

Emotion Recognition for Intelligent Tutoring

Sintija Petrovica¹ and Hazım Kemal Ekenel²

¹ Faculty of Computer Science and Information Technology, Riga Technical University,
Riga, Latvia

sintija.petrovica@rtu.lv

² Faculty of Computer and Informatics, Istanbul Technical University, Istanbul, Turkey
ekenel@itu.edu.tr

Abstract. Individual teaching has been considered as the most successful educational form since ancient times. This form still continues its existence nowadays within intelligent systems intended to provide adapted tutoring for each student. Although, recent research has shown that emotions can affect student's learning, adaptation skills of tutoring systems are still imperfect due to weak emotional intelligence. To enhance ongoing research related to the improvement of the tutoring adaptation based on both student's knowledge and emotional state, the paper presents an analysis of emotion recognition methods used in recent developments. Study reveals that sensor-lite approach can serve as a solution to problems related to emotion identification accuracy. To provide ground-truth data for emotional state, we have explored and implemented a self-assessment method.

Keywords: Intelligent tutoring systems, Affective computing, Emotion recognition, Self-Assessment Manikin.

1 Introduction

With the progress in the affective computing field and studies carried out in education and psychology revealing a close relationship between emotions and human learning, a new generation of intelligent tutoring systems (ITSs) has appeared – affective tutoring systems (ATs). Since teachers can evaluate students' emotional states with a rather high reliability on the basis of facial expressions, body language, and speech to make changes in the teaching process, similarly, tutoring systems should be capable of assessing students' emotions and use this information to promote their learning and achieve better learning outcomes [1]. However, there still exists a shortage for these systems regarding adaptation skills possessed by human-teachers, particularly, the lack of an emotional intelligence [2]. This is mostly due to the inability of ITSs to accurately and unobtrusively classify emotions during learning process [3].

The paper provides a review of existing ATs summarizing their development purposes, working principles, and architectural differences, as well as applied emotion recognition methods in these systems are analyzed both from the developers' perspective (regarding the implementation difficulty) and their caused inconveniences to students. To provide ground truth for automatic emotion identification, a self-assessment method via Self-Assessment Manikin is designed and implemented.

2 Emotional Intelligence and Intelligent Tutoring Systems

ITSs are a generation of computer systems, which aim to support and improve teaching and learning process in certain knowledge domains. ITSs simulate a human-tutor and provide benefits of one-on-one tutoring. Such systems allow providing more natural learning process by adapting a learning environment (content, feedback, navigation, etc.) to the characteristics of a particular student. Adaptation is possible because of integrated knowledge into traditional architecture, which includes a student diagnosis module collecting and processing information about the student, a pedagogical module responsible for implementation of the teaching process, a problem domain module able to generate and solve problems in the problem domain, and an interface module managing interaction among the system and the student through different input/output devices [1].

Research in the field of ITSs in recent years has gradually shifted its emphasis from cognitive processes to emotionally-cognitive processes [4]. Around a decade ago, ideas from affective computing field [5] came also in the development of tutoring systems creating so called affective tutoring systems. These changes can mostly be explained by increasing attention paid to the relationship between emotions and learning [4]. Research results show that emotions are a significant factor in the learning process and even can affect student's motivation and abilities to learn. Various studies demonstrate that students experience a wide diversity of positive and negative emotions during the learning process, e.g., anxiety, enjoyment, hope, pride, surprise, satisfaction, anger, boredom, frustration, confusion, and shame [3,6], therefore, more attention should be given to these emotions in the ITSs' development process. Summarizing the available information about ATS and their development purposes, e.g., [7,8], we define ATS as *an intelligent tutoring system that imitates human-teacher and his/her adaptation abilities not only to student's knowledge but also to emotional state to intervene (react accordingly) only in those situations, when emotional state can become a threat to student's willingness to engage in learning process leaving negative impact on knowledge acquisition and learning outcomes.*

To support ATS functioning, the ITS architecture must be accompanied by additional components. Commonly, three additional components are incorporated into the ITS architecture to form so called affective behavior model that allows providing appropriate responses considering both student's knowledge and emotions [8,9]. The first component usually is responsible for the automatic identification of the student's emotional state [10]. Emotion recognition is carried out by detecting and analyzing different features, e.g. facial expressions, body motion and gestures, speech, physiological characteristics, etc., and applying various classifiers to identify student's emotions [9,10,11]. The emotion response module or affective (behavior) pedagogical model is often distinguished as the second component [8,12]. This component provides reasoning on the current tutoring situation and allows for further adaption of the tutoring process based not only on the student's current knowledge level and learning characteristics but also on the student's emotional state [9]. By analyzing architecture variations of different existing ATSs, an emotion expression module can be found as a third component. This module can be referred as an extension of interface module that allows ATS to express its own emotions via virtual tutor or pedagogical agent (PA) with its own mood and emotions [13].

3 Related work

Regarding emotion identification, various aspects are examined, e.g., sensors used for the acquisition of data related to emotions, methods used for emotion classification and the most commonly modeled emotions. For this study, different ATSS are selected to cover various taught problem domains – both from “*hard*” sciences, for example, mathematics, physics, natural sciences, computer science, and “*soft*” sciences or humanities, e.g. study of languages. However, it must be noted that, in general, the majority of intelligent tutoring systems are developed for well-defined problem domains, since more rules exist regarding task generation and solving, whereas development of ITSs for ill-defined problem domains still remains a challenge. For this research, existing ATSS are analyzed (see Table 1), e.g., *MathSpring* [14], *Prime Climb* [15], *Easy with Eve* [16], *FERMAT* [17], *Cognitive Tutor Algebra* [14,18] and *PAT2Math* [19] intended for teaching mathematics, *ITSPOKE* [20] and *AutoTutor* [6] for physics, *CRYSTAL ISLAND* [21] and *GuruTutor* [22] for biology, *Inq-ITS* [23] for natural sciences, *INES* [24] and *MetaTutor* [25] for teaching medicine and *VALERIE* [26] for French language.

Determination of the student's emotional state is implemented by analyzing various data sources providing features that can give information about student's emotional state. Ideally, a quantitative and continuous measurement of emotional experience is required in an objective and unobtrusive manner, e.g. analysis of interactional content [3]. Two most commonly used feature categories for emotion identification are:

1. *facial features* –mostly patterns of facial appearance are extracted and analyzed, as well as eye movement is tracked and gaze patterns are acquired indicating regions of interest to which student is paying attention;
2. *features acquired from log files* –features recorded in log files and related mainly to student's interaction with the system. Acquired features include both information linked to the student, e.g., behavior patterns, action history, activity level, and data characterizing his/her current tutoring situation, e.g., task history.

Besides these two most common feature categories, other characteristics are also acquired for emotion classification, e.g. body language, physiological signals (for example, skin conductance, heart rate, and muscle movement), speech features (intensity, volume, duration, etc.) and usage of input devices, such as a mouse.

To perceive these features, various sensors are used. Cooper et al. [27] have grouped these sensors in three possible categories considering the level of discomfort they cause to the student:

1. *physiological sensors* –these sensors cause the greatest discomfort because they require a contact with certain parts of the body (e.g., skin conductivity sensor, heart rate sensor, electromyograph, etc.).
2. *touch or haptic sensors* –these sensors (e.g., pressure-sensitive mouse or chair) induce less discomfort and students very often do not notice them, however, the usage of such sensors for emotion recognition requires a student to touch them thus limiting his/her movement freedom;
3. *observational sensors* – these sensors (e.g., video cameras, eye trackers, and microphones) are not physically intrusive, however, they can distract a student's attention and make him/her feel uncomfortable, knowing that all actions are recorded.

Table 1. Emotion recognition in affective tutoring systems.

ATS	Sensors	Emotional data detection and emotion classification
<i>AutoTutor</i>	Video camera Pressure sensitive chair	Posture and eye pattern extraction, analysis of log files. Classifiers: Naïve Bayes, neural networks, logistic regression, nearest neighbor, C4.5 decision trees.
<i>Cognitive Tutor Algebra</i>	Not used	Analysis of log files recording features related to the student's behavior, event and activity history in the learning process. Classifiers: J48 decision trees, K* algorithm, step regression, JRip, Naïve Bayes, REP-Trees.
<i>CRYSTAL ISLAND</i>	Not used	Analysis of surveys, interviews and log files. Emotions are modeled using a Dynamic Bayesian Network.
<i>Easy with Eve</i>	Video camera	Facial feature extraction. Classifier: support vector machines.
<i>FERMAT</i>	Video camera	Extraction of facial feature points and regions of interest. Classifiers: neural network, a fuzzy expert system.
<i>GURU Gaze Tutor</i>	Eye tracker Video camera	Eye tracking and gaze pattern extraction, analysis of log files. Analysis of the attention time paid to the screen.
<i>Inq-ITS</i>	Not used	Analysis of log files. Classifiers: J48 decision trees, step regression, JRip.
<i>INES</i>	Not used	Analysis of the student's activity level, difficulty of the task, previous progress, number of errors, severity of the error. Emotions are predicted by appraisal rules.
<i>ITSPOKE</i>	Microphone	Extraction of acoustic-prosodic, lexical features (speech intensity, energy, volume, duration, and pauses) and dialogue features (e.g. the accuracy of the answer). Semantic analysis is used for the assessment of answer accuracy and linear regression for confidence evaluation.
<i>MathSpring</i>	Not used	Analysis of log files, self-assessment reports, behavior patterns, etc. Classifier: linear regression.
<i>MetaTutor</i>	Eye tracker	Extraction of gaze data features and features related to areas of interest within system's interface. Classifiers: random forests, Naïve Bayes, logistic regression, and support vector machines.
<i>PAT2Math</i>	Video camera	Analysis of log files and extraction of facial feature points. Emotions are identified based on Facial Action Coding System and psychological model of emotions (OCC model).

<i>PRIME</i> <i>CLIMB</i>	Various physiological sensors	Determination of skin conductivity, heart rate, muscle activity, and analysis of log files. Biometrical data is analyzed via unsupervised clustering.
<i>VALERIE</i>	Video camera Microphone Mouse Phys. sensors	Determination of skin conductivity, heart rate, extraction of facial and speech features, analysis of mouse movement. Classifiers: nearest neighbor, discriminant function analysis, Marquardt Back-propagation algorithm.

Besides sensor usage, emotion identification in some ATs is based on results of students' filled surveys or self-assessment reports, where students report their own feelings, emotions, or mood in a particular situation. This can be considered as "accurate" method for the emotion acquisition, if students are aware of their emotions, however, a possibility exists that students will consider such surveys as redundant and not provide correct information about their emotions.

Considering the most commonly modeled student's emotions, it must be noted that part (although minor) of existing tutoring systems (e.g. *Easy with Eve*, *FERMAT*, *VALERIE*) carry out facial expression recognition to identify so called basic emotions (anger, disgust, fear, happiness, sadness, and surprise) that mostly are not characteristic for the learning process. However, it is only a small part of ATs and in overall, emotion modeling trends are improving and developers mainly focus on emotions that are felt during the learning and directly influence the learning process. Therefore, most of analyzed ATs (e.g. *AutoTutor*, *Cognitive Tutor Algebra I*, *Crystal Island*, *Inq-ITS*, *MathSpring*, and *WaLLis*) are aimed at learning specific emotions and are able to determine, whether the student is, for example, concentrated (interested/in flow state), confused, bored, frustrated, anxious, ashamed, etc.

4 Affect through Self-Assessment

Currently, a research direction regarding the emotion recognition is the analysis of log files recording interaction between students and system, thus using so called *sensor-free* approach [18,23]. Mainly, a new ATs development or existing ATs modification using this trend can be explained by the limited availability of sensors in real learning conditions [14]. Since sensor-free approaches do not provide very high emotion recognition accuracy and can crucially decrease accuracy of the tutoring process adaptation, one of the possible solutions to overcome this problem is so called *sensor-lite* approach, which requires a (minimal) use of available sensors, e.g., (built-in) video cameras or microphones [28]. To achieve the emotion recognition as accurate as possible the first step is the collection of a "ground-truth" emotion data set that can be later used for training and comparing results of automatic measurement of affect [29]. Regarding this issue, one of the most popular self-assessment methods is analyzed – "*Self-Assessment Manikin*" (SAM), which can be used independently from the sensor-based approaches. This type of self-assessment allows getting students' feelings using graphic representation of the three fundamental emotion dimensions, which include **P**leasure, **A**rousal and **D**ominance (PAD) [30]. After carrying out the self-assessment, it is possible to represent all three emotion dimensions in the PAD emotion space,

where each graphic depiction can have its own value in the range of $[-1..1]$. By combining values of all three PAD values, classification of emotions can be done. Complete list of emotions and their PAD values is available in [31].

In this research, it was decided that an initial step for emotion recognition is the implementation of SAM, which will be used as an independent method for the acquisition of emotional data to identify students' emotions, while they are learning and go through various instructional activities (e.g., e.g. starting new topic, solving tasks, receiving feedback, etc.) within the ATS. This collected data can serve as ground-truth for sensor-based emotion classification studies.

One of the existing SAM implementations is *AffectButton* tool, which is freely available and can be customized and used in other research projects to acquire emotional data from systems' users [32]. The *AffectButton* is a measurement instrument that enables a user to give detailed emotion feedback about his/her feelings, mood, and attitudes towards different objects. After clicking the button, three values are generated corresponding to all PAD values. Currently, the source code of *AffectButton* tool is already adapted and integrated in the environment for research requirements. However, since this method provides only PAD values characterizing specific emotions but not "specific" emotions, discrete emotion calculation based on acquired PAD values is implemented as well (see Fig. 1.). For this purpose, Equation (1) is applied to determine the distance " d " between acquired PAD values for two emotions e_j and e_i . The idea of emotion calculation is borrowed from [33], where in a similar way the student's mood is calculated. The less is the distance value, the more similar emotions are. In total, 15 different emotions are incorporated for comparison but only the closest five based on their PAD values are shown.

$$d(e_i, e_j) = \sqrt{(e_{P_i} - e_{P_j})^2 + (e_{A_i} - e_{A_j})^2 + (e_{D_i} - e_{D_j})^2} \quad (1)$$

In general, this emotion self-assessment can be ensured during the whole learning process allowing students to report about their emotional changes when they prefer to do this or when the tutoring system itself prompts them to provide emotion self-assessment during performing particular learning activities.

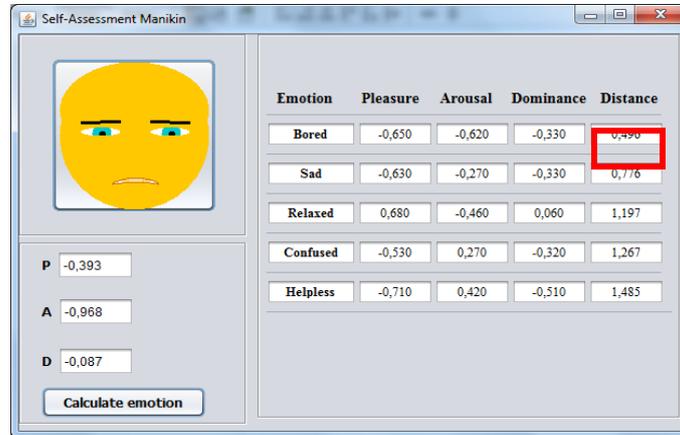


Fig. 1. *AffectButton* and emotion calculation based on generated PAD values

Despite the possible inconveniences, which this method can cause to students (because of extra interventions), it will allow identifying emotions during the learning process. The next step of research will be related to the implementation of automatic identification of emotions and comparing the system with the collected ground-truths. This would contribute to the achievement of a higher goal of ongoing research [1] – improved tutoring adaptation skills based on student's knowledge and emotional state.

5 Conclusions and Future Work

Affective tutoring systems and their functioning principles are studied in this paper. A more detailed analysis of adopted emotion recognition methods is carried out covering sensors used, features acquired and methods applied for feature classification and emotion recognition. Two most common feature categories used for emotion recognition are facial features (e.g. shape of eyes, eyebrows, lips and gaze movement) acquired from video cameras and features extracted from log files that contains saved information about student's behavior during the interaction with ATS, as well as features related to tutoring situation itself.

To provide ground truth for automatic emotion identification, a self-assessment method via Self-Assessment Manikin is designed and implemented. Based on acquired PAD values, discrete emotion classes are calculated. However, more learning specific emotions should be added to the list.

Future work is to develop an automatic emotion identification approach, for example by observing facial appearance variations during learning process, in order to ensure automatic emotion determination without direct students' involvement.

Acknowledgments. This work was supported by the COST Action IC1303 Algorithms, Architectures and Platforms for Enhanced Living Environments Short-Term Scientific Mission grant, by TUBITAK project no. 113E067, and by a Marie Curie FP7 Integration Grant within the 7th EU Framework Programme.

References

1. Petrovica, S., Pudane, M.: Simulation of Affective Student-Tutor Interaction for Affective Tutoring Systems. *Int. J. Educ. Learning Syst.* 1, 99--108 (2016)
2. Thompson, N., McGill, T.J.: Affective Tutoring Systems: Enhancing e-Learning with the Emotional Awareness of a Human Tutor. *Int.J.Inf.Commun.Technol.Educ.* 8, 75--89 (2012)
3. Afzal, S., Robinson, P.: Modelling Affect in Learning Environments - Motivation and Methods. In: *Proceedings of the ICALT'2010*, pp. 438--442. IEEE Computer Society (2010)
4. Petrovica, S.: Tutoring Process in Emotionally Intelligent Tutoring Systems. *Int. J. Technol. Educ. Mark.* 4(1), 72--85 (2014)
5. Picard, R.W.: *Affective Computing*. MIT Press, Cambridge (1997)
6. D'Mello, S.K., Picard, R.W., Graesser, A.C.: Toward an affect-sensitive AutoTutor. *IEEE Intell. Syst.* 22(4), 53--61 (2007)
7. Landowska, A.: Affect-awareness Framework for Intelligent Tutoring Systems. In: *Proceedings of the HIS'2013*, pp. 540--547 (2013)
8. Kaklauskas, A., et al.: Affective tutoring system for built environment management. *Comp. Educ.* 82, 202--216 (2015)
9. Malekzadeh, M., Mustafa, M.B., Lahsasna, A.: A Review of Emotion Regulation in Intelligent Tutoring Systems. *Educ. Technol. Soc.* 18(4), 435--445 (2015)
10. Sarrafzadeh, A., Shanbehzadeh, J., Overmyer, S.: E-learning with Affective Tutoring Systems. In: *Intelligent Tutoring Systems in E-learning Environments*, pp. 129--140 (2011)
11. Rikure, T., Novickis, L.: Building a Learner Psychophysiological Model Based Adaptive e-Learning Systems: A General Framework and its Implementation. In: *Proceedings of ADBIS'2010, LNCS, vol.5968*, pp.31--38 (2010)
12. Hernández, Y., Sucar, E., Conati, C.: An Affective Behavior Model for Intelligent Tutors. In: *Proceedings of the ITS'2008, LNCS, vol. 5091*, pp. 819--821 (2008)
13. Gu, X., Wang, Z., Zhang, J., Wang, W., Zheng, S.: Design of emotional Intelligent Tutor System based on HMM. In: *Proceedings of the ICNC'2010*, pp. 1984--1988 (2010)
14. Wixon, M., Arroyo, I., Muldner, K., Burleson, W., Lozano, C., Woolf, B.: The Opportunities and Limitations of Scaling Up Sensor-Free Affect Detection. In: *Proceedings of the EDM'2014*, pp. 145--152 (2014)
15. Amershi, S., Conati, C., Maclaren, H.: Using feature selection and unsupervised clustering to identify affective expressions in educational games. In: *Proceedings of the ITS'2006, LNCS, vol. 4053*, pp. 21--28 (2006)
16. Sarrafzadeh, A., Alexander, S., Dadgostar, F., Fan, C., Bigdeli, A. How do you know that I don't understand? A look at the future of intelligent tutoring systems. *Comput. Hum. Behav.* 24, 1342--1363 (2008)
17. Zatarain-Cabada, R., Barrón-Estrada, M.L., Camacho, J.L.O., Reyes García, C.A.: Affective Tutoring System for Android Mobiles. In: *Proceedings of the ICIC'2014, LNCS, vol. 8589*, pp. 1--10 (2008)
18. Baker, R.S.J., Gowda, S.M., Wixon, M., Kalka, J., Wagner, A.Z., Salvi, A., Alevan, V., Kusbit, G., Ocumpaugh, J., Rossi, L.: Towards Sensor-free Affect Detection in Cognitive Tutor Algebra. In: *Proceedings of the 5th International Conference on Educational Data Mining*, pp. 126--133. International Educational Data Mining Society (2012)
19. Jaques, P.A., Seffrin, H., Rubi, G., de Morais, F., Ghilardi, C., Bittencourt, I.I., Isotani, S.: Rule-based expert systems to support step-by-step guidance in algebraic problem solving: The case of the tutor PAT2Math. *Expert Syst. with Appl.* 40(14), 5456--5465 (2013)
20. Litman, D., Forbes-Riley, K., Silliman, S.: Towards emotion prediction in spoken tutoring dialogues. In: *HLT/NAACL'2003*, pp. 52--54 (2003)
21. Sabourin, J.L., Rowe, J.P., Mott, B.W., Lester, J.C.: Considering Alternate Futures to Classify Off-Task Behavior as Emotion Self-Regulation: A Supervised Learning Approach. *J. Educ. Data Mining* 5(9), 9--38 (2013)

22. Olney, A., D'Mello, S., Person, N., Cade, W., Hays, P., Williams, C., Lehman, B., Graesser, A.: Guru: A Computer Tutor that Models Expert Human Tutors. In: Proceedings of the ITS'2012, LNCS, vol. 7315, pp. 256--261 (2012)
23. Paquette, L., Baker, R.S., Sao Pedro, M.A., Gobert, J.D., Rossi, L., Nakama, A., Kauffman-Rogoff, Z. Sensor-Free Affect Detection for a Simulation-Based Science Inquiry Learning Environment. In: Proceedings of the ITS'2014, LNCS, vol. 7315, pp. 1--10 (2014).
24. Heylen, D., Nijholt, A., Akker, H.J.: Affect in tutoring dialogues. *J. Appl. AI* 19(3-4), 287--310 (2005)
25. Jaques, N., Conati, C., Harley, J. M., and Azevedo, R.: Predicting Affect from Gaze Data during Interaction with an Intelligent Tutoring System. In: Proceedings of the ITS'2014, LNCS, vol. 7315, pp. 29--38 (2014)
26. Paleari, M., Lisetti, C., Lethonen, M.: VALERIE: Virtual Agent for Learning Environment Reacting and Interacting Emotionally. In: Proceedings of the AIED'2005 (2005)
27. Cooper, D., Arroyo, I., Woolf, B.P.: Actionable affective processing for automatic tutor interventions. *New Perspectives on Affect and Learning Technologies*, pp. 127--140 (2011)
28. D'Mello, S.K., Graesser, A.C.: Feeling, Thinking, and Computing with Affect-Aware Learning Technologies. In: *The Oxford Handbook of Affective Computing*, pp. 419--434 (2015)
29. Gunes, H., Nicolaou, M.A., Pantic, M.: Continuous Analysis of Affect from Voice and Face. In: *Computer Analysis of Human Behaviour*, pp. 255--291 (2011).
30. Bradley, M., Lang, P.: Measuring Emotion: The Self-Assessment Manikin and the Semantic Differential. *J. Behav. Ther. Exp. Psy.* 25(1), 49--59 (1994)
31. Russell, J.A., Mehrabian, A.: Evidence for a Three-Factor Theory of Emotions. *J. Res. Pers.* 11(3), 273--294 (1977)
32. Broekens, J., Brinkman, W.P.: AffectButton: a method for reliable and valid affective self-report. *Int. J. Hum.-Comput. St.* 71(6), 641--667 (2013)
33. Qui-rong, C.: Research on Intelligent Tutoring System Based on Affective Model. In: Proceedings of the MMIT'2010, pp.7--9 (2010)