Easy with Eve: A Functional Affective Tutoring System

Samuel Alexander[1], Abdolhossein Sarrafzadeh[1], Stephen Hill[2]

[1] Institute of Information and Mathematical Sciences, Massey University,
Private Bag 102 904, NSMC, Auckland, New Zealand
{S.T.Alexander, H.A.Sarrafzadeh}@massey.ac.nz
[2] School of Psychology, Massey University, Private Bag 11 222, Palmerston North,
New Zealand
S.R.Hill@massey.ac.nz

Abstract. This paper presents research leading to the development of *Easy with Eve*, an Affective Tutoring System (ATS) in the domain of primary school mathematics. The system adapts to students via a lifelike animated agent called Eve, who is able to detect student emotion through facial expression analysis, and can display emotion herself. 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.

Keywords: Affective tutoring systems, lifelike agents, facial expressions, human tutors

1 Introduction

Many researchers now feel strongly that Intelligent Tutoring Systems (ITSs) would be significantly enhanced if computers could adapt according to the emotions of students [15], [10], [1]. 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 in the same ways that effective human tutors do [17], [19]. It seems that the term "Affective Tutoring System" was first used only several years ago [3], [19], 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* [15]. However, so far as the author is currently aware, no ATS has yet been implemented, although several groups are working towards this goal [10], [1], [11], [7].

During its brief history, the field of ATSs has faced two main barriers: reliably detecting the affective state of students; and knowing how best to adapt to this information once a student's emotions have been detected. The first of these issues seems to have generated by far the most attention, with growing numbers of researchers investigating various forms of facial expression analysis and gesture analysis, voice analysis, wearable computers, and predictive emotion models. In contrast, the second of these issues seems to have suffered serious neglect. However, as the technical obstacles to detecting emotions are being gradually overcome, the

G. Rebolledo-Mendez, E. Martinez-Miron (Eds): Workshop on Motivational and Affective Issues in ITS. 8th International Conference on ITS 2006, pp 38-45, 2006.

pressing relevance of *how* to adapt to student emotion is becoming increasingly obvious.

Therefore, the aim of this research is to investigate *how* to adapt the affective state of students, to develop a method for adapting to the affective state of students, and then to implement this method in a fully functional ATS. The first of these aims has been addressed by an observational study of how human tutors adapt their tutoring based on the affective state of students. This has led to the development of a case-based reasoning method for adapting to student affect, where recommended tutoring actions are generated based upon the ways in which human tutors adapt their tutoring to a student's state. This in turn has been applied to what is perhaps the first ever ATS, *Easy with Eve*; *Eve* is an animated lifelike agent that is able to both express and react to a range of affective states.

This paper will present the case-based method that was developed based on the results of the observational study, the implementation of *Easy with Eve*, and avenues for future work. The next section provides a background to ATSs, and a brief summary of the observational study of human tutors.

2 Background

2.1 Animated Pedagogical Agents

Though few if any existing ITSs can recognise emotions, many ITSs have been developed that can show emotions through an animated pedagogical agent [9], [16]. Animated pedagogical agents are "lifelike autonomous characters that cohabit learning environments with students to create rich, face-to-face learning interactions" [9]. 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 [12]. The persona effect has been shown to increase learner motivation, especially in technical domains, although its overall benefits remain unclear [12].

2.2 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. [17]), gesture analysis (e.g. [6]), voice analysis (e.g. [11]) and wearable computing (e.g. [18]); one example of a predictive emotion model is given by [5].

The current research uses automated facial expression analysis to detect the affective state of the student, chiefly because a facial expression analysis system has been developed in-house at Massey University, Auckland [17].

2.3 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 [12].

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 an existing addition exercise developed by the New Zealand Numeracy Project [13].

Video Analysis. To analyse the videos, a coding scheme was developed expanding on previous work by Person and Graesser [14]. 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; this was the raw data of the study.

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 totaled over two thirds of all tutor turns.
- 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.

- All other expressions, including confusion and apprehension, were very rare by comparison.

These results and their implications are presented in much greater detail in [2]. A more extensive analysis of the observational study, and possible improvements such as gathering data from more tutors and students, and studying human tutor-student interactions in a range of domains that vary in 'emotiveness', will be discussed in detail in a future paper.

3 A Case-based Method for Adapting to Student Affect

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.

3.1 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 [2]. 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.

3.2 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.

4 *Easy with Eve*: 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.

4.1 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. Fig. 1 gives a small sample of the appearance and capabilities of Eve.

Fig. 1. Three examples of Eve in action (from left to right): showing no expression, smiling, and speaking while pointing.

4.2 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 [17], 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 [8]. 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.

4.2 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 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.

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.

5 Conclusion and Future Work

A functioning ATS, *Easy with Eve*, has been implemented at Massey University, and is possibly the first ever 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. 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.

The next step will be to evaluate the effectiveness of the ATS; the system will be tested 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*.

References

1. Alexander, S.T.V., Sarrafzadeh, A.: Interfaces that Adapt Like Humans. Proceedings of Asia-Pacific Computer-Human Interaction 2004, Rotorua, New Zealand (2004)
2. Alexander, S.T.V., Hill, S., Sarrafzadeh, A.: How do Human Tutors Adapt to Affective State?. Proceedings of User Modeling, Edinburgh, Scotland (2005)
3. Alexander, S.T.V., Sarrafzadeh, A., Fan, C.: Pay Attention! The Computer is Watching: Affective Tutoring Systems. Proceedings of E-Learn, Phoenix, Arizona (2003)
4. Conati, C.: Probabilistic assessment of user's emotions in educational games. Applied Artificial Intelligence, Vol. 16. (2002) 555-575.
5. Conati, C., Maclaren, H.: Data-driven Refinement of a Probabilistic Model of User Affect. Proceedings of User Modeling, Edinburgh, Scotland (2005)
6. Dadgostar, F., Ryu, H., Sarrafzadeh, A., Overmyer, S.P.: Making sense of student use of nonverbal cues for intelligent tutoring systems. Proceedings of CHISIG, Canberra (2005)
7. D'Mello, S.K., Craig, S.D., Gholson, B., Franklin, S., Picard, R., Graesser, A.C.: Integrating affect sensors in an intelligent tutoring system. Proceedings of the International Conference on Intelligent User Interfaces (2005)
8. Ekman, P.: Should We Call it Expression or Communication? Innovations in Social Science Research. Vol. 10. (1997) 333-344
9. Johnson, W.L., Rickel, J.W., Lester, J.C.: Animated pedagogical agents: Face-to-face interaction in interactive learning environments. International Journal of Artificial Intelligence in Education, Vol. 11. (2000) 47-78
10. Kort, B., Reilly, R., Picard, R.W.: 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 (2001) 43-48
11. Litman, D.J., Forbes-Riley, K.: Recognizing Student Emotions and Attitudes on the Basis of Utterances in Spoken Tutoring Dialogues with both Human and Computer Tutors. Speech Communication, in press (2006)
12. van Mulken, S., André, E., Muller, J.: The persona effect: How substantial is it? Proceedings of Human Computer Interaction, Berlin (1998)
13. New Zealand Ministry of Education: Book 1. The Number Framework. Numeracy Professional Development Projects, Ministry of Education, Wellington (2003)
14. Person, N.K., Graesser, A.C., The Tutoring Research Group: Fourteen facts about human tutoring: Food for thought for ITS developers. AI-ED Workshop Proceedings on Tutorial Dialogue Systems: With a View Toward the Classroom, Sydney, Australia (2003)
15. Picard, R.W.: Affective Computing. MIT Press, Cambridge, Mass. (1997)
16. Prendinger, H., Ishizuka, M. (eds): Life-Like Characters, Tools, Affective Functions, and Applications. Cognitive Technologies Series, Springer, Berlin Heidelberg (2004)
17. Sarrafzadeh, A., Fan, C., Dadgostar, F., Alexander, S.T.V., Messom, C.: Frown gives game away: Affect sensitive tutoring systems for elementary mathematics. Proceedings of the IEEE Conference on Systems, Man and Cybernetics, The Hague (2004)
18. Strauss, M., Reynolds, C., Hughes, S. Park, K., McDarby, G., Picard R.W.: The HandWave Bluetooth Skin Conductance Sensor. Proceedings of Affective Computing and Intelligent Interaction, Beijing, China (2005
19. de Vicente, A.: Towards Tutoring Systems that Detect Students' Motivation: An Investigation. Ph.D. Thesis, School of Informatics, University of Edinburgh (2003)