Towards Enriching an ITS with Affective Support

Miguel Arevalillo-Herráez[1], Salvador Moreno-Picot[1], David Arnau[2], Paloma Moreno-Clari[1], Jesus G. Boticario[3], Olga C. Santos[3], Raúl Cabestrero[4], Pilar Quirós[4], Sergio Salmeron-Majadas[3], Ángeles Manjarrés-Riesco[3], Mar Saneiro[3]

[1] Department of Computer Science, University of Valencia, Spain
miguel.arevalillo@uv.es, salvador.moreno@uv.es, paloma.moreno@uv.es
[2] Department of Mathematics Education, University of Valencia, Spain
david.arnau@uv.es
[3] aDeNu Research Group. Artificial Intelligence Dept. Computer Science School. UNED, Spain
jgb@dia.uned.es, ocsantos@dia.uned.es, ssalmeron@bec.uned.es,
amanja@dia.uned.es, marsaneiro@dia.uned.es
[4] Department of Basic Psychology, Faculty of Psychology, UNED, Spain
rcabestrero@psi.uned.es, pquiros@psi.uned.es

Abstract. Recent progress in affective computing is having an important impact on the development of Intelligent Tutoring Systems (ITS). Many ITS use action logs to record user's interaction with the system, such as to discover important information that help instructional designers to improve the ITS performance. However, finer grain interaction data as well as emotional information gathered from external sources is required to determine affective or mental states that can be used to enrich learner's experience with affective support. In this paper, we discuss what is needed to design an experiment for capturing relevant information from an ITS to improve the learner's competence in solving algebraic word problems considering learners' emotional and mental states.

Keywords: Affective computing, ITS, Multimodal emotions detection.

1 Introduction

User's affective state features a strong relationship with the cognitive process [1-4]. In the MAMIPEC and MARES projects we aim at exploring potential applications of affective computing in the context of accessible and personalized learning systems. To this end, we consider a user context that includes a wide range of appliances and devices to enrich the user's interaction. To study possible ways to detect user's emotions in a learning context, a number of experiments focused on emotional data gathering have been carried out. A total of 92 subjects with different profiles and backgrounds, including people with functional diversity [5], were asked to solve a collection of mathematical exercises through dotLRN Learning Management System (LMS) while emotional information was gathered both from sensors and questionnaires.

In order to further understand the learning implications of affective states, identify possible applications of affect detection in tutoring systems, and reinforce some of the conclusions drawn from the above study, we are currently following two research directions: 1) investigating potential applications of affective computing to improve an ITS developed in the context of the MARES project [6, 7]; 2) extending the dotLRN open source LMS and related software modules to include the required adaptive affective support through affective educational oriented recommendations [8].

This paper describes some of the actions adopted by both research groups to improve the existing ITS and endow it with adaptive and affective support through recommendations. This ITS is deployed as a standalone application that provides tutoring features on a mathematical topic. In particular, the application aims at improving the learner's competence in solving algebraic word problems. The algebra domain has been chosen because of the many possibilities that it offers, regarding potential responses to specific mental states. Next, the ITS is described. After that, we discuss how to enrich the ITS with affective information based on the analysis of results carried out to date on the aforementioned experiments.

2 ITS description and position within the state of the art

The ITS emulates the behavior of a human tutor by tracking the current resolution path that the student is following, and adapts feedback accordingly. To this end, expert knowledge on the structure of word problems is codified by using hypergraphs that represent the relations between quantities in the different analytical readings associated with each problem [6, 7]. The system is able to provide feedback and hints on demand. In both cases, the most likely analytical reading is computed and used to adapt the system response, which is given in natural language.

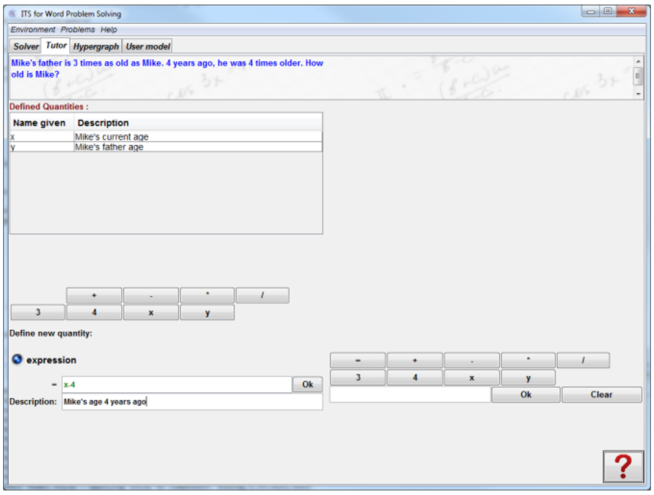

Fig. 1. A screenshot of the ITS in tutoring mode

Fig. 1 shows a screenshot of the system in tutoring mode. The panel on the left hand side is used to define quantities, either by using a letter or as a function of other quantities that have already been defined. In the figure, the student has already used letters x and y to designate the ages of Mike and his father, respectively; and is currently defining Mike's age 4 years ago as x-4. The panel on the right hand size is used to build equations that relate several existing quantities. To encourage a systematic problem solving approach, calculator-like components are used in both cases. These contain the basic operators and one button per quantity already defined. The component used to build equations includes an additional button for the equals sign. In this way, quantities need to be defined before they are used to either define another quantity or set an equation. The question mark button at the right-bottom corner of the screen is used to request a hint. If this button is pressed, a hint is displayed on a floating window. This window is also used to provide feedback to incorrect actions. In Fig.1, a sample help box is also shown on top of the main application window.

The ITS has been designed so that action logs are dynamically produced as the user interacts with the system. Student actions are written to a file in natural language. Fig. 2 shows an example of the output generated. In this file, it can be observed that after defining the two letters, the student requested a hint. Again, the student felt unable to carry out the recommended action and asked for further help. The system reacted by giving further details on the first action suggested. Still, the student did not know what to do and abandoned the application without finishing the resolution. Apart from other obvious uses of such visual information (e.g. files can be inspected to study the student's performance in detail), we are currently working on applying machine learning algorithms to the logs in order to draw relevant conclusions regarding situations that may demotivate the learner and cause abandonments.

```
NEW PROBLEM LOADED: Ages
    STATEMENT: Mike's father is 3 times as old as Mike. 4 years ago, he
was 4 times older. How old is Mike?
    USER ACTION: DEFINING LETTER.
      - x to represent Mike's current age
    SYSTEM ACTION: ACCEPTED
    USER ACTION: DEFINING LETTER.
      - y to represent Mike's father age
    SYSTEM ACTION: ACCEPTED
    SYSTEM ACTION: HINT GIVEN.
      4 years ago, Mike was four years younger than today
      You may try to define
        Mike's age 4 years ago
        as a function of
          - 4
          - Mike's current age (x)
    SYSTEM ACTION: HINT GIVEN.
      4 years ago, Mike was four years younger than today
      Hence
        Mike's age 4 years ago = Mike's current age less 4
        You may try to define
        Mike's age 4 years ago
        as x-4
    USER ACTION: EXITING WITHOUT FINISHING
```

Fig. 2. An example of the high-level log produced by the application

Despite the possibilities offered by the high level information in the logs, finer grain interaction data may have a relatively higher importance to determine affect or mental states. For example, inactivity times, mouse movements or the time elapsed between clicks when defining an expression may provide important indicators relevant for the learning process. Combined with other (ideally non-invasive) sources of information (webcams, eye tracking hardware), interaction data can be used to detect specific emotional situations such as concentration, boredom, confusion or frustration [9-11]. In turn, this information can be directly used by the ITS to adapt common responses and/or handed to a recommender system to act in consequence [2].

3 Issues to consider for emotions detection in the ITS

Currently, we are trying to take advantage of the ITS tracking capabilities to enrich the multimodal emotional data mining detection approach [12] by gathering more detailed interaction and emotional information from the ITS and further exploit its adaptive features. In a previous experience [5], which was planned by a multidisciplinary team that includes experts in different fields (mathematics education, psychology, programming, data mining, machine learning and modeling), participants had to solve multiple choice mathematical exercises. Affective states were elicited at predetermined moments during the experience and gathered through several sources as follows: i) physiological sensors (hearth rate, breath rate, temperature, galvanic skin response, blood pressure) in order to detect significant variations related to certain changes in learner's affective state, ii) video recording (web cams, Kinect device, eye tracker) to find characteristic emotional meaningful facial gestures and attention foci, iii) interaction records (from mouse, keyboard and desktop) to identify behavioural changes, iv) standardized questionnaires (e.g. Big Five Inventory, General Self-Efficacy Scale, Positive and Negative Affect Schedule) to take into account certain aspects of participant's personality and emotions and v) self-reports and scales (e.g. Self-Assessment Manikin) on their feelings and thoughts.

In order to elicit several affective states, three groups of questions were prepared. The first one was easy if paper and pencil could be used, but participants were not allowed to do so. The next group of questions was limited in time, allowing less time than needed in order to cause stress in the participants (they were told that time was sufficient enough to fulfill the task). In the third group of questions the difficulty level was lowered and the type of problem was changed to logical series in order that participants could finish the session with a sensation of joy and happiness.

Although the experimental design allowed for a coherent data capture, the platform (being an LMS) lacked of some relevant functionality that could have enriched the quality of the information gathered in order to dynamically adapt the system behavior according to the user's input. Moreover, dotLRN does not provide in-built support for capturing interaction data. Mouse clicks, keyboard strokes and other interaction data had to be captured by using independent software, hence requiring a careful post-processing step to ensure that all data were adequately synchronized with the additional sensors used, namely video recorded information and physiological sensors.

In order to plan the experimental design to be carried out with the ITS based on the experience of previous experiments, the following issues are being addressed:

- **Determination of the affective and mental states that are relevant from a learning perspective.** Boredom, interest and frustration are some of these relevant states. Here, the involvement of educators with experience in providing affective support through virtual learning scenarios is needed. The purpose here is to help these educators to identify relevant situations that require emotional support, and ultimately define the support to be provided. TORMES methodology, which adapts the ISO standard 9241-210, can be used to guide educators to elicit and describe affective recommendations with educational value in their scenarios [13].
- **Selection of the most adequate devices to capture appropriate data that can be used to infer the user's affective/mental state**. Although non-intrusive devices are preferred, other more intrusive are still of interest. Despite that they may not be directly applicable in current practical setups, research conclusions may highlight their importance and encourage the construction of non-intrusive devices to capture the same type of signal. Microsoft Kinect technology, eye tracking, webcams, physiological sensors were already used at previous experiments [5, 12].
- **Plan the data gathering process.** This includes deciding on the most relevant variables, the format used to record the data produced by the different devices, and the synchronization mechanism that will be used to be able to combine information coming from the multiple sources. The ITS runs as a standalone application, and may easily be extended to capture low level interaction data related to keystrokes and mouse clicks and movements, and/or modified to adapt the higher level information that is currently recorded to the objectives of the experience. However, other devices are not integrated into the ITS and may require further development to ease the subsequent analysis, as well as their synchronization.
- **Elicitation of affective/mental states.** During the interaction, affective and mental states need to be provoked. The analysis of this type of interventions is focused on inferring changes in the user's affective state from the reactions detected by the input devices. The ITS is currently able to assess the difficulty of each problem by examining the analytical reading associated with them; and it is also capable of providing adapted feedback by using parameterized templates. These two features can be exploited to devise specific instructional designs aimed at eliciting emotions rather than maximizing learning.

4 Major Challenges

Apart from the intrinsic difficulty associated with detecting metal states in a non-intrusive way, there are many other aspects that make the work on enriching the ITS with affective support specially challenging. Currently, the ITS does not incorporate a recommender system, which could issue appropriate responses when particular affective and mental states are detected. The integration of the many aspects involved in the construction of a recommender system based on the detection of mental states requires expertise in several application fields. For this reason, a multidisciplinary team with experience in different areas has been built. Group members coming from the psychology field have previously copped with the problems involved in the affective states detection. In particular, their experience will serve to collect and integrate the great variety of sources of information –cognitive, behavioral and physiological– present in the study of emotions. This will be specifically valuable to maximize the amount, accuracy and relevance of emotional information, along with the minimization of intrusiveness to yield more accurate information. The expertise of other members in artificial intelligence is needed to construct data models which are appropriate for the problem at hand, to design the inference that will support the system and to combine the multiple information sources which will be fed into the recommender. Psycho-educational expertise is another fundamental ingredient, mainly related to the identification of situations where recommendations may have a positive impact in learning. Some team members have extensive experience in this subject and have developed an entire methodology to support the elicitation of educational oriented recommendations (TORMES), which can be used to identify opportunities where affective based recommendations could be offered [13]. Emotional support in e-learning platforms is currently a widely addressed issue in order to take advantage of the role emotions play in learning and cognitive process [1]. Affective state detection is a necessary open issue widely addressed, where the use of many different data sources (drawing a multimodal approach) is being applied in order to get new and richer information about the learner [14]. Due to the huge amounts of data a multimodal approach can lead to, the use of data mining techniques to extract affective information from the data gathered surfaces.

5 Conclusions and Future Work

Results from first experience on emotional states detection is the basis for a new experiment that aims to incorporate affective support to an existing ITS in the algebra domain. Shortcomings identified in the first experience have been considered to be used in a more flexible standalone tutoring application (from the adaptation point of view) than an LMS. Experiments are being designed on the same platform as affective support will be provided (i.e. the ITS).

Once the new experiment is run, data mining will proceed in a similar way as in the previous experiments. Conclusions will be used to include affective information into the user's model, and to adapt the ITS to react to disruptive emotional and mental states. The effect of incorporating affective support will then be evaluated in a real environment, with the participation of students at secondary education. High level logs will also be analyzed to identify factors that may contribute to promoting positive and negative mental states. This information will be used by instructional designers to improve the ITS. Furthermore, it can also be used to define affective educational oriented recommendations that can deliver affective feedback provided that the required adaptive infrastructure has been developed. This project implies far more than simply detecting mental states, but the development of a module that uses information related to the user's mental states to improve learning is a challenging issue on itself.

Acknowledgments

Authors would like to thank the Spanish Ministry of Economy and Competitivity for funding the MAMIPEC and MARES projects (TIN2011-29221-C03-01 and TIN2011-29221-C03-02), which have supported this research work.

References

1. Porayska-Pomsta, K., Mavrikis, M., and Pain, H., "Diagnosing and acting on student affect: the tutor's perspective," *User Modeling and User-Adapted Interaction*, vol. 18, no. 1–2, pp. 125–173, Feb. 2008.
2. Shen, L., Wang, M., and Shen, R., "Affective e-Learning: Using 'Emotional' Data to Improve Learning in Pervasive Learning Environment," *Educational Technology & Society*, vol. 12, no. 2, pp. 176–189, 2009.
3. Blanchard, E. G. , Volfson, B., Hong, Y.-J. and Lajoie, S. P. ,"Affective Artificial Intelligence in Education: From Detection to Adaptation," in *Proceedings of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling*, Amsterdam, The Netherlands, The Netherlands, 2009, pp. 81–88.
4. Forgas, J. P. , Baumeister, R. F., and Tice, D. M. *Psychology of self-regulation: Cognitive, affective, and motivational processes.*, vol. 11. New York, NY US: Psychology Press, 2009.
5. Santos, O. C., Rodríguez-Ascaso, A., Boticario, J. G., Salmeron-Majadas, S., Quirós, P., Cabestrero, R. Challenges for inclusive affective detection in educational scenarios. In proceedings of HCI International 2013. Lecture Notes in Computer Science, in press (2013)
6. Arevalillo-Herráez, M., Arnau, D., Marco-Giménez, L., Domain-specific knowledge representation and inference engine for an intelligent tutoring system, Knowledge Based Systems (2013), in press.
7. Arnau, D., Arevalillo-Herráez, M., Puig, L., González-Calero, J.A., Fundamentals of the design and the operation of an intelligent tutoring system for the learning of the arithmetical and algebraic way of solving word problems. Computers & Education 63: 119-130 (2013)
8. Santos, O.C., Boticario, J.G., Manjarrés-Riesco, A. An approach for an Affective Educational Recommendation Model. Recommender Systems for Technology Enhanced Learning: Research Trends & Applications. Manouselis, N., Drachsler, H., Verbert, K., Santos, O.C. (eds.). Springer, 2013 (in press).
9. Elliott, C., Rickel, L., Lester, J.. 1999. Lifelike pedagogical agents and affective computing: an exploratory synthesis. In Artificial intelligence today, Michael J. Wooldridge and Manuela Veloso (Eds.). Lecture Notes In Computer Science, Vol. 1600. Springer-Verlag, Berlin, Heidelberg 195-212.
10. de Vicente, A., Pain, H., "Motivation Diagnosis in Intelligent Tutoring Systems" Intelligent Tutoring Systems. Lecture Notes in Computer Science Volume 1452, 1998, pp 86-95
11. D'Mello, S.; Picard, R.W.; Graesser, Arthur, "Toward an Affect-Sensitive AutoTutor," Intelligent Systems, IEEE , vol.22, no.4, pp.53,61, July-Aug. 2007

12. Santos, O.C., Salmeron-Majadas, S., Boticario, J.G. Emotions detection from math exercises by combining several data sources. In proceedings of the 16th International Conference on Artificial Intelligence in Education (AIED 2013), in press
13. Santos, O.C., Boticario, J.G. (2012) Affective Issues in Semantic Educational Recommender Systems. In Proceedings of the 2nd Workshop on Recommender Systems for Technology Enhanced Learning. CEUR-WS proceedings, vol. 896, p. 71-82.
14. D'Mello, S. K., & Kory, J. Consistent but Modest: A Meta-Analysis on Unimodal and Multimodal Affect Detection Accuracies from 30 Studies. In L. P. Morency et al (Eds)Proceedings of the 14th ACM International Conference on Multimodal Interaction (2012). (pp. 31-38). New York:ACM