Toward Empathetic Agents in Tutoring Systems

Jessica Faivre*, Roger Nkambou* and Claude Frasson**

*University of Quebec in Montreal, Department of Computer Science
mail box 8888, succursale Centre-ville, H3C 3P8 Montréal (Québec) Canada
**University of Montreal
{faivre.jessica@courrier.uqam.ca; roger.nkambou@uqam.ca; frasson@iro.umontreal.ca}

Abstract

This paper presents a way of improving computer-based with lifelike presence in learning environment. The approach combines Intelligent Tutoring Systems with research on human emotion in Cognitive Sciences, Psychology and Communication. Considering the relations between emotion, cognition and action in contextual learning, we propose an architecture of a multiagent-based instructional system in which two adaptive emotional agents have been integrated. One manifests the tutor's emotional expressions trough a 3D embodied agent, whereas the second is designed to elicit and analyse the learner's emotional experiences during the interactions with the system. We present here the system's architecture and its first implementation.

Keywords Intelligent Tutoring Systems, Human-Computer Interaction, Emotions, Multiagent System, Learner Modeling

1. Introduction

Our interest in the role of emotional agents integrated in Tutoring Systems is motivated by what is now common wisdom in social cognitive theory, i.e that learning takes place through a complex interplay involving both cognitive and affective dimensions (Pintrich and Schrauben 1992). Researchers in cognitive sciences argue that emotions enable people to communicate efficiently by monitoring and regulating social interaction, and by evaluating and modifying their emotional experiences (Damasio 1999).

Tutoring Systems are computer-based learning systems inspired by new methods of teaching and learning based on one-to-one interactions. To be classified as "intelligent", they must present "human-like" tutoring capabilities, that means that they have to be able to adjust the content and delivery of the lesson to the students' characteristics and needs by analysing and/or anti-

cipating theirs responses and behaviours. Given what we said earlier about the role of emotions, the human teachers usually monitor student emotions in order to take relevant decision regarding the interaction with the student. An ITS should be able to do the same.

In order to provide those social interactions and provide a more effective learning environment in computer-based system (Baylor 2001), we focus our attention on developing pedagogical emotional agents with capabilities of recognition student's emotions and of modelling and synthesis their own emotion through gesture expression. We propose an Intelligent Tutoring System (ITS) model based on a multiagent architecture in which two adaptive emotional agents have been integrated. One is designed to elicit and analyse the learner's emotional experiences through his interactions with the system, whereas the second manifests the tutor's emotional non-verbal expressions using a 3D embodied agent.

This paper is organized as follow. Section 2 briefly introduces the current internal structure of the multiagent tutoring system. The two emotional agents integrated in that architecture are presented in depth (section 3 and 4). A summary and prospects for future research are presented in section 5.

2. Multiagent Architecture of the ITS

Tutoring Systems traditionally include modules devoted to instructional session such as tutor, learner model and virtual laboratory. The open JAVA platform we designed (figure 1) integrates three artificial agents named Tutor Agent (TA), Tutor's Adaptive Emotional Agent (TAEA) (grouped in a Tutoring Subsystem) and Learner's Adaptive Emotional Agent (LAEA). These agents interact with flexibility and communicate in accordance with Fipa-ACL (FIPA 2000). The Learner Model (LM) keeps track of the learning path of the student in order to identify the student's learning style including both cognitive and emotional styles that will be stored respectively in Cognitive State (Cstate) and Emotional State (Estate) modules. The first one monitors the integrity and the coherence of the student's knowledge structure (Knowledge Management System) and the
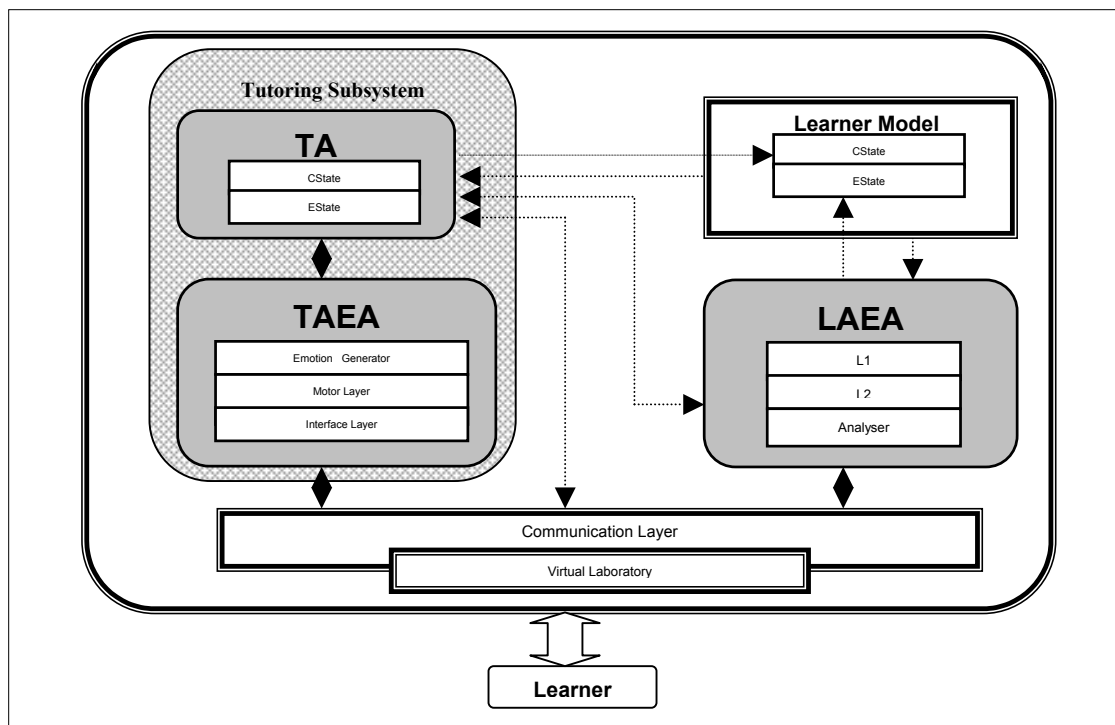
Figure 1. Architecture of the system

second consists of temporal indexation of emotion (Emotional Memory). The Virtual Laboratory is a micro-world, which contains definite primitives that permit the manipulation of environmental objects in a learning context where the students must perform tasks or solve problems using interactive virtual simulations. The Communication Layer permits learner's actions syntactic validation and to communicate them to the agents. For more details concerning the different modules inside the system see (Faivre, Nkambou, and Frasson 2002).

3. Tutor's Adaptive Emotional Agent

The Tutor Agent monitors the learning session by making decisions on the contents and the relevant resources to present. It has its own Cognitive State (tutorial strategies, plans, scenarios, pedagogical goals, knowledge), which is used to analyse the student's actions and results comparatively to its own desires or beliefs. It includes an Emotional State based on the OCC model (Ortony, Clore, and Collins 1988) that has been simplified to represent student's emotions at any given moment through a combination of 24 non-overlapping emotion types with assigned values. Events and data from other sources concerning the student (both cognitive and emotional) given by the LAEA and LM, are what triggers the TA's emotional variations

(experience). For example, a successful action event creates a Happy-For feeling depending on how many mistakes the learner made trying before. The variations carried by this event are also influenced by the degree of difficulty to accomplish this action compared to the ones previously encountered. The TA's plans give an overview of how difficult the coming action will be. These relations are expressed in the form of "if-then" rules implemented in an expert system using *Java Expert System Shell* (JESS). This simplifies the task of specifying how different factors should influence the tutor agent's emotional state with a low level of formalism. The Tutor Agent interacts with the learner via the Communication Layer and expresses itself physically via the Tutor's Adaptive Emotional Agent. Its exchanges with the LAEA and LM are useful to update the LM's cognitive state (information about the concepts mastered by the learner), and to adjust itself to the difficulty level associated with an event so it reflects more accurately the difficulty really encountered by the learner.

The TAEA focuses on the tutor's emotional expression displayed on computer screen and embodied in EMILIE, a 3-D agent (figure 2). The TAEA provides three layers, which are used to generate, represent and express tutor's body gestures, facial expression and eye gaze (Nkambou et al. 2002). The Emotion Generator is a set of relations that define how events, plans and records of past events that induce variations in the TA's EState, should influence variation to the TAEA's appearance. To ensure a good bodily concordance of expression with the homologous
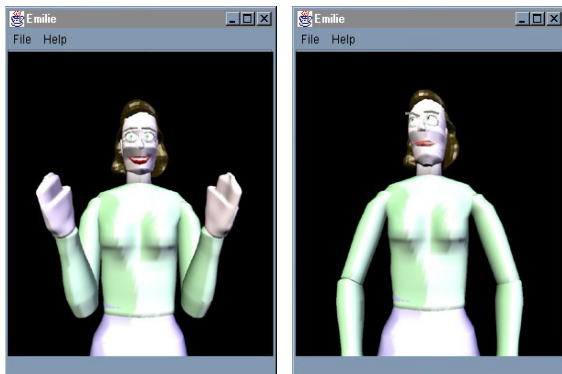
Figure 2. Examples of emotional expressions

internal data, there should be a specification of symbolic gesture associated to semantic representations. Thus, the Motor layer is a set of relations that define how emotions expressed in the Emotion Generator are translated in a representation in the agent's interface. This choice allows the elaboration of emotional and behavioural knowledge that we structure to build a collection of propositions characterizing the different behaviours in a human being with facial expressions and gestures physiologically linked. The behavioural expression is inferred directly from emotional representation; for example, a direct relation is established between the emotion of Joy and smiling and inversely with Distress. The Interface Layer is a definition of the agent's appearance, free of geometrical considerations, used to produce a visual output of the agent.

In order to split TA's emotional states into TAEA's behavioural units, a set of meaningful finite situations was considered, onto which some rules for reasoning were applied to build up a collection of propositions characterizing various system behaviours. To simulate emotional responses to learner's manipulations in a virtual laboratory, EMILIE was designed using a process similar to qualitative reasoning, which permits to formalize its modeling process relying on representation of continuous aspects of the world such as space, time and quantity, while enabling reasoning based on a small quantity of information (Forbus 1996). Such a model is appropriate to assign relations between different emotions, values, and properties outside of the emotional model (for example, relation between a feeling of joy and the amount of smiling) and between two factors without knowing exactly how much and following what function the first factor influences the second. When the tutor's emotional state changes, the agent's visual representation can be affected in two ways. Firstly, its facial expression is inferred directly from its representation of emotions with

relations specified between different emotions and facial characteristics. Secondly, changes in the EState module can influence the visual appearance by initiating gestures when changes exceed a certain specified threshold (the fire small pre-recorded movements for example, a sudden raise in the Sorry-For emotion triggers a gesture of the agent shaking its head bent forward looking down). A lower variation would have initiated a gesture of the agent briefly smiling down, lowering the outer eyebrows and slightly closing the eyes. These gestures amplify the tutor's emotional feedback, increase the power of representation and make transitions between different emotional states that are more obvious.

4. Learner's Adaptive Emotional Agent

The principal goal of this agent is to detect, analyse the learner emotional state during an instructional interaction and to adapt himself by becoming increasingly specific to the student. Is the user satisfied, more confused, frustrated, or simply sleepy? It acts like a "behavioural planner" by adapting its own behavioural rules according 1) to current learner's "emotional actions" transmitted by video image and interpreted by the Analyser (student's face analysis), 2) to information stocked in LM and 3) to learner's performance delivered by the Tutor Agent. Three layers constitute this agent. The deep layer (L1) contains general rules of hypothetic behavioural actions accompanying emotional experiences (emotional actions) induced by any specific stimulation and directed towards emotional regulation/adaptation according to valid norms and rules expected by the pedagogical goals. The second layer (L2) (empty at session's beginning) contains new rules corresponding to L1's old rules revised, adapted to the learner's current emotional state and stored as contextual-specific rules to identify schematic and conceptual processing and further to predict emotional actions. The Analyser is dedicated to the examination of student's emotional behaviours and computes the value of the difference (Δe) between what is "expected", predicted by the layer L1 and what is really "observed" and obtained from the learner. When Δe is relevant (i.e., $\Delta e \neq 0$), the Analyser transmits the information to correct the initial rules, to reduce the value of Δe. If $\Delta e = 0$, L2 keeps one empty set.

Emotion-recognition decisions can be modelled using collections of production rules that specify classes of external situations that turn on particular emotions (Blascovich and Mendes 2000). However, it is very difficult to exactly know which attributes are relevant when differentiating between many emotional states, because in natural interaction, prototypic expressions of basic emotions occur relatively infrequently and facial cues unconsciously perceived make it difficult to elicit emotion detection knowledge. In order to add those

capabilities of automatic facial expression recognition to the LAEA, we are designing a neuro-fuzzy system providing semantic interpretation outputs and parameters intensity of the expressiveness (very low, medium, very high) corresponding to five relevant emotions experienced in learning context (satisfaction, confidence, surprise, confusion, and frustration). Artificial neural networks traditionally employed have learning, non-linear classification, and generalization abilities and fuzzy logic provides a natural framework for the creation of emotion diagnosis rules with linguistic variables dealing with uncertainty and imprecision. Thus a hybrid neural network with fuzzy inference system relies heavily on high-speed number-crunching computation to find rules or regularity in data sets and present close resemblance to human-like decision-making dealing with flexible information in real-life ambiguous situations (Castro et al. 2002).

In order to initialize emotional action rules, LAEA also addresses queries to the learner after the session about his own estimation of his emotional states via a self-evaluation scale and a visual support (figure 3) (harmonization between what the system detected and what the learner means). This debriefing window proposes the key snapshots of the activity windows and their real-time corresponding learner's face expression. This recall test requires learner's self-inference in particular moments: before, during and immediately after decision-making or responses during the session and also after tutor's feedbacks. Even if it could seem subjective, this explicit diagnosis approach can serve as an outcome variable (how various experiences affect the way the student feels about himself) and as a mediating variable (self-esteem needs are presumed to motivate a wide variety of psychological processes) (Brown, Dutton, and Cook 2001). For that initialization, LAEA can also refer to a third expert party (teacher, educator, psychologist) willing and able to evaluate and infer emotional states by observation (De Vicente and Pain 2002). LEAE can dynamically change it's own data structures to achieve maximum efficiency and capacity to adapt itself to the particular student's profile during execution time.

Figure 3. Debriefing window

5. Conclusion

Intelligent Tutoring System must enrich human-machine interactions with more personalized communications and interrogative methods in order to provide flexible tutoring process, context-sensitive help and explanation in a current learning session.We have proposed a multiagent system with emotional agents that, we hope, will contribute to raise the student's productivity by involving him in a constructive interaction that reveals aspects of his learning states (both cognitive and emotional). Engagement on the meta-cognitive level by the self-assessment (debriefing) achieves significantly better results and should also promote generative and reflective learning (White, Shimoda, and Frederiksen 1999).

In order to improve the learning context, empathic pedagogical agents endowed with affective anticipation and planning capacity could be able to optimize the learner behaviour inducing a particular mood state to him (emotional contagion) (Golman 1995), or at least a positive feeling. An inappropriate behaviour would lead to unpleasant learner's emotions, indicating deviance or inappropriate actions, whereas the tutor's enthusiasm would induce student's enthusiasm. New agent functionalities will be integrated in our system. The TAEA should be able to converse with the user and give him positive or negative self-relevant feedback in a natural, fluent prosodic context and his visual interface will benefit of a new 3D model more cartoon-like. By adding and/or removing some functionalities and/or agents to the platform, the distributed architecture we propose will permit experimental studies to evaluate and to ascertain student perceptions of the usefulness of AEA-based learning, their personal comfort and confidence or enjoyment (Massaro et al. 2000).

Further works will consist in experiencing the current system with real university courses. This will permit not only to evaluate the impact of each type of emotion in learning process, but also to select only those emotions that can positively influence (improve) the learning quality.

References

Bailor, A. 2001. Permutations of Control: Cognitive Gui-delines for Agent-Based Learning Environments. *Journal of Interactive Learning research* 12(4):403-425.

Blascovich, J.; and Mendes, W.B. 2000. Challenge and threat appraisals: the role of affective cues. In Forgas, J. ed. *Feeling and thinking: The role of affect in social cognition*. New York: Cambridge University Press 201-222.

Brown, J.D.; Dutton, K.A.; and Cook, K.E. 2001. *From the top-down: Self-esteem and self-evaluation.* Cognition and Emotion *15(5):615-631.*

Castro, J. L.; Mantas, C. J.; and Benitez, J. M. 2002. Interpretation of Artificial Neural Networks by the means of fuzzy rules. *IEE Transactions on Neural Networks* 13(1): 101-116.

Damasio, A. 1999. *The Feelings of What Happens. Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace & Compagny.

De Vicente, A.; and Pain, H. 2002. Informing the Detection of the Students' Motivational State: An Empirical Study. In *Proceedings of Intelligent Tutoring Systems*, Cerri, S.A., Gouardères, G. and Paraguaçu, F. eds. Springer-Verlag Berlin Heidelberg, 933-943.

Faivre, J.; Nkambou, R.; and Frasson, C. 2002. Adaptive Emotional Agents in Learning Context. In *Proceedings on Artificial Intelligence and Soft Computing* , Banff, Alberta. Acta Press.

FIPA 2002. FIPA. FIPA Communicative Act Library Specification.

Forbus, K.D. 1996. *Qualitative reasoning*. Chapter for the CRC Handbook of Computer Science. Qualitative Reasoning Group.

Golman, D. 1995. *Emotional intelligence*. New York: Bantam Books.

Massaro, D.W.; Cohen, M.M.; Beskow, J.; and Cole, R.A. 2000. Developing and evaluating conversational agents. In Cassell, J., Sullivan, J., Prevost, S., and Churchill, E. eds. *Embodied Conversational Agents*. Cambridge: MIT Press 287-318.

Nkambou, R. ; Laporte, Y. ; Yatchou, R. ; and Gouardères, G. 2002. Embodied Emotional Agent in Intelligent Training Systems. In Abraham, A., Jain, L. and Kacprzyk, J. eds. *Recent Advances in Intelligent Paradigms and Applications*, chap. 11. Springer Verlag, Berlin, 235-255.

Ortony, A.; Clore, G.L.; and Collins, A. 1988. *Cognitive Structure of Emotions*. Cambridge University Press.

Pintrich, P. R.; and Schrauben, B. 1992. Students' motivational beliefs and their cognitive engagement in classroom tasks. In Schunk, D., and Meece, J. eds. *Student perceptions in the classroom: Causes and consequences*. Hillsdale, NJ: Erlbaum, 149-183.

White, B. Y.; Shimoda, T. A.; and Frederiksen, J. R. 1999. Enabling Students to Construct Theories of Collaborative Inquiry and Reflective Learning: Computer Support for metacognitive Development. *International Journal on Artificial Intelligence in Education* 10:151-182.