Evaluating Learner Experience from Affective Cues

Shazia Afzal

Computer Laboratory, University of Cambridge,
15 JJ Thomson Avenue, Cambridge, CB3 0FD, UK
Shazia.Afzal@cl.cam.ac.uk

Abstract. Learning with computers is typically self-paced. Appraisal of the learner's experience is therefore crucial for making machine-learner interactions 'truly intelligent' and instinctive. As a significant aspect of Affective Learning, affect recognition needs to be considered in light of its implications and appliance to computer based learning environments. Drawing on this motivation, I propose to investigate the potential of using facial expression analysis to model affect in learning as a means for evaluating learner state.

Keywords: Affective Learning, Intelligent Tutoring, Facial Expression Analysis.

1 Motivation

The growing emphasis on affect sensitive human-computer interaction [1] finds an obvious application in computer based learning environments [2, 3]. Despite the recognition of affect as a vital component in the learning process, computer based learning has mostly concentrated on modelling the behaviour of a learner congruent to a particular instructional strategy [2, 4]. While researchers seek to emulate the effectiveness of expert human tutors in the design and functioning of learning technologies, efforts to make this interaction more natural motivates the ambitious field of Affective Learning. It reflects the growing understanding of the centrality of emotion in the teaching-learning process and explores ways to design emotionally aware computer tutors.

Effective tutoring by humans is an interactive and guided process where learner engagement is constantly monitored [5]. Formative assessment and feedback is therefore a crucial aspect of effectively designed learning environments, and should occur continuously and unobtrusively, as part of the instruction [6]. Naturalistic settings facilitate a constant evaluation of learner experience by virtue of both verbal and non-verbal communication channels. A significant challenge in equipping computer tutors with such a mentoring capability is giving them skills of affect perception. However, the absence of a theory or model of affect in learning, at a level of detail that is amenable to implementation, magnifies the overarching difficulty of affective inference. A synergistic approach, to study affect in learning with computers, in conjunction with suitable affect recognition technologies becomes therefore necessary. The objective is to reasonably emulate the social dynamics of human teacher-learner interactions in models that capture the essence of effective learning strategies like one-one tutoring [7, 8].

1.1 Outline

This paper outlines the framework of a proposed study to understand affect in learning and to work towards development of a real-time, unobtrusive sensing technology that can be deployed in an average computing environment.

The relevant background is briefly introduced in Section 2 with a review of some recent works on affect sensing in learning environments. Section 3 builds on this to propose a research methodology and outlines objectives that could be helpful in advancing the field of Affective Learning. Throughout this paper the terms emotion, affect and mental states will be used interchangeably.

2 Background

Given the diverse perspectives on conceptualising the term *emotion* itself, there is an inherent complexity in any study that investigates affective user-interfaces. However, a context driven application like computer based learning gives the advantage of modelling aspects that are characteristic to such an environment [3]. Affective Learning has been approached as a 3 stage problem – exploring underlying theories and models of affect in learning and deriving some sort of emotion subset or taxonomy to start with, developing sensing techniques that reliably identify these, and finally, utilizing this information effectively in context of learner's knowledge state for appropriate feedback or diagnosis [9, 10, 11]. Though each is a substantial research area in itself, the nature of the problem relies on insights from each of these.

2.1 Learning, Emotions and Motivation

There is no dissent on the importance of emotions in learning or their role in providing a basis for healthy cognitive functioning, but learning theories have generally positioned the affective domain as occupying a separate realm [12, 13]. Led by empirical studies and supporting evidence from neuroscience, there has been a surge of interest amongst educational psychologists and technologists alike for studying affective phenomena in learning. For a selected discussion on previous works see Craig et al [9], Kort et al [10] and, Afzal and Robinson [14].

The motivational quality and engagement value of machine-learner interactions significantly affects the success or failure of a learning experience. But motivation, like emotions can neither be directly observed nor assessed. Emotional state is an important component of motivation and has been used to infer the motivational state of learners [15, 16]. In all practical sense affective state influences motivation towards a task and inferring affective state may prove to be a vital indicator of a learner's motivation level.

2.2 Review of Affective Learning Technologies

Despite the prospects, there are relatively few studies on affect sensing in learning environments. These can be broadly categorised into two groups: methods that predict emotions based on an understanding of their causes, and those that detect emotions based upon their physical effects [17]. The first method is based on reasoning from direct input behaviour like state knowledge, self-reports, navigation patterns or outcomes to actions. The second approach relies on the understanding that non-verbal behaviour through bodily gestures, facial expressions, voice, etc, is instinctively more resourceful and aims to infer affective cues with the aid of sensors. Hybrid approaches as in Conati [18] and Zhou & Conati [19] are also promising as they impart contextual significance to affective cues and can aid in a better interpretation of affect state.

The Affective Computing Group at MIT is involved in a series of projects towards the building of a *Learning Companion.* Kapoor et al [20] use a novel method of self-labelling to classify data automatically, observed through a combination of sensors, into 'pre-frustration' or 'not-pre-frustration'. In a related work, Kapoor and Picard [21] and Kapoor et al [22] use multi-sensor classification to detect interest in children solving a puzzle by utilising information from the face, posture and current task of the subjects. Posture patterns are also used by D'Mello and Graesser [23] along with dialogue, to discriminate between affect states during interaction with an intelligent tutoring system. Sarrafzadeh et al [24] employ a fuzzy approach to analyse facial expressions for detecting a combination of states like happiness/success, surprise/happiness, sadness/disappointment, confusion and frustration/anger. In another study, de Vincente and Pain [15] use concrete aspects of learner interactions like mouse movements, quality of performance, etc to formalize inference rules for diagnosis of motivation. Litman and Forbes [25] propose a method of affect modelling from acoustic and prosodic elements of student speech.

The field is still in a formative stage and current technologies need to be fine tuned and validated for reliability outside controlled experimental conditions [26, 27].

3 Proposed Work

The growing understanding and the current state of art in Affective Learning substantiates research in this domain. The broad objective of this research is to examine how affect sensing can give insights into learner experience and how this can be exploited to make learning with computers more immersive and engaging.

It is known that experienced human teachers make use of "students facial expressions, body language, intonation, and other paralinguistic cues" for detecting learner state [28]. My aim is to explore methods that can reasonably approximate this human perceptual skill by restricting the bandwidth of communication to what is practically available in an average computer based learning environment. In order to model the dynamics of machine-learner interaction as closely and naturally as possible, I hope to base my work on detailed observational analysis of empirically collected data. The following subsections outline the steps in which this research is being undertaken.

3.1 Conceptualising Emotion

The definitional debate on the rather fuzzy concept of emotion poses a fundamental problem for distinguishing affect states in learning. In absence of a standard theory or model of affect in learning, the choice of conceptualising emotion is principally ad-hoc. A common approach is recourse to folk concepts of emotions derived from natural languages [29].

Such a lexical taxonomy underlies the Mind Reading DVD [30] which is an interactive computer-based guide to emotions. Based on a taxonomic classification by Baron-Cohen [31] it groups 412 emotion concepts into 24 distinct emotion groups. Each group encompasses the finer shades of that emotion concept and therefore gives the flexibility of choosing the right level of semantic distinction. The DVD in itself is a rich corpus of labelled video and can serve as a powerful tool for preliminary analysis. Using this taxonomy, I have selected a set of affect categories that are representative of some important affective states linked to learning [14]. These are – Afraid, Angry, Bored, Happy, Interested and Unsure. Figure 1 illustrates some of the emotion concepts they encompass.

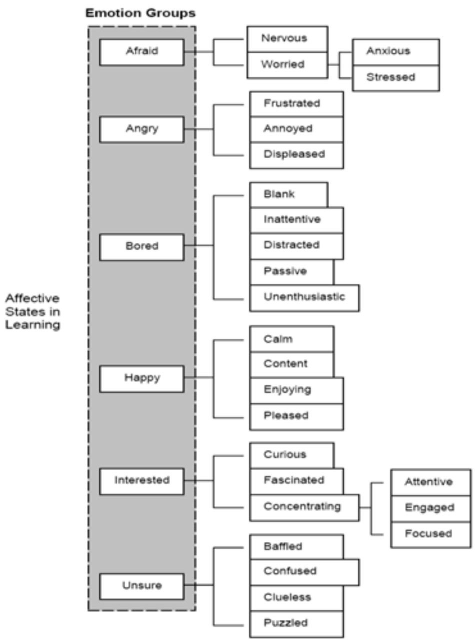

Fig. 1. Selected view of the Emotion Groups selected for this study and the emotion concepts they encompass.

3.2 Choice of Modality – Detection and Suitability in situational context

Lack of a consistent mapping between observable aspects of learner behaviour and actual affective states, technical feasibility, and practical issues complicate the choice of modality for sensing. The design, use and deployment of appropriate sensing technologies is further constrained by issues of ethics, privacy and comfort. The use of tactile sensing in particular is challenging. Though relatively easy to detect physiological sensing has its own shortcomings [26]. Speech analysis may not always be suitable as not all learning environments are dialogue based.

Given the pre-eminence of facial signs in human communication [3, 32, 33, 34] the face is a natural choice for inferring affective states. Facial information can be detected and analysed unobtrusively and automatically in real-time requiring no specialised equipment except a simple video camera. However, facial expressions are not simple read-outs of mental states and their interpretation being context-driven is largely situational. Computer tutors can exploit this aspect to infer affective states from observed facial expressions using the knowledge state and navigation patterns from the learning situation as a guide.

3.3 Empirical Data Collection and Analysis

Though current techniques of facial affect analysis show promising recognition rates, context-sensitive human affect analysis still remains a significant challenge [33]. A major impediment is the lack of a standardised database for testing and evaluation [33, 35]. Though a number of face databases exist, these are mostly posed or recorded in scripted situations that may not be entirely relevant in a learning situation. Therefore, in addition to looking at existing corpora [30, 35], it is essential to observe the nature of affect displays as they occur naturally in learning environments, for a more meaningful interpretation.

For this purpose I conducted a preliminary experiment to record a subject's face while doing an interactive map-based tutorial [36] and a card sorting puzzle [14]. Interesting patterns of facial expressions could be observed from this data. Using a split-screen view of similarly recorded facial behaviour and screen capture along with transcript analysis, I hope to get a rich test-bed of data for labelling and analysis. In addition to subjective descriptions, self-report from subjects will be obtained using a dimensional response format [37]. This self-labelling will be compared to labelling on the same measure by human experts, ideally teachers. Figure 2 shows the annotation tool devised for this study.

3.4 Computational Modelling Approach

Perception and interpretation of facial expressions is inherently interdisciplinary and thus, computationally challenging. Adopting a modular perspective by separating measurement of facial behaviour from interpretation of affect states can significantly reduce the complexity [34]. This separation can also facilitate subjective evaluations of data to be abstracted into measurable features for implementation.

The mental state inference tool developed by El Kaliouby [38] supports such a modular approach. This tool has been validated to perform real-time inference of complex mental states from head and facial displays in video streams. It combines bottom-up vision-based processing of face and head displays with top-down predictions of mental state models using a multi-level probabilistic graphical model. While data from experiments is collected, labelled and annotated, some groundwork using the available corpus like the Mind Reading DVD [30] is being done with the aid of this tool. This automated system is being trained using videos from [30] to examine how it generalises to the selected affect categories (see Section 3.1). Eventually, data collected from experiments will be processed using the trained system and based on inference results the framework will be validated. Figure 2 below shows the facial affect analysis by this tool for a video labelled as *Puzzled*.

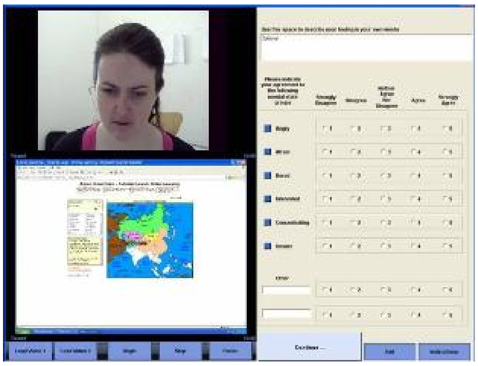

Fig. 2. Screenshot of the annotation tool showing a subject's facial expressions while doing an online tutorial [36].

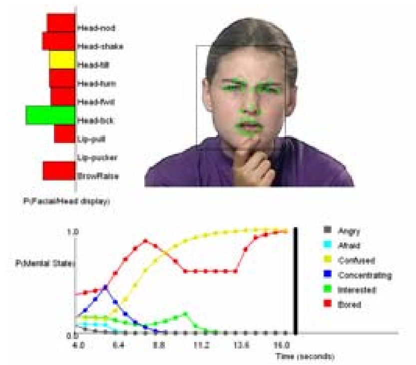

Fig. 3. Facial affect analysis [38] for a video labelled *Puzzled* (emotion group *Unsure*) of the Mind Reading DVD [30].

4. Conclusions

Learning is associated with a strong affective element that impacts overall performance, memory, attention, decision-making and attitude [2, 3]. To optimally balance this quality the first step is to reliably detect learner state. This research intends to explore how facial expression analysis can help in interpreting affective cues that may arise in a computer based learning environment. Data analysis using judgement and component approaches [34, 37] shall form the basis of this study.

Acknowledgments. I am grateful to my supervisor, Peter Robinson, for his guidance and support, and to the Gates Cambridge Trust for generously funding my research.

References

1. Picard, R.W.: Affective Computing. The MIT Press, Cambridge; 1997
2. Picard R.W., Papert S., Bender W., Blumberg B., Breazeal C., Cavallo D., Machover T., Resnick M., Roy D. & Strohecker C.: Affective Learning-a manifesto. BT Technology Journal, 2004, 22, 253-269
3. Lisetti C. & Schiano D.: Facial expression recognition: Where Human Computer Interaction, Artificial Intelligence and Cognitive Science Intersect, Pragmatics and Cognition, 8(1): 185-235, 2000
4. du Boulay B., Luckin R.: Modelling Human Teaching Tactics and Strategies for Tutoring Systems, International Journal of Artificial Intelligence in Education, 2001, 12, 235-256
5. Merrill D.C., Reiser B.J., Ranney M. & Trafton J.G.: Effective Tutoring Techniques: A Comparison of Human Tutors and Intelligent Tutoring Systems. Journal of the Learning Sciences, 1992, 2, 277-305
6. Bransford J.D., Brown A.L. & Cocking R.R.: How People Learn: Brain, Mind, Experience and School. National Academy Press, Washington, DC; 1999
7. Bloom B.S.: The 2 sigma problem: The search for method of group instruction as effective as one-to-one tutoring. Educational Researcher, 1984, 13, 3-16
8. van Vuuren S.: Technologies that power pedagogical agents and visions for the future. To appear in Special Issue of Educational Technology, 2006
9. Craig S.D., Graesser A.C., Sullins J. & Gholson B.: Affect and learning: an exploratory look into the role of affect in learning with AutoTutor. Journal of Educational Media, 2004, 29, 241-250
10. Kort B., Reilly R. & Picard R.W.: An affective model of interplay between emotions and learning: reengineering educational pedagogy-building a learning companion. In Proceedings of International Conference on Advanced Learning Technologies (ICALT), 2001, Madison Wisconsin
11. D’Mello S.K., Craig S.D., Gholson B., Franklin S., Picard R.W. & Graesser A.C.: Integrating Affect Sensors in an Intelligent Tutoring System. In Proceedings of the Workshop on Affective Interactions: The Computer in the Affective Loop Workshop, International Conference on Intelligent User Interfaces, 2005, 7-13
12. Schutz P.A., Lanehart S.L.: Introduction: Emotions in Education. Educational Psychologist, 2002, 37(2), 67-68
13. O’Regan K.: Emotion and e-learning. Journal of Asynchronous Learning Networks, 2003, 7(3), 78-92
14. Afzal S., Robinson P.: A Study of Affect in Intelligent Tutoring, In Proceedings of the Workshop on Modelling and Scaffolding Affective Experiences to Impact Learning, International Conference on Artificial Intelligence in Education, 2007, Los Angeles
15. de Vicente A., Pain H.: Motivation Diagnosis in Intelligent Tutoring Systems. In Goettl B.P., Halff C., Redfield C.L. & Shute V.J. (Eds.): In Proceedings of the Fourth International Conference on Intelligent Tutoring Systems, 1998, 86-95, Texas
16. Elliott C., Lester J.C. & Rickel J.: Integrating Affective Computing into Animated Tutoring Agents. In Proceedings of the Workshop on Animated Interface Agents: Making Them Intelligent, IJCAI, 1997, 113-121
17. Alexander S.T.V., Hill S. & Sarrafzadeh A.: How do Human Tutors Adapt to Affective State? Proceedings of User Modelling, 2005, Edinburgh, Scotland
18. Conati C.: Probabilistic assessment of user’s emotions in educational games. Journal of Applied Artificial Intelligence, 2002, 16, 555-575
19. Xhou X., Conati C.: Modelling students’ emotions from cognitive appraisal in educational games. Intelligent Tutoring Systems, 2002, 944-954
20. Kapoor A., Burleson W. & Picard R.W.: Automatic Prediction of Frustration, Int. J. Human-Computer Studies, 2007, 65(8), 724-736, doi:10.101016/j.jhcs.2007.02.003

21. Kapoor A., Picard R.W.: Multimodal Affect Recognition in Learning Environments**.** Proceedings of the 13th Annual ACM International Conference on Multimedia, 2005, Singapore
22. Kapoor A., Picard R.W. & Ivanov Y.: Probabilistic combination of multiple modalities to detect interest. Proceedings of International Conference on Pattern Recognition, 2004, Cambridge, England
23. D'Mello S., Graesser A.: Mind and Body: Dialogue and Posture for Affect Detection in Learning Environments. Proceedings of International Conference on Artificial Intelligence in Education, 2007, Los Angeles
24. Sarrafzadeh A., Fan C., Dadgostar F., Alexander S.T.V. & Messom C.: Frown gives game away: Affect sensitive tutoring systems for elementary mathematics. In Proceedings of the IEEE Conference on Systems, Man and Cybernetics, 2004, The Hague
25. Litman D., and Forbes, K.: Recognising emotions from student speech in tutoring dialogues. In Proceedings of the IEEE Automatic Speech Recognition and Understanding Workshop (ASRU), 2003
26. Ward R.D., Marsden P.H.: Affective computing: problems, reactions and intentions, Interacting with Computers, 2004, 16, 707-713
27. Batliner A., Fischer K., Huber R., Spilker J., Noth E.: How to Find Trouble in Communication, Speech Communication, 2003, 40(1-2), 117-143
28. Lepper M.R., Woolverton M., Mumme D.L., & Gurtner J.: Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In Lajoie S.P. & Derry S.J. (Eds.): Computers as Cognitive Tools, 75–105. Lawrence Erlbaum, Hillsdale, New Jersey; 1993
29. Scherer K.R.: Trends and developments: research on emotions, Social Science Information, 2005
30. Baron-Cohen, S., Golan O., Wheelwright S., and Hill J.: Mind Reading: The Interactive Guide to Emotions. Jessica Kingsley Publishers, London, 2004
31. Baron-Cohen, S., Golan O., Wheelwright S., and Hill J.: A New Taxonomy of Human Emotions (under review), 2004
32. Carroll J.M., Russell J.A.: Do Facial Expressions Signal Specific Emotions? Judging Emotion from the Face in Context. Journal of Personality and Social Psychology, 1996, 70, 205-218
33. Pantic M.: Toward an Affect-Sensitive Multimodal Human-Computer Interaction, Proceedings of the IEEE, 2003, 91, 1370-1390
34. Ekman, P.: Emotion in the Human Face. Cambridge Univ. Press, Cambridge; 1982
35. Cowie, R. et al.: Beyond emotion archetypes: Databases for emotion modelling using neural networks, Neural Networks, 2005, 18, 371-388
36. Sheppard Software–educational games: http://www.sheppardsoftware.com/Geography.php
37. Scherer K., Ekman P.: Handbook of Methods in Nonverbal Behaviour Research. Cambridge Univ. Press, Cambridge, UK; 1982
38. El Kaliouby R.: Mind-Reading Machines: Automated Inference of Complex Mental States. PhD Dissertation, Computer Laboratory, University of Cambridge, UK; 2005