"How do you know that I don't understand?" A look at the future of intelligent tutoring systems

Abdolhossein Sarrafzadeh[1], Samuel Alexander[1], Farhad Dadgostar[1], Chao Fan[1], Abbas Bigdeli[2]
[1]*Institute of information and Mathematical Sciences, Massey University, Auckland, New Zealand*
[2]*Safeguarding Australia Program, National ICT Australia (NICTA) Ltd, QLD Lab, Brisbane, Australia*

Abstract

Many software systems would significantly improve performance if they could adapt to the emotional state of the user, for example if Intelligent Tutoring Systems (ITSs), ATM's, ticketing machines could recognise when users were confused, frustrated or angry they could guide the user back to remedial help systems so improving the service. Many researchers now feel strongly that ITSs would be significantly enhanced if computers could adapt to the emotions of students. This idea has spawned the developing field of Affective Tutoring Systems (ATSs): ATSs are ITSs that are able to adapt to the affective state of students. The term "Affective Tutoring System" can be traced back as far as Rosalind Picard's book *Affective Computing* in 1997.

This paper presents research leading to the development of *Easy with Eve*, an ATS for primary school mathematics. The system utilises a network of computer systems, mainly embedded devices to detect student emotion and other significant bio-signals. It will then adapt to students and displays emotion via a lifelike agent called Eve. Eve's tutoring adaptations are guided by a case-based method for adapting to student states; this method uses data that was generated by an observational study of human tutors. This paper presents the observational study, the case-based method, the ATS itself and its implementation on a distributed computer systems for real-time performance, and finally the implications of the findings for Human Computer Interaction in general and e-learning in particular. Web-based applications of the technology developed in this research are discussed throughout the paper.

Keywords: Affective tutoring systems, lifelike agents, emotion detection, facial expressions, human computer interaction, affective computing

1. Introduction

Intelligent Tutoring Systems (ITS) provide individualised instruction, by being able to adapt to the knowledge, learning abilities and needs of each individual student. Existing ITS build a model of the student's current state of knowledge and individualise instruction based on that model (Sarrafzadeh, 2002). Intelligent tutoring systems offer many advantages over the traditional classroom scenario: they are always available, non-judgmental, and provide tailored feedback (Anderson, Corbett, Koedinger, Pelletier, 1995; Johnson, Shaw, Marshall and LaBore, 2003; Self, 1990). They have been proven effective, resulting in increased learning (Aleven, Koedinger and Cross, 1999; Aleven and Koedinger, 2000; Anderson et al., 1995; Conati and VanLehn, 2001). However, they are still not as effective as one-on-one human tutoring. We believe that an important factor in the success of human one-to-one tutoring is the tutor's ability to identify and respond to affective cues.

Human communication is a combination of both verbal and nonverbal interactions. Human teachers may not know the knowledge state of all the students, however, by looking at the facial expressions, body gesture and other non-verbal cues, a human teacher may change his/her teaching strategy or take some other appropriate measures. Puzzled or bored faces might mean that there is no sense in continuing with the current teaching strategy. When it comes to one to one tutoring these cues may be even more useful.

As teacher shortages loom in both rural and urban schools, especially in mathematics and science, any contribution to alleviate this problem becomes important (The Urban Teacher Collaborative, 2000).We are proposing a new generation of intelligent tutoring systems that model not only the knowledge state of the student but also his/her cognitive and emotional state. Estimating the emotional state of a learner may involve analysing facial expressions, voice tone, heartbeat and other bio-signals.

This paper discusses how intelligent tutoring systems can be enhanced to include learners' affective state in its student model. It gives an overview of a new type of ITS proposed by the authors, *Affective Tutoring Systems (ATS)*, which detect non-verbal behaviour and use this information to individualise interactions with the student. For the ATS systems to be effective the non-verbal behaviour are to be detected in real-time. In order to achieve real-time performance, a network of computer systems mostly embedded is required to pre-process various bio-signals in a distributed fashion. This paper presents an implementation platform compromising low-cost embedded systems utilised to improve the overall system speed and performance. This paper includes a discussion of two primary research foci. First, we introduce a facial expression and gesture analysis system that forms the basis of affective state detection. We then present an affective mathematics tutoring system Easy with Eve which detects affective state detection and a case based reasoning system to react to the emotions of the learner through a life-like agent called Eve.

1.1. Networked applications of emotion detection

The ability to detect nonverbal behaviour in real-time has tremendous implications for networked applications. Detecting the behaviour of individuals in a group or the collective behaviour of the group and feeding this information back to the system can help the system adapt its behaviour to not only the individuals but also the average mood of the users. On the other hand, in a distributed processing environment, the nonverbal behaviour of an audience can be detected in real-time by processing the individual faces in parallel. This information can then be fed back to performers for improving their interaction with the audience. An example of such a scenario would be feedback to panellists in a political debate. In fact the technology developed for "Easy with Eve" ATS is being used to develop a performance stage interface intended to help actors in live shows adapt to the mood of the audience.

Satisfied customers is key to the success of any business. Another important area of application of the technology described in this paper is sales and marketing applications. Emotion detection techniques used in Easy with Eve are being used in the development of an online sales assistant that is capable of guiding buyers to suitable products in an online sales environment. It is believed that a sales assistant that is capable of providing such service will result in increased sales (Shergil, Sarrafzadeh, Diegel, Shekar, 2004). Call centres can greatly benefit from emotion detection from voice. Detecting emotions from recorded conversations of call centre staff with customers can be used to guide supervisors to recordings where frustration or anger is detected in customer's voice. These applications are only some of business applications of automatic emotion detection.

2. A closer look at ITS

Computer based instructional material tends to be presented in the didactic mode, which means information is presented to the student without any feedback about the level of student understanding. This didactic mode can be enhanced with self-test question as well as quiz questions allowing the student and teacher to identify the level of knowledge and understanding achieved. Most computer-based learning systems, particularly multimedia presentations do not enter a full Socratic mode as a student model is not developed during interaction and the responses of the system are not adapted to the abilities of the student.

ITSs are so called for their 'intelligent' ability to adapt to the needs of individual students, by being able to adapt to the knowledge, learning abilities and needs of each individual student. Traditionally, ITS research has assumed that students are modelled according to their answers to questions. These models represent different aspects of the student's cognitive state.

An ITS applies artificial intelligence techniques to computer-assisted instruction allowing a fully Socratic mode to develop with individualised instruction. Figure 1 shows the main components of an intelligent tutoring system. An ITS is traditionally said to comprise of four interdependent components: the student model, the pedagogical module, domain knowledge and the communication model. The student model stores information specific to individual learners, upon which the pedagogical module devises appropriate teaching strategies for the ITS to employ. These strategies are applied to the domain knowledge, generating a subset of knowledge to be presented to the learner using the communication model, which acts as an interface between the learner and the ITS. As the learner responds to the system, the student model is updated and the cycle repeats.

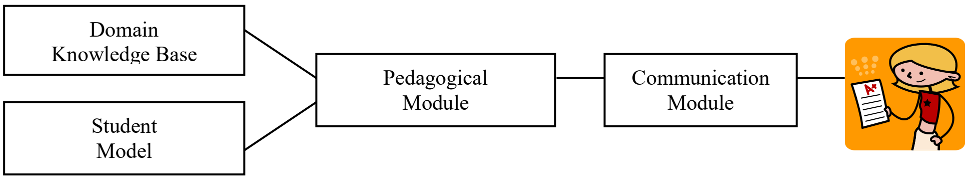

Figure 1. Major Components of an Intelligent Tutoring System

ITSs individualise instruction by maintaining a model of the student. A student model containing knowledge about the learner, is thus the most important component of an ITS. The aim of student modelling in existing ITS's is to construct a model of the learner depicting the learner's state of knowledge. Existing student models only maintain information about student's knowledge and use that to individualise instruction. Student models may contain information about what the student knows, has incorrect knowledge of, or knows to an extent giving rise to overlay, perturbation and fuzzy student modelling.

However, many researchers now agree that to restrict student modelling to simply interpreting answers is to overlook one of the human tutor's greatest allies, an appreciation of the student's non-verbal behaviour (e.g. Picard, 1997; Kort, Reilly, and Picard, 2001). Such is the nature of human communication, that tutors unconsciously process a continuous stream of rich non-verbal information that can suggest improved tutoring strategies. Competent human tutors adapt their tutoring according to the real time non-verbal behaviour of their students, as well as their answers to questions.

Since adapting to the non-verbal behaviour of students is key to the success of human tutoring, it follows that perhaps ITSs could be significantly improved if they too could recognise and adapt to the affective information carried largely by the non-verbal behaviour of students. This conclusion has spawned a developing field of artificial intelligence: Affective Tutoring Systems. Affective Tutoring Systems adapt to the affective state of students just as effective human tutors do (de Vicente, 2003; Alexander, Sarrafzadeh, and Fan, 2003; Sarrafzadeh, Fan, Dadgostar, Alexander and Messom, 2004).

3. Why Affective Tutoring Systems

3.1 Overview

The extent to which emotional upsets can interface with mental life is no news to teachers. Students who are anxious, angry, or depressed don't learn; people who are caught in these states do not take in information efficiency or deal with it well (Golman, 1996). Skilled humans can assess emotional signals with varying degree of accuracy. Within the Human-Computer Interaction community, there is a growing agreement that traditional methods and approaches for user interface design need to become more human-like (Falangan, Huang, Jones and Kasif, 1997). One aspect of developing such capability is to recognise the user's emotional or mental state and respond appropriately (Reeves and Nass, 1996). Adding such capability to machines can reduce the gap between human thinking and machine *'thinking'*, although, this gap is still very large.

There are two crucial issues on the path to what Picard (1997) has coined "affective computing", which are: 1) Providing a system with a mechanism to infer the emotional state and personality of the user, 2) Providing a mechanism generating behaviour in an application, consistent with a desired personality and emotional state.

Many researchers now feel strongly that Intelligent Tutoring Systems (ITSs) would be significantly enhanced if computers could adapt according to the emotions of students (Picard, 1997; Kort, Reilly and Picard, 2001; Alexander and Sarrafzadeh, 2004). ATSs have a very short history: it seems that the term "Affective Tutoring System" was first used only a few years ago (Alexander, Sarrafzadeh and Fan, 2003; de Vicente, 2003), although the popular concept of an ITS adapting to perceived emotion can be traced back at least as far as Rosalind Picard's book *Affective Computing* (1997). However, so far as the author is currently aware, no ATS has yet been implemented, although several groups are working towards this goal (Kort, Reilly and Picard, 2001; Alexander, 2004; Litman and Forbes, 2006).

3.2. Why facial expressions are important

Facial expressions are most visible reflexes of our emotions. It is possible that by measuring muscular actions by the visible changes they produce in bulges, bags, pouches, wrinkles, shapes and positions of facial features (Hager and Ekman, 1995), and we can judge about their related emotions. Some emotions including sadness, happiness, anger, disgust, surprise and fear, are relatively automatic, and involuntary (Ekman, 1994).

3.3. Why gestures are important

There is a broad agreement on the scale of the emotion lexicon. Approximately 93% of human affective communications is conveyed through nonverbal means (Mehrabian, 1981). Surprisingly, there is little knowledge about the extent to which body movements and gestures provide reliable cues to emotions across the life span, although such information may have the advantage of primacy during social interactions because it is the first thing observed when people approach each other. Recent results with psychophysics indicate that in natural circumstances, the eye, the head, and hands are in continual motion in the context of ongoing behaviour (Hayhoe, 2000). In particular, body movements (e.g. gait, body posture) and gesture (e.g. hand positions) have been found to reliably communicate emotional messages (Montepare, Koff, Ziatchik and Albert, 1999).

3.4. Comparing emotion recognition by man and machine

Measuring emotions using physiological signal sources, like automatic arousal, heart rate, blood pressure, skin resistance and some facial electromyography activities, is the attractive prospect that, physiological measurement might offer a way to accessing a person's emotional state. On the other hand, the real discriminative power of physiological measures is limited. However, automatic detection of human emotions is an area that is still maturing and researchers in various disciplines are still making progress. To understand the current state of technology, it is appropriate to compare the performance of human and machine.

In order to get an idea about the effectiveness of machine-based emotion recognition compared to humans, a review of research done by Huang, Chen and Tao (1998) follows. They investigated the performance of machine based emotion employing both video and audio information. Their work was based on human performance results reported by De Silva, Miyasato and Nakatsu (1997). It was indicated that including both the video information (as extracted facial expressions) and the audio information (as prosodic features) improves the performance significantly. Huang, Chen and Tao's research indicated that the machine performance was on average better than human performance with 75% accuracy. In addition, comparing detection of confusions indicated similarities between machine and human. These results are encouraging in the context of our research for intelligent tutoring systems.

4. An architecture for Affective Tutoring Systems

4.1. Overview

There will be four main components in an Affective Tutoring System: a student model, a set of tutoring strategies, domain knowledge, and a tutoring module that will interface with the student.

The student model of an ATS will be divided into two main parts: one will analyse the student's answers to questions, and the other will analyse the student's non-verbal behaviour. Non-verbal behaviour will be identified by analysing images of the student's upper body and face to detect gestures or facial movements. The tutoring module will select the most appropriate tutoring strategies on the basis of the state of the student model; material from the domain knowledge component will then be presented on the basis of these tutoring strategies. The architecture of a complete

Affective Tutoring System is shown in Figure 2, which shows the scope of the current research (the bold box) and how it will be extended in future to include other channels for detecting nonverbal behaviour of the learner.

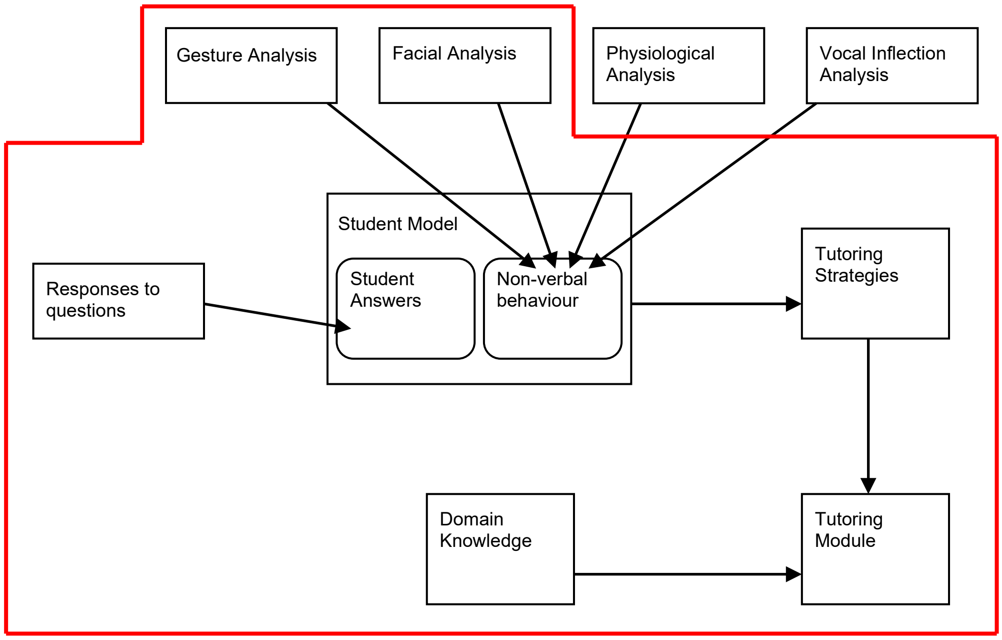

Figure 2. Architecture of an Affective Tutoring System.

4.2. Detecting affective state via gesture and facial expression

4.2.1. Facial expression analysis for ITS

A major branch of image processing is automated facial expression analysis. Using a simple web-cam, automated facial expression analysis systems identify the motion of muscles in the face by comparing several images of a given subject, or by using neural networks to learn the appearance of particular muscular contractions (Fasel and Luettin, 2003; Pantic and Rothkrantz, 2003). This builds on the classic work of Ekman and Friesen (1978), who developed the Facial Action Coding System for describing the movement of muscles in the face. An affective state can be inferred from analysing the facial actions that are detected (Pantic and Rothkrantz, 1999).

A state-of-the-art automated facial expression analysis system was developed in-house at Massey University in Auckland, New Zealand in 2003. Figure 3 shows a screenshot of the system's output (Fan, Johnson, Messom, and Sarrafzadeh, 2003). There are three main components to the system: face detection using an Artificial Neural Network, facial feature extraction and a fuzzy facial expression classifier. The system is capable of accurately identifying seven affective states: surprise, happiness, sadness, puzzlement, disgust, and anger, plus a neutral state. A new improved system has recently been developed which is more suitable for use in ATS. Test results show a high degree of accuracy and real-time performance as presented in the following section.

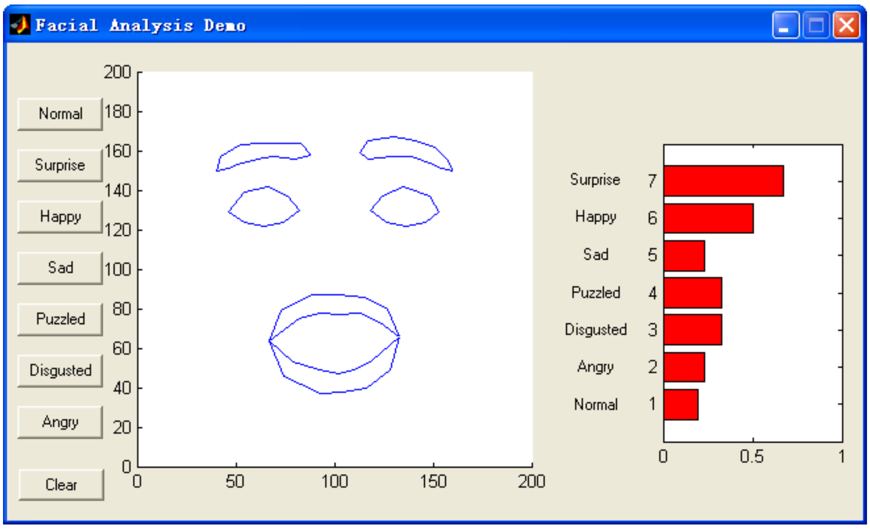

Figure 3. A screenshot taken from the in-house facial expression analysis software

4.2.1.1. Test results for the improved facial expression recognition component

We recently improved our facial expression recognition system using a new approach to facial expression recognition and obtained very promising results. The new approach is based on support vector machines (SVM). A support vector machine is a supervised learning algorithm based on statistical regression (Vapnik, 2000). The SVM algorithm operates by mapping the training set into a high-dimensional feature space, and separating positive and negative samples (Cristianini and Shawe-Taylor, 2000).

We represented the facial expression database that we used by connection features. Each raw image is 200 pixels in width and 200 pixels in height. By connection extraction, we reduced the image size to 50 pixels by 50 pixels, which is 2500 connection features. This makes the training process significantly faster than the pixel-wise analysis of the image which we used in the earlier version.

An SVM was trained using the summarised facial images using our connection features algorithm. The training was obviously considerably shorter using this approach. Memory requirements were also considerably lower. To train the SVM, a

Table 1. The result of applying different kernel models and correct detection for each facial expression

	Linear Kernel Model	Polynomial Kernel Model	RBF Kernel Model
Normal	89%	85%	92%
Disgust	78%	82%	93%
Fear	83%	86%	90%
Smile	91%	92%	93%
Laugh	85%	92%	96%
Surprised	87%	93%	94%

5-fold cross-validation was applied. Different Kernel methods are used in calculating the SVM to classify non-linear functions. We tested different kernel models, namely a linear model, a polynomial model and an Radial Basis Function (RBF) Kernel with the SVM, and have presented the results in Table 1. The test image database contains 1000 image for each facial expression making up 6000 images in total. The results obtained using an RBF kernel were the most promising.

Video sequences are recorded and analyzed at a rate of 12 frames per second which is real-time. Our SVM analyzer is triggered each 0.0833 seconds. These results indicate that the system is capable of operating in real-time and with a high rate of accuracy.

4.2.2. Gesture analysis for ITS

An affective tutoring system should be able to analyse the body movements of the learner just as a human tutor would. A child counting fingers when given an addition or subtraction is in a totally different developmental stage than one who doesn't. Detecting counting on fingers and various other head and hand movements is a slightly different case, as these are gestures, and not facial expressions. However, similar techniques employed by automated facial expression analysis can also be applied to automated gesture analysis. We have developed a system that detects various hand and body movements which is being extended to include a student counting on his/her fingers.

McNeill (1992) and Cassell (1998) categorise gestures in communication activities into three types: *deictic, iconic, and metaphoric*. These three types of gestures have different roles in communication. Deictic gestures, also called pointing gestures, highlight objects, events, and locations in the environment. They have no particular meaning on their own end, but frequently convey information solely by connecting a communicator to a context. Deictic gestures generally spatialise or locate the physical space in front of the communicator, with aspects of the discourse. An example of this type of gesture might be pointing left and then right, saying "well, Jane (left) was looking at Peter (right) across the table. . ." (Cassell, 1998).

Iconic gestures, on the other hand, can convey much clearer meaning out of context than deictic gestures. These gestures represent information about such things as object attributes, actions, and spatial relations. Iconic gestures may specify the manner in which an action is carried out, even if this information is not given in accompanying speech. As Cassell (1998) exemplified, only in gesture does the speaker specify the essential information of how the handle of the caulk gun is to be manipulated. Specifically, recent work on children communication behaviour shows that deictic and iconic gestures are pervasive in children's speech. Interestingly, children produce deictic gestures before they begin to talk (Goldin-Meadow S, Singer, 2003).

Finally, metaphoric gestures are more representational, but the concept they represent has no physical form. Instead, the form of the gesture comes from a common metaphor. An example is "the meeting went on and on" accompanied by a hand indicating rolling motion (Cassell, 1998). There need not be a productive metaphor in the speech accompanying metaphoric gestures; sometimes the "metaphors" that are represented in gesture have become entirely conventionalised in the language, e.g., describing the solution of a mathematical equation or a physics problem by students. Kwon, Ju, Park, Cho and Ewha (2003) showed that the student's gesture is often transformed from a pictorial metaphoric/iconic gesture to a deictic gesture of simple pointing.

The use of the three types of gestures would vary with different learning contexts. Recent research within psychology and mathematics education has looked at the role of gesture and embodiment in the different problem domains such as counting (Alibali and diRusso, 1999) or arithmetic problem solving (Goldin-Meadow, 2003; Edwards, 2005), suggesting that different learning situations might allow different use of nonverbal cues. Together with this understanding, we have raised one research hypothesis. The gesture use may vary with the learner's level of skill. However, none of the research has identified how the learner's level of skill may affects on the different use of gestures.

We have developed a novel approach to 2D gesture trajectory recognition for use in our affective tutor. This approach taken consists of two main steps: i) gesture modelling, and ii) gesture detection. The gesture modelling technique which we have developed is robust against slight rotation, requires a small number of samples, is invariant to the start position of the gesture and it is device independent. For gesture recognition, we used one classifier for detecting each gesture signal. We implemented the gesture recognition system using both multilayered feed-forward Artificial Neural Networks (ANNs) and support vector machines. We evaluated the neural network approach in comparison with a SVM with radial basis kernel function. The results showed high accuracy of 98.27% for ANN and 96.34% for SVM in gesture signal recognition. The results also show that the overall performance of the ANN classifier is slightly better than the SVM classifier. We have therefore adopted the ANN approach. The gesture recognition system we have developed is based on movement trajectory information. Therefore, it can be used with a variety of front-end input systems and techniques including vision-based hand and eye tracking, digital tablet, mouse, and digital gloves.

Table 2 presents the results of testing the gesture recognition system which is currently capable of recognizing 13 gesture signals. The accuracy of the detection is satisfactory as shown in the table. Figure 4 shows a screen shot of the gesture recogniser tracking a hand.

Table 2. The evaluation of gesture recognition system

Gesture No.	Correctly detected	Total correct	Correct positive detection	Falsely detected	Total false	False positive detection	False positive to Correct positive detection
1.	80	80	100%	23	402	5.72%	0.05714
2.	110	112	98.21%	0	397	0%	0
3.	101	103	98.05%	35	397	8.81%	0.089907
4.	96	103	93.2%	29	401	7.23%	0.077592
5.	175	176	99.43%	15	602	2.49%	0.025059
6.	176	182	96.7%	23	619	3.72%	0.038423
7.	93	93	100%	15	440	3.41%	0.034091
8.	82	84	97.62%	27	407	6.63%	0.067957
9.	96	96	100%	0	409	0%	0
10.	89	90	98.89%	13	417	3.11%	0.031525
11.	103	105	98.09%	6	387	1.55%	0.015805
12.	105	105	100%	18	411	4.38%	0.043796
13.	116	118	98.30%	12	436	2.75%	0.027997
	1422	1447	98.27%	216	5725	3.77%	0.0383

Figure 4. Gesture recogniser tracking the hand

4.2.3. Reacting to affect

Though few if any existing ITSs can recognise emotions, many ITSs have been developed that can show emotions through an animated pedagogical agent (Johnson, Rickel and Lester, 2000; Prendinger and Ishizuka, 2004). Animated pedagogical agents are “lifelike autonomous characters that co-habit learning environments with students to create rich, face-to-face learning interactions” (Johnson, Rickel and Lester, 2000). Animated agents carry a persona effect, which is that the presence of a lifelike character can strongly influence students to perceive their learning experiences positively (van Mulken, André and Muller, 1998). The persona effect has been shown to increase learner motivation, especially in technical domains, although its overall benefits remain unclear (van Mulken, André and Muller, 1998). Assuming that the affective state of students can be reliably identified, animated agents are able to show timely empathy towards students through their own facial expressions and gestures.

4.3. Real-time Performance

As stated in section 2, the ATS that is developed currently utilises two main inputs namely Facial Expression and Body Gesture. As more and more bio-signals of both physiological as well as vocal inflection are used to enhance the student model in the system, the computational requirements for such analysis increases. This increase in computational process can quickly hinder the practicality of such system. Furthermore, a general purpose computer will fall short of performing in real-time once more subtle and distinctive features such as skin resistance, vocal vibration, eye movement etc is to be processed.

In order to address the performance requirement of a practical ATS and to allow seamless integration of additional modules as they are developed into the current system, a modular and distributed network of computing devices is proposed. In the proposed architecture, the main computer hosting the domain knowledge and the brain of the ATS also acts as the central controller for a range of embedded computing devices each processing different bio-signal and only communicating the processed data to the main ATS engine.

Figure 5 depicts the distributed computing systems developed for a modular Affective Tutoring Systems. Embedded Processors in the figure depending on the signal being processed can compromise one or more of the following: Microcontroller, ASIC, Microprocessor, Digital Signal processor and Field Programmable Gate Arrays (FPGA).

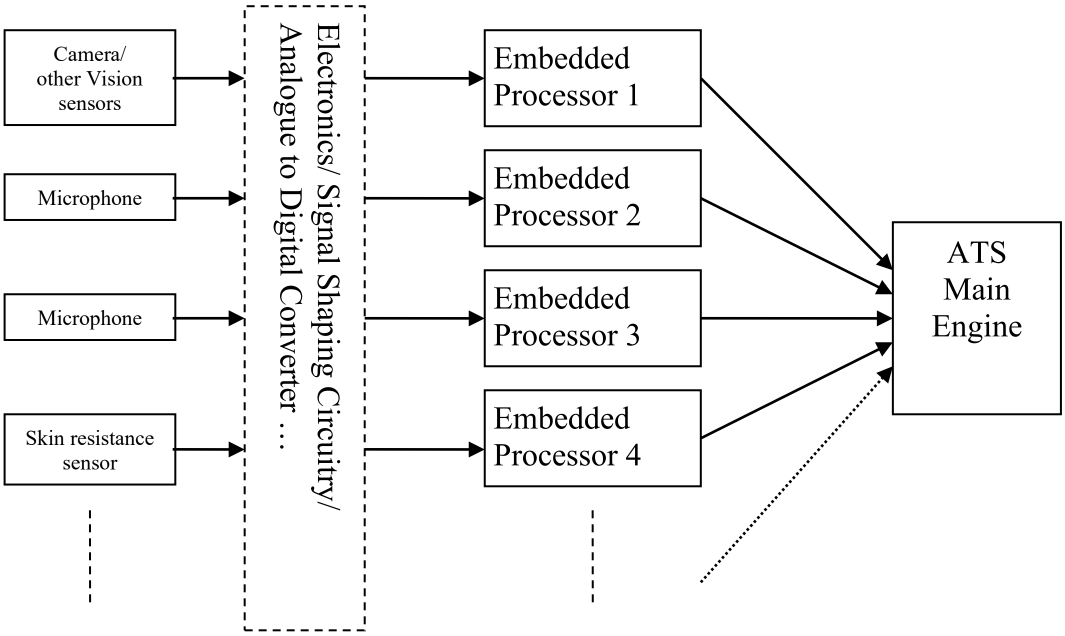

Figure 5. A distributed Computer System for implementation of a real-time ATS

4.4. Pedagogical Uses for Affective Knowledge

Efforts toward affective state recognition in our research are focused on enhancing learning, within intelligent tutoring systems. Specifically, we are developing new models of computational pedagogy that augment the traditional approach of monitoring a student's current subject knowledge with active utilization of knowledge of a student's affective state. There are a number of specific pedagogical support tasks that we are exploring, including:

- Detection of boredom, confusion, inattention, and anxiety
- Detection of a student's affinity for and anthropomorphisation of the pedagogic persona
- Transformation of counterproductive affective states using content modification and affective agents

4.4.1. Detection of boredom, confusion, frustration, inattention and anxiety

We believe that fundamental to the successful use of learners' affective state in intelligent tutoring systems, is accurately estimating the state (or combination thereof), and using the knowledge to adapt content and presentation. The affective states of bored, confused, and frustrated are detectable from facial expressions. This is supported by our own work, as well as that of others (D'Mello, Craig, Gholson, Franklin, Picard and Graesser, 2005).

For example, upon detecting a persistent bored affective state, we might change the material being presented. Alternatively, we might present the user with a knowledge assessment to determine if the material is already known. Confusion, on the other hand, might be mitigated by the presentation of the next previous topic, or the one prior to that, if necessary, depending on the user's knowledge state. Affective state alone is unlikely to be conclusive with respect to the pedagogical goals of the lesson without confirmation of knowledge state.

4.4.2. Detection of affinity and anthropomorphisation

This category of support could more rightly be termed "feedback on the success of affinity seeking strategies", since what we seek to attain here is a knowledge of whether or not efforts to generate affinity for the pedagogic persona embodied by the ITS has been successful. In a survey of 293 undergraduate students, Beebe and Butland (1994) demonstrated that teachers who utilise affinity seeking strategies significantly increase student cognitive learning, student liking, and feelings of pleasure and arousal (and consequently more learning). This study was done using student questionnaires (i.e., self-report), however, our system would be able to more directly measure learning and affect to confirm the authors' findings.

The idea here is to attempt to employ affinity seeking strategies within the intelligent tutoring system, and then determine the effects on student learning by sensing affective response and by cross-referencing student knowledge state. This is assuming that the notion of immediacy, originally advanced by Mehrabian (1972), extends to intelligent tutoring systems. There is ample evidence to support this notion, however, given the degree to which users anthropomorphise their computers, and treat them like other humans (Reeves and Nass, 1996).

4.4.3. Transformation of counterproductive affective states

This category of support can be thought of as maintaining affective states that support learning. One recent approach to this area is detecting affective state, and given personality type, attempting to generate the "optimal" emotional state for learning on an individual basis (Chaffar, Frasson, 2004). This is an interesting concept; however, we think that this approach is premature. In the study cited (Chaffar, Frasson, 2004), emotional state was obtained by subject's self-report as was the connection between personality type and "optimal" emotional state for learning. In addition, the authors' measurement of affective state was based on colour selection on the part of the users, and was only shown to be about 58% accurate, or close to chance performance. In our research, we directly measure affective state via facial expression and gesture recognition, and will use these measurements to determine both "momentary" and "sustained" affective state. This knowledge will be used to decide when it is necessary to help the learner modify their own affective state to promote learning.

New research directions in this area are reported by Burleson and Picard, who are using affective learning agents to help learners maintain a sense of "flow", and to help them over periods in which they are "stuck" on a particular problem (Burleson and Picard, 2004). In this work, they propose to use what they call "affective partners" as a collaborative learning agent, using the learner's affective state to direct the behaviour of the learning agent. We believe that this is a fruitful direction for affective tutoring systems research, and intend to explore this area as part of our research as well.

5. The development of an Affective Tutoring System for mathematics

5.1. Easy with Eve

Using the facial expression and gesture recognition systems explained above combined with lifelike agents and a case-based system, we have developed Easy with Eve, an affective tutoring system that is capable of detecting and expressing affect while teaching primary school mathematics. The following subsections explain the tutor and the background studies leading to its development.

5.2. Identifying Affective State

As discussed above, there are several different ways that computers can attempt to identify the affective state of users. These can be divided into two main groups: methods that aim to detect emotions based upon their physical effects, and methods that aim to predict emotions based upon understanding their causes. Methods that aim to detect emotions based upon their physical effects include facial expression analysis (e.g. Sarrafzadeh, Fan, Dadgostar, Alexander and Messom, 2004), gesture analysis (e.g. Dadgostar, Ryu, Sarrafzadeh and Overmyer, 2005), voice analysis (e.g. Litman and Forbes-Riley, 2006) and wearable computing (e.g. Strauss, Reynolds, Hughes, Park, McDarby and Picard (2005); one example of a predictive emotion model is given by Conati and Maclaren (2005).

5.3. An Observational Study of Human Tutors

However, even if an ATS could perfectly identify the affective state of students, it would still need to know what to do with this information before it could adapt its tutoring in a genuinely useful manner – this is the key issue that this study addresses. As good human tutors *can* effectively adapt to the emotions of students, the most obvious way to learn about how to adapt to the affective state of students is to study human tutors.

The ways in which human tutors adapt to the affective state of students have yet to be fully explained. Therefore the aim of the observational study was to take a step towards designing an ATS that can sensibly make use of affective state information.

Secondly, if the affect-based adaptations of the animated agent are based on human tutors then this should help to increase the agent's believability. This is advantageous because if the animated agent is especially lifelike, then this will maximise the persona effect (van Mulken, André and Muller, 1998).

5.4. Methodology

The observational study of human tutors involved videoing several tutors as they tutored students individually. There were three tutors altogether, and nine student participants, all of whom were 8 or 9 year old students at a school in Auckland, New Zealand; each participant was tutored for about 20 minutes.

The domain that was chosen for the observational study was the concept of part-whole addition. The study used an existing exercise developed by the New Zealand Numeracy Project (2003) that teaches students to add numbers by transforming the initial equation to make the first addend up to the next 10 – hence the phrase "part-whole addition". For example, 17 + 6 would become 17 + 3 (to make 20) + 3 = 23. Students learn this reasoning by manipulating tens frames and counters, as shown in Figure 6: in this example the student should move three counters from the tens frame on the right over to the tens frame in the middle to simplify the equation.

As students progress through the exercise the tasks become increasingly abstract, until by the end of the exercise students need to apply the principle of part-whole addition to equations where using physical tens frames and counters would not be practical. For instance, students would have to solve an equation like 87 + 6 in their head, because they would not have that many counters to use.

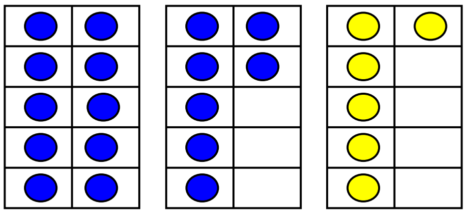

Figure 6. Tens frames and counters in the maths exercise.

5.5. Video Analysis

To analyse the videos, a coding scheme was developed expanding on previous work by Person and Graesser (2003). This scheme was used to extract data from each tutoring video to describe the behaviours, facial expressions and expression intensities of students and tutors.

Each tutoring video was divided into several hundred clips, with each clip being either a student or tutor turn in the tutoring dialogue – student and tutor turns describe the behaviour of the actor in any given clip. For each clip the actor (either "student" or "tutor"), the turn, the facial expression of the actor, and the intensity of the expression (either "low" or "high") was recorded, thus generating the raw data of the study.

5.5.1. Results

The nine tutoring videos were divided into over 3000 sequential clips of student and tutor turns. The coded data from the clips showed several main results:

- Almost all student turns were related to answering questions.
- The occurrences of tutor turns were more varied, although actions "ask new question", "pump for additional information", "positive immediate feedback" and "neutral immediate feedback" between them totalled over two thirds of all tutor turns.

- As Figure 7 shows below, neutral expressions were by far the most common, especially for tutors, but also for students.
- The second and third most commonly appearing expressions were also the same for both students and tutors, with smiling (low intensity) the second and smiling (high intensity) the third most common expressions. Smiles were much more common for students than tutors. See Figure 7.
- All other expressions, including confusion and apprehension, were very rare by comparison. See Figure 7.

These results, along with their implications and shortcomings, are discussed in much greater detail in Alexander, Hill and Sarrafzadeh (2005).

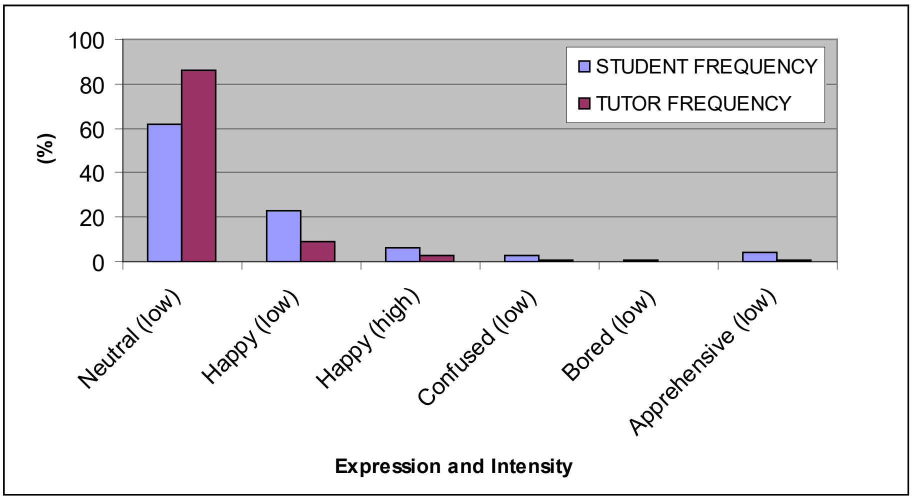

Figure 7. Frequencies of student and tutor facial expressions (at either low or high intensity) in the observational study of human tutors.

5.6. A case-based method for adapting to student affect

5.6.1. Background

The data from the observational study of human tutors contain a wealth of information about the *interaction* between tutors and students during the tutoring process. For any given combination of student turn, facial expression and intensity of facial expression, the following information is readily available:

- the frequencies in the data of all the tutor turns that immediately follow this combination of student states, and
- the frequencies in the data of all the tutor facial expressions (and intensities) that immediately follow this combination of student states.

However, a human tutor's response to a tutoring scenario is influenced by the *history* of his/her interactions with the student throughout the tutoring session. Implicit in the data is the way that a human tutor's adaptations vary according to the immediate history of interactions with a student.

Thus, a case-based reasoning program has been written that searches the data based upon the sequence of interactions in a given scenario, and outputs a weighted set of recommended tutoring actions and facial expressions.

5.6.2. Searching for similar sequences

The earliest version of the case-based program only searched the data for exact matches with the given scenario; if no matches were found for a given sequence, then the sequence would be iteratively shortened by one interaction until a match was found in the data (Alexander, Hill and Sarrafzadeh, 2005). However, this approach had two main shortcomings: firstly, only a relatively small amount of data would ever be relevant to a specific scenario; and secondly, a lot of very relevant data would be completely overlooked in most cases. For instance, "ask new question" with a low-intensity smile is very nearly the same as "ask new question" with a high-intensity smile. Similarly, "give neutral feedback and ask new question" is very nearly the same as "ask new question" – and thus a sequence of interactions in the data containing the former might be extremely relevant to a scenario containing the latter. By including similar sequences in the search, it should be possible for the program to make a more balanced recommendation of appropriate tutor responses. A fuzzy approach such as this would also make the data go much further, as much more of the data would be relevant to any given search than would otherwise be the case.

5.6.3. Implementation of the Fuzzy Approach

The case-based program takes as input a (hypothetical) sequence of interactions between a tutor and a student, which is coded using the same scheme that was used in the observational study of human tutors. It generates a weighted set of similar sequences that are relevant to this input scenario, and searches the data from the observational study for each of these sequences. Whenever a match in the data is found, the tutor's next action is recorded, again in the format of the coding scheme. Each of these tutor's next actions has a cumulative score; each time a match in the data is found, the score for the tutor's next action is increased by the weight of the sequence that matched the data.

Similar sequences are generated in three different ways: by varying the turns in interactions, by varying the expressions in interactions, and by varying the lengths of the interactions. Each student and tutor turn is linked to a set of other turns with specific weights (between 0, low and 1, high), and each combination of expression and intensity is linked to a set of other expressions and intensities with specific weights. So the first step is to generate new sequences using the turns and expressions that are linked to each of the interactions in the current scenario – the overall weight of each sequence is the product of each of the weights of the interactions. A new set of sequences is then generated by varying the lengths of all the sequences that have now been generated; the overall weight of these shortened sequences falls exponentially as their lengths decrease. All of these sequences are then searched for in the data from the observational study.

To keep a lid on the number of sequences that are generated, the maximum length of a sequence is currently restricted to the last 15 interactions between the tutor and the student (the minimum sequence length is 1). Generated sequences with a weight so low as to render them insignificant are also discarded.

The recommendations of the case-based program are the tutor's next actions that are found to follow matches with any of the sequences. Each of these recommendations carries a score – this is a function of the weightings and frequencies of the sequences that preceded the recommended tutor actions. A sample of the output of the system is given in Figure 8: given the sequence of interactions "student answers incorrectly, neutral expression", "tutor gives a hint, neutral expression", "student answers correctly, low intensity happy expression", the recommended tutor action with the highest score is "pump for additional information, neutral expression". In this particular case there were over 100 different recommendations generated (not all shown in the screenshot), ranging in score from 21.76 to 0.06; well over half of these recommendations had a score of less than 1.

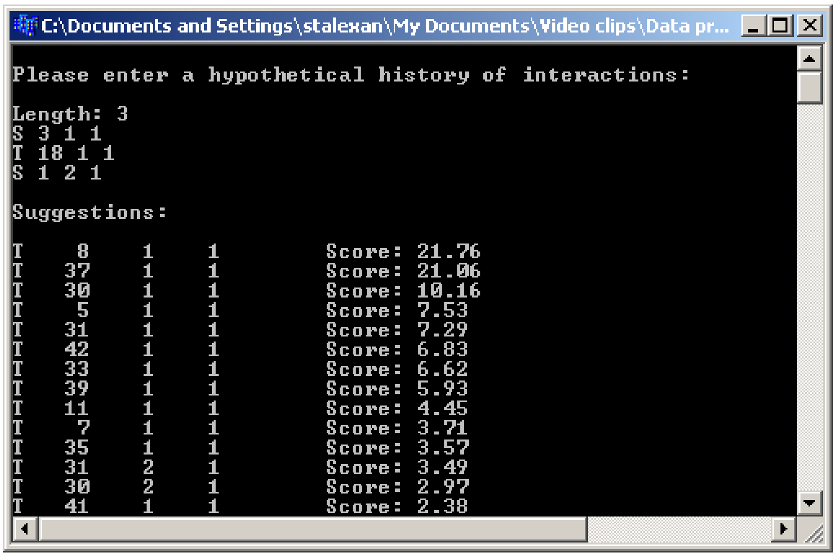

Figure 8. Screenshot of the case-based program for recommending tutoring adaptations.

5.7. An ATS for addition

The case-based method for adapting to student affect has been applied in an ATS, *Easy with Eve*, which is designed to help primary school students with the same New Zealand Numeracy Project exercise that was used in the observational study of human tutors.

5.7.1. Eve: An affect-sensitive animated tutor

Eve displays a comprehensive range of emotions through facial expressions; she is also able to deliver teaching content through realistic lip-synching and gestures. All possible tutoring actions of Eve were first animated, and saved as videos that could be imported into the ATS. These tutoring actions include: giving positive or neutral feedback, asking questions, discussing problems or solutions, giving hints, or answering her own questions if need be. Whenever Eve is waiting for a response from a student, a looping 'idle' video seamlessly gives the impression that she is patiently waiting. Figure 9 gives a small sample of the appearance and capabilities of Eve.

Figure 9. Three examples of Eve in action (from left to right): showing no expression, smiling, and speaking.

5.7.2. Emotion detection

Emotion detection in the ATS is achieved using a real time facial expression analysis system that has been developed in-house at Massey University (Sarrafzadeh, Fan, Dadgostar, Alexander and Messom, 2004), as discussed above. The emotion classification is achieved by using support vector machines, and is able to detect the 6 basic facial expressions that are defined by Ekman (1997). The module uses a facial feature extraction algorithm that is not only able to extract all important facial features, but is also fast enough to work in real time. Unlike other facial expression analysis systems, facial information is automatically detected without the need to manually mark facial features.

5.7.3. Integrating eve with emotion detection and the case-based program

Eve's responses to the student are driven by the case-based program discussed above. This program requires the ATS to keep a running history of the interactions between Eve and the student: student turns are determined by their responses to Eve's questions; student expressions and intensities are determined by the facial expression analysis system; and tutor turns, expressions and intensities are kept track of as Eve does them. Thus each time the student does something, the history is updated, and each time Eve does something, the history is updated too.

The case-based program generates a set of recommended tutor actions, each with weighted scores; Eve randomly selects one of the recommendations to follow according to the weights of the recommendations. Thus the recommendations with the higher scores are more likely to be selected.

The facial expression analysis system runs independently of the rest of the ATS. The expression and intensity of the student is continuously updated in a data file – this file is accessed by the ATS whenever it needs to update the history, i.e. the student has just responded to a question that Eve has asked, or the system has timed out.

5.7.4. An order of events

So whenever the student responds to a question, or the system times out (the student has taken too long to answer the question), the following is the order of events:

- the data file generated by the facial expression analysis system is accessed to find out about the current facial expression of the student;
- the history is updated classifying the student's response to the question and using the expression information from the data file;
- the case-based method generates a set of weighted recommendations for Eve's next action;
- based on their weights, a response is chosen from the set of recommendations;
- the tutoring action is carried out by Eve;
- the history is updated with what Eve has just done;
- Eve waits for the next student response.

6. Conclusion

A functioning ATS, *Easy with Eve*, has been implemented at Massey University, and is possibly the first of its type. Emotion detection is carried out using facial expression analysis, and the system adapts to the student via an emotionally expressive animated lifelike agent, Eve. Tutoring actions are guided by a case-based method for adapting to student states that recommends a weighted set of tutor actions and expressions. The data that this case-based program uses was generated by an observational study of human tutors.

The implications of this system are potentially very significant: it represents another step towards tutoring systems that are fully aware of the cognitive and affective states of the student, and are fully capable of adapting to these states wisely. Future work will build on the current ATS, and work towards making Eve increasingly lifelike in her interactions with students.

7. Future Work

The next step for this research will be to evaluate the effectiveness of the ATS in a range of learning situations including both young and adult learners; the system will be tested with primary school children at the same place where the observational study of human tutors was carried out. The data gathered from these tests should provide valuable information on the effectiveness of including affective state in a tutoring strategies module, and also on the significance of empathy on the persona effect of an animated pedagogical agent. The testing should also provide important direction for improvements to be made in the next version of *Easy with Eve*.

The nonverbal communication component of the system including facial expression which will soon be augmented with a gesture recognition components will be improved and extended to include other input channels such as vocal and physiological input.

The other improvement to be made to the ATS is to iron out memory problems relating to the case-based method for recommending tutor actions; the large number of sequences that are searched for in the data has caused the system to run slowly as the history of interactions increases in length. Another avenue for future work is the addition of a learning component to the case-based program. It would be very useful if Eve could learn from her interactions with students, rather than relying on the existing data that was collected from the observational study of human tutors.

References

Alexander, S. T. V., Sarrafzadeh, A. & Fan, C. (2003). Pay attention! The computer is watching: Affective Tutoring Systems. Proceedings of E-Learn 2003, Phoenix, Arizona.

Alexander, S. T. V., & Sarrafzadeh, A. (2004). Interfaces that adapt like humans. Proceedings of Asia-Pacific Computer-Human Interaction 2004, Rotorua, New Zealand.

Alexander, S. T. V., Hill, S., & Sarrafzadeh, A. (2005). How do human tutors adapt to affective state?. Proceedings of User Modeling, Edinburgh, Scotland.

Aleven, V., Koedinger, K. R., & Cross, K. (1999). Tutoring answer explanation fosters learning with understanding. In: Lajoie, S.P. and Vivet, M.(eds.), Proc. AIED'99, IOS Press, 199-206.

Aleven, V., & Koedinger, K. (2000). Limitations of student control: Do students know when they need help? Proc. ITS'2000, Springer-Verlag, 292-303.

Alibali M. & diRusso A. (1999). The function of gesture in learning to count: More than keeping track. Cognitive Development, 14, 37-56.

Anderson, J.R, Corbett, A.T., Koedinger, K.R., & Pelletier, R. (1995). Cognitive tutors: Lessons learned. The Journal of the Learning Sciences, 4(2), 167-207.

Beebe, S.A. & Butland, M. (1994). Emotional response and learning: Explaining affinity seeking behaviours in the classroom. In Proceedings of the Annual Meeting of the International Communication Association Conference, Sydney, Australia.

Burleson, W., & Picard, R. W. (2004). Affective agents: Sustaining Motivation to learn through failure and a state of stuck. Social and Emotional Intelligence in Learning Environments Workshop In conjunction with the 7th International Conference on Intelligent Tutoring Systems, Maceio - Alagoas, Brasil, August 31st, 2004.

Cassell J. (1998). A framework for gesture generation and interpretation. In: Computer Vision for Human-Machine Interaction. R. Cipolla & A. Pentland eds., Cambridge University Press.

Chaffar S., Frasson, C. (2004). Inducing optimal emotional state for learning in Intelligent Tutoring Systems, In: Proceedings of the 7th International Conference on Intelligent Tutoring Systems, ITS 2004, Maceió, Alagoas, Brazil, August 30 - September 3, 2004, Lecture Notes in Computer Science, Volume 3220, 45 – 54.

Conati, C., & VanLehn, K. (2001). Providing adaptive support to the understanding of instructional material, Proc. Intelligent User Interfaces '01, Santa Fe, New Mexico, 2001, 41-47.

Conati, C. (2002). Probabilistic assessment of user's emotions in educational games, *Applied Artificial Intelligence*, 16, 7-8, 555-575.

Conati, C. & Maclaren, H. (2005). Data-driven Refinement of a Probabilistic Model of User Affect. Proceedings of User Modeling, Edinburgh, Scotland.

Cristianini, N., Shawe-Taylor, J. (2000). *An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press, Cambridge, UK.

Dadgostar, F., Ryu, H., Sarrafzadeh, A. & Overmyer, S. P. (2005). Making sense of student use of nonverbal cues for intelligent tutoring systems. Proceedings of CHISIG, Canberra, Australia.

De Silva L. C., Miyasato T., & Nakatsu R. (1997). Facial emotion recognition using multimodal information, In Proc. IEEE International Conference on Information, Communications and Signal Processing, USA, 397–401.

D'Mello, S. K., Craig, S. D., Gholson, B., Franklin, S., Picard, R. & Graesser, A. C. (2005). Integrating affect sensors in an intelligent tutoring system. In Affective Interactions: The Computer in the Affective Loop Workshop at 2005 International conference on Intelligent User Interfaces.

Edwards, L. D. (2005). Metaphors and gesture in fraction talk. In: The 4th Congress of the European Society for Research in Mathematics Education; Spain, Sant Feliu de Guixols.

Ekman, P. (1994). Strong evidences for universals in facial expressions: A reply to Russel's mistaken critique. *Psychological Bulletin*, 115(1), 268-286.

Ekman, P. (1997). Should we call it expression or communication?. *Innovations in Social Science Research*, 10, 4, 333-344.

Ekman, P., & Friesen, W.V. (1978). *Facial action coding system*. Consulting Psychologists Press.

Falangan, J., Huang, T., Jones P., & Kasif, S. (1997). Final report of the NSF Workshop on Human-Centers Systems: Information, Interactivity, and Intelligence. Washington D.C.: NSF.

Fan, C., Johnson, M., Messom, C. & Sarrafzadeh, A. (2003). Machine vision for an Intelligent Tutor. Proceedings of the International Conference on Computational Intelligence, Robotics and Autonomous Systems, Singapore.

Fasel, B., & Luettin, J. (2003). Automatic facial expression analysis: A survey. *Pattern Recognition*, 36(1) 259-275.

Goldin-Meadow S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Belknap.

Goldin-Meadow, S, & Singer, M. A. (2003). From children's hands to adult's ears: Gesture's role in the learning process. *Development Psychology*, 39(3), 509-520.

Golman S. D. M. (1996). Teaching a smarter learner. *Journal of Computer and System Sciences*, 52, 255-267.

Hager, J. C. & Ekman, P. (1995). Essential behavioral science of the face and gesture that computer scientists need to know. International workshop on automatic face and gesture recognition; Zurich.

Hayhoe M. (2000). Vision visual routines: A functional account of vision. *Visual Cognition*, 7, 43-64.

Huang T. S., and Chen L. S., & Tao H. (1998). Bimodal emotion recognition by man and machine, ATR Workshop on Virtual Communication Environments, Japan.

Johnson, W. L., Rickel, J. W. & Lester, J. C. (2000). Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education*, 11, 47-78.

Kort, B., Reilly, R. and Picard, R. W (2001) An affective model of interplay between emotions and learning: Reengineering educational pedagogy - building a learning companion. Proceedings of IEEE International Conference on Advanced Learning Technologies, 43-48.

Kwon O. N., Ju M.K., Park J. S., Cho K. H. & Ewha K. H.. (2003). Gesture in the context of mathematical argumentation. Proceedings of the 27th conference of the International group for the Psychology of Mathematics Education held jointly with the 25th Conference of PME-NA; Honolulu, USA.

Litman, D. J. & Forbes-Riley, K. (2006). Recognizing student emotions and attitudes on the basis of utterances in spoken tutoring dialogues with both human and computer tutors. Speech Communication, in press.

McNeill, D. (1992). *Hand and mind*. Chicago, The University of Chicago Press.

Mehrabian, A. (1972). *Nonverbal communication,* Chicago: Aldine-Atherson.

Mehrabian A. (1981). Silent messages: Implicit communication of emotions and attitudes. Belmont, CA: Wadsworth.

D'Mello, S. K., Craig, S. D., Gholson, B., Franklin, S., Picard, R. W., & Graesser, A. C. (2005), Integrating affect sensors in an intelligent tutoring system. In Affective Interactions: The Computer in the Affective Loop Workshop at 2005 International conference on Intelligent User Interfaces (pp. 7-13) New York: AMC Press

Montepare, J., Koff, E., Ziatchik, D., & Albert, M. (1999). The use of body movements and gestures as cues to emotions in younger and older adults. *Journal of Nonverbal Behavior*, 23(2), 133-152.

Mitrovic, A., Martin, B., & Mayo, M. (2002). Using evaluation to shape ITS design: Results and experiences with SQL-Tutor. *Int. J. User Modeling and User-Adapted Interaction* 12 (2-3), 243-279.

van Mulken, S., André, E. and Muller, J. (1998). The persona effect: How substantial is it?. Proceedings of Human Computer Interaction, Berlin.

New Zealand Ministry of Education (2003). Book 1. The Number Framework, Numeracy Professional Development Projects, Ministry of Education, Wellington.

Pantic, M., Rothkrantz, L. J. M. (1999). An expert system for multiple emotional classification of facial expressions. IEEE International Conference on Tools with Artificial Intelligence, Chicago, 113-120.

Pantic, M., & Rothkrantz, L. J. M.(2003). Toward an affect-sensitive multimodal human-computer interaction. Proceedings of the IEEE, 91(9), 1370-1390

Person, N. K., Graesser, A. C. & The Tutoring Research Group (2003). Fourteen facts about human tutoring: Food for thought for ITS developers. AI-ED 2003 Workshop Proceedings on Tutorial Dialogue Systems: With a View Toward the Classroom, Sydney, Australia.

Picard, R.W. (1997). *Affective computing*. MIT Press, Cambridge, Mass.

Prendinger, H. & Ishizuka, M. (Eds) (2004). *Life-like characters, tools, affective functions, and applications*. Cognitive Technologies Series, Springer, Berlin Heidelberg.

Reeves, B. and Nass, C. I. (1996). The Media equation: how people treat computers, television and new media like real people and places. Cambridge University Press.

Sarrafzadeh, A. (2002). Representing domain knowledge structure in Intelligent Tutoring Systems, Proceedings of the International Conference on Information and Communication Technologies in Education, Spain, November 02, 665-9.

Sarrafzadeh, A., Fan, C., Dadgostar, F., Alexander, S. T. V. & Messom, C. (2004) Frown gives game away: Affect sensitive tutoring systems for elementary mathematics. Proceedings of the IEEE Conference on Systems, Man and Cybernetics, The Hague.

Shergil, G., Sarrafzadeh, A., Diegel, O. & Shekar, A. (2004). Computerized Sales Assistants: The application of computer technology to measure consumer interest: A conceptual framework, Working Paper Series, Massey University, Issue 4.25, September 2004, ISSN 1174-5320

Strauss, M., Reynolds, C., Hughes, S. Park, K., McDarby, G., & Picard R. W. (2005). The HandWave Bluetooth Skin Conductance Sensor. Proceedings of Affective Computing and Intelligent Interaction, Beijing, China.

The Urban Teacher Collaborative. (2000). The urban teacher challenge: Teacher demand and supply in the Great City Schools. A study by The Urban Teacher Collaborative: Council of the Great City Schools, Recruiting New Teachers, Inc. and Council of the Great City Colleges of Education. Washington, DC: Council of the Great City Schools. [Online]

de Vicente, A. (2003). Towards tutoring systems that detect students' motivation: an investigation, Ph.D. Thesis, School of Informatics, University of Edinburgh.

Vapnik, V.N. (2000). *The nature of statistical learning theory*. Springer-Verlag New York, Inc., New York,NY, USA.

Please cite this article in press as: Sarrafzadeh, A. et al., ''How do you know that I don't understand?''
..., Computers in Human Behavior (2007), doi:10.1016/j.chb.2007.07.008