1457–1468
17. Pantic, M., Rothkrantz, L.J.M. Toward an affect-sensitive multimodal human-computer interaction, Proceedings of the IEEE, Institute of Electrical and Electronics Engineers 91 (2003) 1370–1390
18. Picard, R.W. Affective computing: Challenges, International Journal of Human-Computer Studies 59(1–2) (2003) 55–64
19. Rich, E. Users are individuals: Individualizing user models, International Journal of Man-Machine Studies 18 (1983) 199–214
20. Stathopoulou, I.O., Tsihrintzis, G.A. Detection and expression classification system for face images (FADECS), IEEE Workshop on Signal Processing Systems, Athens, Greece (2005)
21. Stathopoulou, I.O., Tsihrintzis, G.A. Automated processing and classification of face images for human-computer interaction applications, Intelligent Interactive Systems in Knowledge-based Environments, 104 (2007)
22. Vincke, P. Multicriteria Decision-Aid. Wiley, New York (1992)
23. Virvou, M. Software engineering of environmental computer-assisted learning, Proceedings of the International Conference on Ecological Protection of the Planet Earth, 2 (2001) 927–934
24. Zeng, Z.H., Tu, J.L., Liu, M., Huang, T.S., Pianfetti, B., Roth, D., Levinson, S. Audio-visual affect recognition, IEEE Transactions on Multimedia 9(2) (2007) 424–428